Hardware requirements for realizing a quantum advantage with deterministic single-photon sources

Patrik I. Sund,[1] Ravitej Uppu,[1,2] Stefano Paesani,[1,3] and Peter Lodahl[1,*]

[1]*Center for Hybrid Quantum Networks (Hy-Q), Niels Bohr Institute University of Copenhagen, Blegdamsvej 17, DK-2100 Copenhagen, Denmark*

[2]*Department of Physics & Astronomy, University of Iowa, Iowa City, Iowa 52242, USA*

[3]*NNF Quantum Computing Programme, Niels Bohr Institute, University of Copenhagen, Blegdamsvej 17, Copenhagen 2100, Denmark*

(Received 16 October 2023; accepted 12 March 2024; published 8 April 2024)

Boson sampling is a specialized algorithm native to the quantum photonic platform developed for near-term demonstrations of quantum advantage over classical computers. While clear useful applications for such near-term pre-fault-tolerance devices are not currently known, reaching a quantum advantage regime serves as a useful benchmark for the hardware. Here, we analyze and detail hardware requirements needed to reach quantum advantage with deterministic quantum emitters, a promising platform for photonic quantum computing. We elucidate key steps that can be taken in experiments to overcome practical constraints and establish quantitative hardware-level requirements. We find that quantum advantage is within reach using quantum emitters with an efficiency of 60%–70% and interferometers constructed according to a hybrid-mode-encoding architecture, constituted of Mach-Zehnder interferometers with an insertion loss of 0.0035 dB (a transmittance of 99.92%) per component.

DOI: 10.1103/PhysRevA.109.042613

I. INTRODUCTION

Devices based on quantum systems can potentially outperform the capabilities of classical computers [1,2]. Quantum technologies are rapidly progressing towards this goal and new computational regimes are being explored [3–5]. While fault tolerance is generally thought to be necessary to reach most practical applications, reaching this regime necessitates hardware requirements that are far from current capabilities, limiting demonstrations to small-scale experiments [6,7]. Quantum advantage (QA), where specialized algorithms can demonstrate speedups over classical computers, has been identified as an intermediate milestone computational regime amenable for near-term hardware using readily available quantum hardware components [3,8,9]. While it is currently not known if any practical applications are possible in this regime, it serves as an entry point to beyond-classical capabilities and an important benchmark for developing scalable platforms that can evolve towards fault tolerance. In this context, we analyze the hardware requirements for achieving QA using photonic quantum hardware, where fusion-based approaches for fault-tolerant quantum computing have been proposed [10,11].

Aaronson and Arkhipov proposed boson sampling of photonic quantum states [8] as a route for demonstrating QA with near-term quantum hardware. The key insight is the connection between the correlations induced by linear interference operations on N indistinguishable photons in an M-mode linear optical interferometer and the matrix permanent, a quantity that is #P-hard to compute on a classical machine [12,13]. However, imperfections in photon sources and losses in optical interferometer networks and detectors rapidly diminish the degree of quantum correlations and overthrow the quantum advantage. The diminishing effect of imperfections on QA is quantified in models of "noisy" boson samplers for which efficient classical computation algorithms have been proposed, thereby imposing strict bounds on the indistinguishability of photons and the overall optical loss [14–16]. As quantum photonic hardware continues to rapidly advance [17–20], the formulation of quantitative benchmarks for realizing QA is a critical need as they provide milestones for guiding the hardware development.

In this paper, we present a comprehensive analysis quantifying the performance metrics of the constituent building blocks essential for surpassing efficient classical algorithms with boson sampling. By benchmarking our framework against state-of-the-art single-photon sources and photonic integrated circuits, we identify a realistic regime for conducting QA experiments. Our analysis focuses on the case of boson samplers based on single-photon sources (discrete variable photonic qubits), where several proof-of-concept experiments have been carried out to date [21–25].

Recent advances in deterministic photon sources employing semiconductor quantum dots have demonstrated the generation of long strings of nearly identical photons [26,27], setting the stage for scaling up from proof-of-concept experiments. A remaining challenge is to realize a photonic platform with sufficiently low loss, requiring efficient collection from the source, as well as large-scale and low-loss photonic circuits such that the large photonic resource can be processed and measured to demonstrate QA [14,19]. As the optical loss in an interferometer circuit is highly dependent on its design, i.e., the spatial arrangement of Mach-Zehnder interferometers (MZI), we analyze the requirements of the individual MZIs and their integration into an optimal architecture. Combining the analysis of the source imperfections and optical circuit

*lodahl@nbi.ku.dk

loss, we identify two key optimizations, a rectangular circuit architecture and the encoding of modes in multiple degrees of freedom, that could enable an unequivocal demonstration of QA. We determine that an insertion loss of 3.5 mdB per MZI interferometer, i.e., a transmittance of 99.92 % is sufficient for the optimal architecture using state-of-the-art photon sources and detectors. This sets a clear target metric for ongoing advances in photonic integrated circuits [28,29], and is already within reach for the specialized, fixed circuits employed in Refs. [25,30].

II. BOSON SAMPLING ALGORITHM AND ITS VALIDATION

Boson sampling is the task of sampling from the output photon distribution after multiple independent photons have interfered in a multimode linear optical interferometer. The setup for implementing boson sampling is schematically illustrated in Fig. 1(a) highlighting the key components: an input consisting of multiple indistinguishable photons, a large multimode interferometer module, and single-photon detectors.

The computational complexity of simulating boson sampling arises from the connection between multiphoton correlations and the calculation of matrix permanents. Analyses of the computational complexity of boson sampling typically assume that both input states and output states are collision free [8,15], meaning that each mode contains at most one photon. To preserve the computational complexity, it is crucial to ensure collision-free states at both the inputs and outputs [8], meaning that each mode contains at most one photon. Collision-free input states can be guaranteed by choosing the initial condition, where no more than one photon is injected into each input mode. Ensuring collision-free outputs demands the interferometer to possess a large number of modes per photon. Specifically, the number of modes m must scale at least quadratically with the number of photons p, i.e., $m \propto p^2$, a requirement arising from a phenomenon called the *bosonic birthday paradox* [8,31]. In experiments, collision-free outputs can be ensured through postselection of events where photons are detected in the same number of output modes as input modes while discarding all other events. This postselection strategy remains applicable even when the detectors lack number-resolving capabilities, thereby enabling near-term implementations of QA with efficient single-photon detectors [25,30,32].

To demonstrate QA, it is essential that the output samples can be validated as being computationally hard to produce by classical means [14]. Due to the computational hardness, direct validation of the samples by comparison with exact distributions is infeasible. Within these constraints, the validation of QA through boson sampling requires two steps. First, we must require that deviations in the experimental setup are small enough that approximate classical algorithms cannot simulate the output distribution efficiently. Second, instead of validating that the samples are produced from the exact distribution, one verifies that the outputs are not reproduced by a computationally efficient distribution [14,25]. Specifically, statistical tests performed on the output samples obtained in a boson sampling experiment verify that the experimentally observed distribution differs significantly from a set of

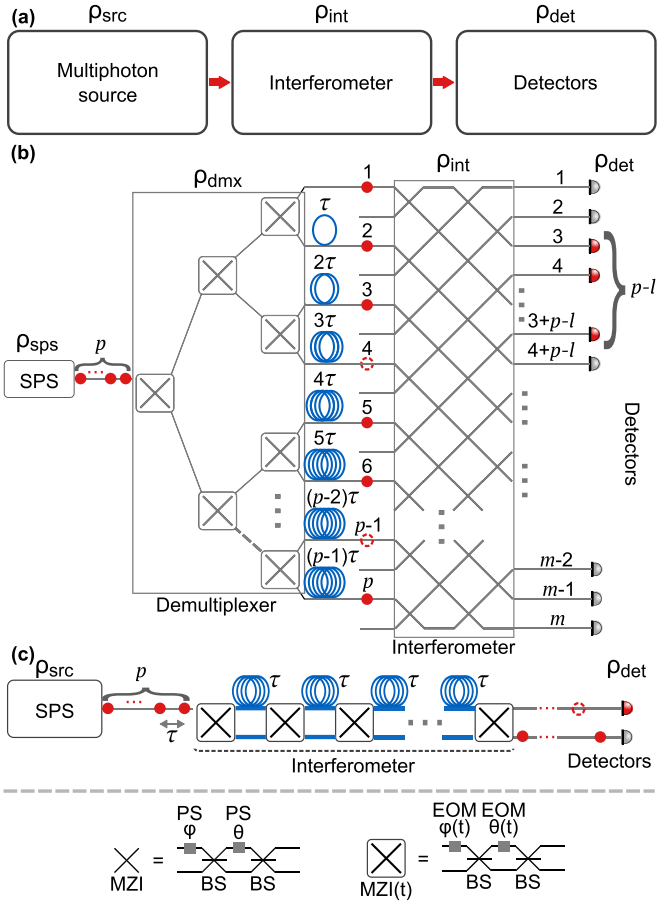

FIG. 1. (a) A general boson sampling setup, consisting of a source of multiphoton input states, a multiport interferometer, and detectors. The associated system losses ρ are indicated for each subcomponent. (b) A boson sampling setup based on a deterministic single-photon source (SPS) and a spatially encoded multimode interferometer. The single-photon source emits a string of single photons in predetermined time bins, illustrated as filled red circles. A demultiplexer converts the photon stream into a p-photon source by deterministically switching each photon into a separate mode. The spatially encoded input state is sent into an interferometer constructed from a network of MZIs. Each MZI is illustrated as a cross, where each cross consists of two phase shifters (PSs) and two beam splitters (BSs) arranged according to the illustration at the bottom. (c) A boson sampling setup based on a deterministic single-photon source (SPS) and a time-bin interferometer. The SPS is identical to the case in (b) where τ denotes the time separation between subsequent photons. The interferometer consists of reconfigurable time-dependent MZIs, where time-dependent phase shifters are implemented using electro-optic modulators (EOMs), which are connected by delay lines illustrated as black lines, where one loop corresponds to a single time bin τ of delay. For both (b) and (c), it is assumed that up to l photons may be lost throughout the setup, illustrated as a red circle with a dashed outline, due to residual optical loss associated with each component.

efficiently computable distributions. These statistical tests provide a termination condition for the experiment, whereby the boson sampler is run until a sufficient number of output samples are generated to establish the statistical tests' convergence unequivocally. Thus, a QA demonstration will

be feasible if a sufficient number of samples can be produced over an experimentally viable integration time, proportional to the sample acquisition rate r_{sample}. The sample acquisition rate in an experiment is equal to the product of the generation rate of the multiple-photon input state r_{input} and the probability of the state reaching the detectors P_{sample}:

$$r_{\text{sample}} = P_{\text{sample}} r_{\text{input}}. \tag{1}$$

The probability of sampling a p-photon coincidence at the output is related to the total per-photon efficiency of the system P_{sys} as

$$P_{\text{sample}} = P_{\text{sys}}^{p}. \tag{2}$$

For convenience, we express these probabilities in decibels, i.e., $\rho_{\text{i}} = -10\ \log_{10}(P_{\text{i}})$. For brevity, we refer to the decibel probability ρ_{i} as loss, while P_{i} is referred to as efficiency.

III. EXPERIMENTAL SETUP AND IMPERFECTIONS

As illustrated in Fig. 1(a), the loss in the boson sampling architecture can be broken down into component-level losses as

$$\rho_{\text{sys}} = \rho_{\text{src}} + \rho_{\text{int}} + \rho_{\text{det}}, \tag{3}$$

where ρ_{src} is the source loss, ρ_{int} is the interferometer loss, and ρ_{det} is the detector loss. In this section, we detail the implementation and requirements on the components and discuss strategies to tackle experimental limitations in each component.

A. Input-state preparation

The input state in the boson sampling algorithm consists of multiple indistinguishable single photons, each generated and encoded separately. Below, we discuss these two steps and any imperfections introduced in real experiments.

An important figure of merit for the photons employed in a boson sampling experiment is the pairwise indistinguishability x^2, defined as the overlap integral of photons with wave functions ψ and ψ', $x^2 = |\langle\psi'|\psi\rangle| = |\int \psi'^{*}(t)\psi(t)\mathrm{d}t|$. In experiments, the indistinguishability is quantified through the Hong-Ou-Mandel visibility in an interference experiment [33]. To ensure computational hardness, photon indistinguishability approaching near-unity visibility is necessary.

Highly indistinguishable photons have been generated employing two approaches: (1) nonlinear optics and (2) single quantum emitters. The former exploits the energy-time correlations in optical processes in nonlinear media like spontaneous parametric down-conversion and spontaneous four-wave mixing to generate correlated photon pairs. The resulting squeezed state can be used directly for Gaussian boson sampling [4,9,30] or, alternatively, the detection of one photon in the pair can be used to herald the presence of the other. Despite being probabilistic, the heralded nature can be exploited through active feedforward and multiplexing to realize a near-deterministic photon source [34]. As the sources suffer from an intrinsic tradeoff between the photon-number purity (the probability of having one and only one photon pair per pulse) and the photon-pair generation rate, this requires massive multiplexing of many probabilistic sources to reach near-deterministic operation with high number purity, which is an active area of research [35]. These challenges and tradeoffs could be overcome by leveraging the deterministic light-matter interactions enabled by coupling a highly efficient single-photon emitter (e.g., a semiconductor quantum dot) to a nanophotonic structure [36]. Single-photon sources based on such deterministic light-matter interfaces produce near-identical photons and are operational at a rate of up to a GHz, as defined by the lifetime of the emitter, thus highlighting an avenue for realizing boson sampling in the QA regime [26].

Photons generated using such deterministic light-matter interfaces are naturally encoded in different time bins as determined by the excitation process of the quantum emitter. While these time-bin encoded photonic states can be readily employed for boson sampling [37,38], compatibility with photonic integrated circuits requires converting the time-bin encoded photons to spatial encoding using a demultiplexer [25,39]. The demultiplexer takes an incoming stream of p single photons separated in time and converts it to p simultaneous photons in separate spatial modes. In both time-bin [see Fig. 1(c)] as well as demultiplexed spatial-bin encoding [Fig. 1(b)], in the demultiplexed spatial-bin-encoded case, a single-photon source emitting at a rate of $r_{\text{single-photon}}$ is converted to a source of p single photons emitted at a rate of $r_{\text{single-photon}}/p$. With a time-bin encoding, the rate will depend on the time-bin interferometer architecture employed, but for the case we will analyze, the input-state generation rate will be $2r_{\text{single-photon}}/m$, where m is the number of modes, such that $m/2$ is the number of time bins, as will be explained in more detail later.

A variety of quantum emitters are being explored as candidates for on-demand single-photon sources [40–42], with semiconductor quantum dots leading the field. As quantum dot sources have proven capable of producing long strings of highly indistinguishable photons [26], the scalability of this approach is mainly determined by the overall efficiency of generating single photons and delivering them to the demultiplexer, as well as the actual efficiency of the demultiplexer setup. We specifically consider and benchmark the case of a single deterministic source but note that access to multiple deterministic quantum dot sources would allow for spatially encoded experiments with higher input-state generation rates and lower demultiplexer losses. The simultaneous use of multiple sources relies on the development of local tuning methods for overcoming intrinsic inhomogeneities of quantum dot sources, and important progress has recently been reported both for quantum dots in bulk samples [43] and in nanophotonic waveguides [44].

A time-to-spatial-mode demultiplexer can be realized by sending the emitted photon stream through a binary tree of switches. Each step in the tree doubles the number of spatial modes, as illustrated in Fig. 1(b), such that the full demultiplexer requires a depth of $\lceil \log_2(p) \rceil$. Each output mode from the demultiplexer requires a specific optical delay to synchronize all photons. We associate a loss of ρ_{switch} with each switching operation, such that the overall loss of the demultiplexer is

$$\rho_{\text{dmx}} = \lceil \log_2(p) \rceil \rho_{\text{switch}}.$$

Finally, there will be a coupling loss ρ_{coupling} associated with connecting the source to the input of the demultiplexer, and the output of the demultiplexer to the input of the interferometer. Thus, the overall source loss is

$$\rho_{\text{src}} = \rho_{\text{sps}} + \rho_{\text{dmx}} + \rho_{\text{coupling}},$$

where ρ_{sps} is the loss associated with the single-photon source itself, accommodating for inefficiencies associated with the generation of single photons and subsequent coupling from the cavity or waveguide applied [36].

We note that the number of photons generated should ideally be kept as low as possible (for the targeted computational complexity) while maintaining intractability, as the sample acquisition rate will decrease exponentially with system efficiency for an increasing number of input photons in accordance with Eq. (2). In order to increase acquisition rates, experimental efforts typically employ a related algorithm called *Aaronson-Brod* boson sampling [45], where an additional l photons are added to the p input photons, while the outputs are postselected to contain the same number of photons as before. The probability of detecting the correct number of photons, i.e., the probability of generating a sample $P_{\text{sample}}(p, l)$ can then be expressed as

$$P_{\text{sample}}(p, l) = P_{\text{sys}}^{p-l}(1 - P_{\text{sys}})^l \binom{p}{l}, \tag{4}$$

where the factor $\binom{p}{l}$ is the number of combinations in which one can lose l photons from p input photons. This leads to a speedup in the sample acquisition rate, which increases combinatorially with the number of lost photons l. The downside is that postselection increases the deviation from the ideal case and lowers the computational complexity. In practice, the algorithm in Ref. [15] can produce samples from an approximate distribution where the error of the approximation E is bounded by

$$\frac{x^{2(k+1)}\left(\frac{p-l}{p}\right)^{k+1}}{1 - \left(x^2 \frac{p-l}{p}\right)} \geqslant E^2, \tag{5}$$

where x^2 is the indistinguishability of the photons. Thus, the number of lost photons allowed depends on the indistinguishability of the photons in the input state.

B. Interferometer design and operation

Multimode interferometers are typically constructed using cascaded arrays of Mach-Zehnder interferometers (MZIs) and implement an m-dimensional unitary matrix operation, with the circuit accommodating m input modes and m output modes. To validate boson sampling through statistical sampling of random unitary circuits [8], experiments require either several static circuits or a programmable circuit of MZIs with adjustable beam splitters and phase shifters. The latter approach is favored, as it offers access to a substantial number of random unitary operations [46]. Universal interferometer architectures enable the implementation of all unitary transformations on m modes, effectively generating any $m \times m$ unitary matrix [47,48].

The per-photon loss for an interferometer architecture depends on the number of optical components, i.e., MZIs, a photon passes through from the input to the output. We will refer to this number of MZIs as the *optical depth* $D(m)$, such that the interferometer loss can be written as

$$\rho_{\text{int}} = D(m)\rho_{\text{MZI}}.$$

The depth will increase for a higher number of modes where the exact dependence is given by the specific architecture employed. As such, there are two main strategies that are employed to reduce loss: 1. reducing the number of modes, and 2. employing interferometer architectures with a lower optical depth for a given number of modes.

C. Reducing the number of modes

Due to the challenges in scaling up low-loss interferometers, experiments involving a large number of photons ($p > 10$) have employed interferometers with the number of modes m smaller than p^2. Consequently, this choice yields a proportionally reduced depth $D(m)$. Although this reduction in the number of modes mitigates optical loss within the interferometer, the bosonic birthday paradox no longer holds, i.e., the outputs cannot be assumed to be collision free. In this scenario, the output event where multiple photons occupy the same mode will be indistinguishable from photon loss unless detectors with number-resolving capabilities are utilized. As for the computational complexity, Ref. [49] established that the computational complexity of simulating of boson sampling with linear modes is similar to the case with quadratic modes, even without postselecting on collision-free outputs.

Assuming non-number-resolving detectors are used with a postselection criterion that the detection of photons is restricted to d ($< p$) modes, with p representing the number of input photons, Eq. (4) can be reformulated to derive the total sample efficiency $P_{\text{sample, lin}}$ for a linear number of modes ($m \propto p$)

$$P_{\text{sample, lin}}(p, d) = \sum_{l=0}^{d} P_{\text{sys}}^{(p-l)}(1 - P_{\text{sys}})^l \binom{p}{l} P_{\text{ps}}(p, d, l, m). \tag{6}$$

Here, $P_{\text{ps}}(p, d, l, m)$ is the probability of detecting an output state with $d - l$ collisions and $p - l$ photons from an interferometer with m modes, essentially quantifying the effective postselection efficiency. This equation involves summing the probabilities of all possible collision and photon-loss configurations that result in photodetection in d modes. These probabilities are then multiplied by the occurrence probability of the given combination of losses and collisions. The computation of Eq. (6) relies on the knowledge of the effective postselection efficiency $P_{\text{ps}}(p, d, l, m)$. To estimate this quantity, we assume uniform sampling of Haar-random scattering matrices in the Hilbert space. Thus, we can estimate the effective postselection efficiency as the ratio between the size of the Hilbert space with $p - l$ photons in m modes with $p - l - d$ collisions, essentially the postselected portion of the Hilbert space, and the full Hilbert space for $p - l$ photons in m modes. This is expressed as

$$P_{\text{ps}}(p, d, l, m) = \frac{\binom{m}{d}\binom{p-l-1}{p-l-d}}{\binom{m+p-l-1}{p-l}}, \tag{7}$$

where $\binom{a}{b}$ represents the binomial coefficient for a and b. For more details, we refer to Appendix B.

D. Path-encoded boson sampling circuit architectures

The physical design of the interferometer architecture depends on the mode encoding employed. We will first present the architectures for spatial encoding. While theoretical works typically assume the use of universal interferometer architectures [8], larger-scale experimental endeavors have so far featured interferometers constructed from nonuniversal architectures [4,25,30] with lower optical depth. Formally, the complexity arguments from Refs. [8,45] are valid only in the universal case, but practically, quantum advantage experiments are hard to simulate even in nonuniversal cases. To provide an overview of how the different approaches compare, we will examine three interferometer architectures: a universal architecture, a fully connected architecture with single-mode encoding, and a fully connected hybrid-mode-encoded architecture. Here, "fully connected" signifies that all of the output ports of the interferometer are connected to all of the input ports, generally resulting in a unitary matrix where all elements are nonzero, but that in general can be nonuniversal. In all three cases, the architectures were chosen to balance the loss per photon to the best possible degree for all input-output configurations. For the universal architecture, and to a lesser extent the nonuniversal architecture for a given mode encoding, paths along the edges traverse fewer MZIs. Since the number of such edge cases is negligible, we assume uniform loss for all input-output configurations in our analysis.

The universal architecture introduced by Clements *et al.* [48], which we will refer to as the "Clements" architecture for simplicity, is constructed from m columns of MZIs, as shown in Fig. 1(a). The interferometer loss in the Clements architecture is given by $\rho_{\text{int}} = m\rho_{\text{MZI}}$. The interferometer loss can be optimized by employing nonuniversal architectures. We propose a nonuniversal "rectangular" interferometer architecture [see Fig. 1(b)], which maintains full connectivity but reduces the depth D. This is achieved by reducing the number of input modes to be equal to the number of input photons p, while maintaining the same number of output modes m. The interferometer consists of an initial section where two modes are added at the edges of each additional column of MZIs, and a second section fully connecting every input mode to all output modes. In practice, the rectangular architecture is equivalent to starting a Clements architecture partway in, distributing input modes starting from the middle, and removing unused MZIs. The interferometer loss ρ_{int}, given by the number of MZI columns multiplied by the MZI insertion loss, is $(\lceil \frac{m}{2} \rceil + \lceil \frac{p}{2} \rceil - 1)\rho_{\text{MZI}}$. Notably, when the number of output modes is much larger than the number of input modes, i.e., the number of input photons, such that $m \gg p$, ρ_{int} is approximately halved compared to the Clements architecture.

Next, we introduce a "hybrid-mode interferometer architecture" inspired by recent experiments [25,30]. These interferometers encode modes over multiple degrees of freedom, e.g., path or polarization, resulting in a hybrid-mode encoding. For instance, one degree of freedom might represent spatial modes separated in the x direction, while another represents spatial modes separated in the y direction as shown in Fig. 2(c). Concatenating fully connected interferometers in each direction results in an interferometer that is fully connected across all modes. The power of this approach lies in the way the number of modes and the depth scale with degrees of freedom. The total number of modes in the interferometer is equal to the product $m = \prod_i m_i$, where m_i is the number of modes encoded over the ith degree of freedom. The optical depth, however, is equal to the sum of individual optical depths $\rho_{\text{int}} = \sum_i D_i \rho_{\text{MZI},i}$, where D_i is the depth for the interferometer connecting all modes for the ith degree of freedom, and $\rho_{\text{MZI},i}$ is the MZI insertion loss for the ith degree of freedom. As an example, if we encode modes over two degrees of freedom, with $m_x = \sqrt{m}$ modes in the x direction and $m_y = \sqrt{m}$ modes in the y direction, the total number of modes remains as $m_x m_y = m$. A fully connected interferometer can then be constructed from Clements interferometers over the m_x modes followed by Clements interferometers over the m_y modes, as illustrated in Fig. 2(c). The total optical depth will then be $D_x + D_y = m_x + m_y = 2\sqrt{m}$. This approach allows for efficient scaling of both modes and depth, and it is worth noting that the Clements interferometers in each mode encoding can be replaced with rectangular interferometers to reduce depth further. The optical depth scalings for the different architectures can be summarized as follows:

$$\rho_{\text{int}} = \begin{cases} m\rho_{\text{MZI}}, & \text{Clements} \\ \left(\lceil \frac{m}{2} \rceil + \lceil \frac{p}{2} \rceil - 1\right)\rho_{\text{MZI}}, & \text{rectangular} \\ \sum_i D_i \rho_{\text{MZI},i}, & \text{hybrid.} \end{cases} \tag{8}$$

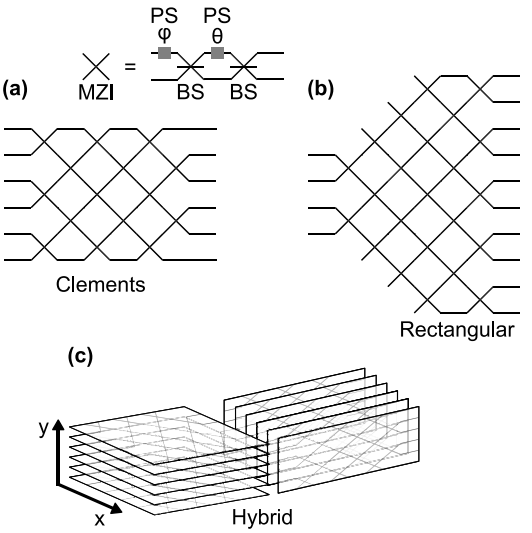

FIG. 2. (a) Illustration of the Clements interferometer architecture, as detailed in Ref. [48]. Each cross corresponds to an MZI, which can be constructed from two 50:50 beam splitters (BSs) and two phase shifters (PSs). (b) Illustration of a rectangular interferometer with a larger number of output modes than input modes, the latter of which is equal to the number of input photons. The interferometer can be described by a rectangular matrix, hence the name. (c) Illustration of an interferometer with multiple-mode encodings, i.e., a hybrid-mode encoding. In this case, one mode encoding has spatial modes separated in the x direction, while the other has modes separated in the y direction.

E. Time-bin-encoded interferometer architectures

Time-bin interferometers make use of time-dependent MZIs and fiber delays to implement multimode interferometers with significantly fewer physical resources than their spatially encoded counterparts. This is achieved by reconfiguring each physical MZI to implement a different transformation for every time bin. For this to be possible, the MZIs must be rapidly reconfigurable, i.e., the bandwidth must be substantially higher than the repetition rate of the single-photon source. Several interferometer architectures have been proposed and implemented [37,38,50,51], which differ in the number of spatial modes, and number of physical interferometers. We will analyze the architecture in Ref. [38], with one modification: instead of using a single MZI, we employ a cascaded series of multiple physical MZIs based on the interferometer employed in Ref. [50]. This choice is due to the lower total propagation loss and higher-input-generation rate of cascaded interferometers, as explained in Appendix C. The multiphoton-state input into this time-bin interferometer is encoded over two spatial modes of the MZI and $m/2$ time bins. Each physical MZI in the cascade implements operations equivalent to a column of MZIs in the Clements or rectangular interferometer, as illustrated in Fig. 3. The number of physical MZIs in the cascade then determines the depth $D(m)$ of the interferometer. In addition, a relative time delay of one time bin is introduced between the two output arms of the MZI between each step in the cascade, which ensures interference between photons in separate time-bins.

Figure 3 shows how certain spatial interferometers can be converted to time-bin interferometers. Specifically, we can construct the equivalent of spatial interferometers by combining three different MZI column types as shown in Fig. 3(a). All of the interferometer columns have the exact same construction for the time-bin interferometer architecture, a time-dependent MZI with a delay in one of the output modes, with the only difference being the sequence of transformations. The first MZI column type, labeled C_1, interferes all modes pairwise, starting with the first mode. The time-bin implementation of C_1 increases the total number of time bins at the output by 1 due to the time delay, where the first mode and last mode will occupy their own time bins. The second interferometer column type C_2 differs from C_1 in the nature of the input state, where the first and last time bins of the input states are occupied by only one mode, as shown in Fig. 3(c). The operation of the MZI on this asymmetric input state results in two additional modes (time bins) at the output in addition to the time delay. The final interferometer column type C_3 takes an input state where the first and final time bins are only occupied by one of the modes and enacts a swap transformation

$$\mathrm{U}_{\mathrm{swap}} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

on the first and last time bins, as shown in Fig. 3(d). This reduces the number of time bins by one compared to the input state. This type of column effectively interferes the modes pairwise starting with the second mode, such that the first and last modes do not interfere with any other mode.

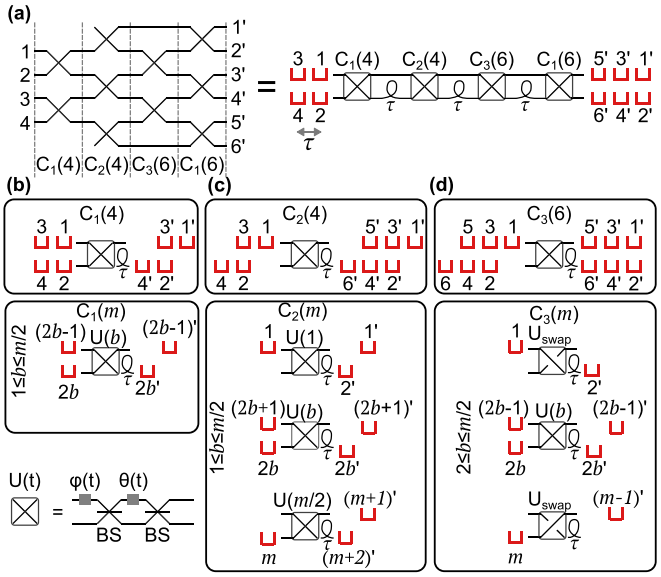

FIG. 3. Illustration of how time-bin interferometers can be constructed to implement specific features in interferometer architectures. The red bin symbols ⊔ correspond to the modes of the interferometer, and τ is the temporal separation between photons, which is equal to the inverse of the SPS emission rate $\tau = r^{-1}_{\text{single-photon}}$. Numbers correspond to input modes, and numbers with primes correspond to output modes. All of the time-bin interferometer architectures make use of reconfigurable MZIs, illustrated as boxed crosses. As is shown in the bottom-left corner, these MZIs are constructed using time-dependent phase shifters $\phi(t)$ and $\theta(t)$, which determine the unitary transformation U(t) effected by the interferometer. (a) Illustration of a rectangular interferometer architecture implemented with spatial modes on the left-hand side and with a time-bin interferometer on the right-hand side. The MZI columns are split into three categories $\mathrm{C}_1(m)$, $\mathrm{C}_2(m)$, and $\mathrm{C}_3(m)$, where m is the number of modes which is assumed to be an even number. (b) Upper: illustration of a specific MZI column for four input modes $\mathrm{C}_1(4)$. Lower: operation protocol of the $\mathrm{C}_1(m)$ MZI column for a general even number of input modes m. (c) Upper: illustration of a specific MZI column for four input modes $\mathrm{C}_2(4)$. Lower: operation protocol of the $\mathrm{C}_2(m)$ MZI column for a general even number of input modes m. This type of MZI column increases the total number of modes in the time-bin interferometer by two. (d) Upper: illustration of a specific MZI column for six input modes $\mathrm{C}_3(6)$. Lower: operation protocol of the $\mathrm{C}_3(m)$ MZI column for a general even number of input modes m. In the first and last time bins, the MZI must be configured to perform the operation $\mathrm{U}_{\mathrm{swap}}$, which swaps the two input modes. This ensures that the number of output modes from the column is equal to the number of input modes m.

In order to construct a specific architecture, we only need to identify the order and type of MZI columns that must be implemented. A Clements interferometer can be implemented by combining $m/2$ pairs of $\mathrm{C}_1(m)$ and $\mathrm{C}_3(m)$ columns, such that the optical depth, i.e., the number of physical MZIs in the interferometer, is m. The rectangular interferometer consists of a $\mathrm{C}_1(2p)$ column followed by $(m-2p)/2$ $\mathrm{C}_2(m')$ columns where m' is increased from $2p$ at the input of the interferometer to m at the interferometer by increasing the number of modes by 2 for every column. This is followed by $p-1$ pairs of $\mathrm{C}_3(m)$ and $\mathrm{C}_1(m)$ columns, resulting in an optical depth

of $(m+2p)/2-1$. Note that we have increased the number of input modes from p in Eq. (8) to $2p$ as one SPS can only populate one of the spatial input modes as shown in Fig. 1, i.e., half of the input modes of the rectangular interferometer.

To account for the effect of delay lines when establishing hardware requirements for time-bin interferometers, we can adjust the MZI insertion loss to include the propagation loss of the output with the longest delay

$$\rho_{\text{MZI, time bin}} = \rho_{\text{MZI}} + \tau\rho_{\text{prop}} + \rho_{\text{coupling}}, \tag{9}$$

where τ is the separation between time bins, ρ_{prop} is the propagation loss per unit time for the delay lines, and ρ_{coupling} is the coupling loss associated with going from one MZI to the next, i.e., coupling into and out of delay lines.

Lastly, we have to account for the empty time bins that are required to implement an interferometer where the number of modes (output modes in the case of rectangular interferometers) exceeds $2p$. The interferometer can only process $m/2$ time bins at a time, which necessitates a temporal separation of $m/2$ time bins between the start of two input states. Therefore, the input-state generation rate r_{input} is limited in comparison to the single-photon generation rate as

$$r_{\text{input}} = \frac{2r_{\text{single-photon}}}{m}. \tag{10}$$

IV. SYSTEM BENCHMARK FOR IMPLEMENTING THE AARONSON-BROD BOSON SAMPLING ALGORITHM

Before delving into the component-level hardware benchmarks, we address the overall system requirements, focusing on the two key system parameters, photon indistinguishability x^2 and system loss ρ_{sys}. We analyze the interplay between these two parameters in implementing the Aaronson-Brod boson sampling algorithm with lost photons. To provide a practical assessment of hardware performance, we choose realistic experimental conditions for run time and error rate of approximate classical algorithms for QA demonstrations with $p \geqslant 50$ photons.

We use the coupon collectors problem [26,52] to estimate the number of samples required for validation

$$r_{\text{sample}} t_{\text{integration}} \approx m\ \ln(m)/p, \tag{11}$$

where $t_{\text{integration}}$ is the total run time of the boson sampler and $r_{\text{sample}} t_{\text{integration}}$ is the total number of samples acquired. We set the target for r_{sample} to be 100 samples per day, $r_{\text{sample}} = 100/(24\times 3600)$ Hz, which is in line with the measured 6/3600 Hz photon coincidence rate reported in Ref. [25] that allowed validating the boson sampling experiment.

We analyze the boson sampling experiment for a deterministic single-photon source at the input that emits photons at a rate of 1 GHz. We will assume that threshold detectors without number-resolving capabilities are used, where we assume the dark count rate of the detectors is sufficiently low ($\sim$1 Hz) to have a negligible effect on the computational complexity, as explained in Appendix A. We consider both the case with a demultiplexer and a spatially encoded interferometer as well as the case with a time-bin interferometer. In the former, the p-photon input state is generated at a rate of $r_{\text{input, spatial}} = 1/p$ GHz. In the latter, the input state containing $M/2$ time bins is generated at a rate of $r_{\text{input, time bin}} = 2/m$ GHz. We consider two circuit architectures of the interferometer: one with a quadratic number of modes $m = (p-l)^2$ and another with a linear number of modes $m = 10(p-l)$, as detailed in Secs. III A and III C, respectively. By comparing the two cases we can examine the influence of mode scaling on the tradeoff between photon indistinguishability and system-loss tolerance. We combine Eq. (1) with (4) [Eq. (6)] for the quadratic (linear) case, to find the level of loss that results in an r_{sample} of 100 samples per day.

We choose an error rate of $E \leqslant 0.01$, given by Eq. (5), for approximating the noisy boson sampling output using classically computed permanents of order $k = 49$. We plot this relation for a varying number of detected photons $(p-l)$, where the number of lost photons l has been set as high as possible while keeping the error of the approximation below the threshold of $E \leqslant 0.01$.

Results for mode scalings $m = (p-l)^2$ and $m = 10(p-l)$ are shown in the upper and lower rows of Fig. 4, respectively. The lowest indistinguishability $x^2 \approx 0.805$, found by setting $l = 0$ and $p = 50$ in Eq. (5), corresponds to the maximal per-photon loss for Aaronson-Arkhipov boson sampling, i.e., with no photon loss or collisions. Increasing the photon indistinguishability allows for a higher number of lost photons with Aaronson-Brod sampling, increasing the per-photon loss that can be tolerated. We find that >3 dB (i.e., 50%) loss tolerance for realistic degrees of indistinguishability of quantum dot SPSs [36] for all four cases. We find the highest maximal loss values at an indistinguishability of $x^2 = 0.98$ to be, from highest to lowest: 3.78 dB for the quadratic spatial case, 3.46 dB for the quadratic time-bin case, 3.35 dB for the linear spatial case, and 3.20 dB for the linear time-bin case. For comparison, in the limit of perfect indistinguishability, $x^2 = 1$ where up to $l = 12$ photons can be lost with 50 detected photons, the highest maximal loss values would be 3.96 dB for the quadratic spatial case, 3.65 dB for the quadratic time-bin case, 3.53 dB for the linear spatial case, and 3.39 dB for the linear time-bin case.

As the number of detected photons increases the total loss of the system typically increases, leading to a decreased maximum per-photon loss. However, for a fixed degree of indistinguishability, gradually increasing the number of detected photons can lead to abrupt changes when an additional lost photon can be tolerated according to Eq. (5). Thus, the optimum number of lost photons and the detected photons will both depend on the exact photon indistinguishability in the experiment.

In comparing the different plots it is evident that quadratic-mode scalings and spatial interferometers lead to higher overall loss tolerance compared to the linear-mode scaling and time-bin interferometers. The advantage of quadratic-mode scaling can be attributed to the added effective postselection loss associated with linear-mode scaling, described by Eq. (7). Specifically, the average difference between quadratic- and linear-mode scalings is 0.430 dB for spatial interferometers and 0.277 dB for time-bin interferometers, in favor of quadratic interferometers. The advantage of spatial interferometers can be attributed to the lower input-state generation rate for time-bin interferometers. The average difference between spatial and time-bin interferometers is 0.298 dB for a

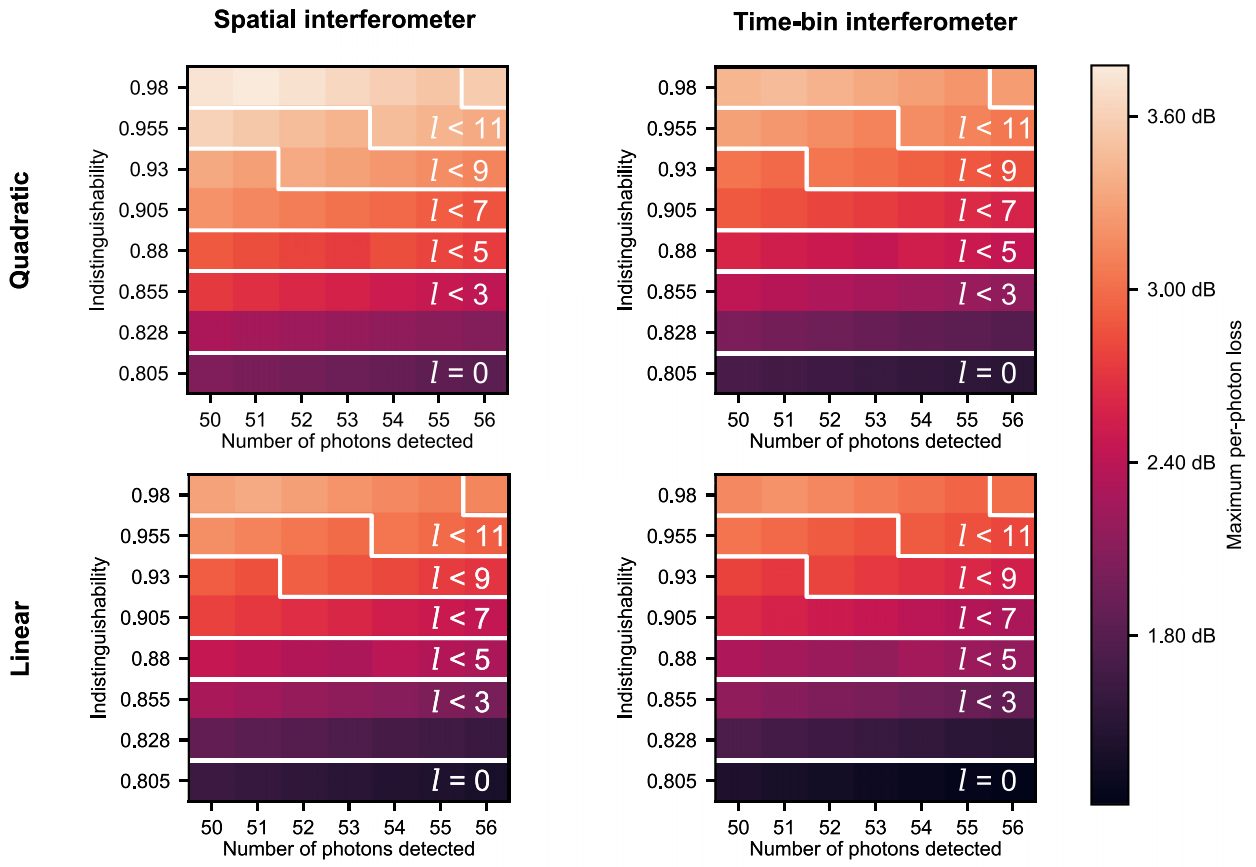

FIG. 4. Maximum per-photon loss for the full circuit to perform boson sampling in the QA regime versus degree of indistinguishability and number of detected photons and for different values l of lost or colliding photons. The two upper plots are for the case where the number of modes scales quadratically with the number of photons $m = (p-l)^2$, and the two bottom plots are for the case where the number of modes scales linearly with the number of photons $m = 10(p-l)$. The two plots to the left are for spatial interferometers with a demultiplexed source, whereas the two plots on the right are for time-bin interferometers. White contours indicate the added number of lost photons l, which increases with the indistinguishability, and detected number of photons.

quadratic number of modes and 0.145 dB for a linear number of modes, both in favor of spatial interferometers.

Although quadratic-mode scalings allow for higher per-photon loss, the interferometers consist of more MZIs. As such, there is a tradeoff between lower effective postselection loss for quadratic interferometers and lower interferometer loss for linear-mode scalings where the MZI insertion loss determines which mode scaling is favored. We note that using photon-number-resolving detectors would allow for the same maximum per-photon loss in both cases, as the postselection would allow for an arbitrary number of photon collisions (i.e., postselecting on $p-l$ detected photons).

Time-bin interferometers have a similar tradeoff, as they have a lower maximal per-photon loss due to the lower input-state generation rate, but do not require the use of a demultiplexer. However, a demultiplexer can be constructed from the same MZIs that are used to construct a time-bin interferometer and, as such, this tradeoff can also be quantified in terms of MZI insertion loss. Specifically, a demultiplexer has an optical depth of $\lceil \log_2(p) \rceil = 6$, where the equality holds for the optimal number of detected and lost photons for all indistinguishabilities considered in Fig. 4. The demultiplexer also involves a delay on all except the last photon, where the first photon has to go through the longest delay of $(p-1)$ time bins. As for the time-bin interferometer, all of the $D(m)$ MZIs in the interferometer will in the worst case include one time bin of delay which is not present in the spatial case. If we compare the added per-photon loss from the demultiplexer in the spatial case with the added delay and lower maximum per-photon loss in the time-bin case, we can find the following inequality for the regime where time-bin interferometers are less favorable implementations than spatial interferometers:

$$6\rho_{\mathrm{MZI}} + (p-1)\rho_{\mathrm{prop}} \leqslant [D(m)-1]\rho_{\mathrm{prop}} + \Delta,$$
$$\Delta = \rho_{\mathrm{sys,\,spatial}} - \rho_{\mathrm{sys,\,time\,bin}}. \tag{12}$$

The average of the value for Δ, i.e., the difference between maximal per-photon loss for spatial and time-bin interferometers, was found to be 0.145 dB for the case where $m = 10(p-l)$ and 0.298 dB for the case where $m = (p-l)^2$. If we neglect propagation loss $\rho_{\mathrm{prop}} = 0$ and insert the average values we can estimate this inequality in the two cases considered in Fig. 4:

$$\rho_{\mathrm{MZI}} \leqslant 0.05\,\mathrm{dB} \quad m = (p-l)^2,$$
$$\rho_{\mathrm{MZI}} \leqslant 0.024\,\mathrm{dB} \quad m = 10(p-l). \tag{13}$$

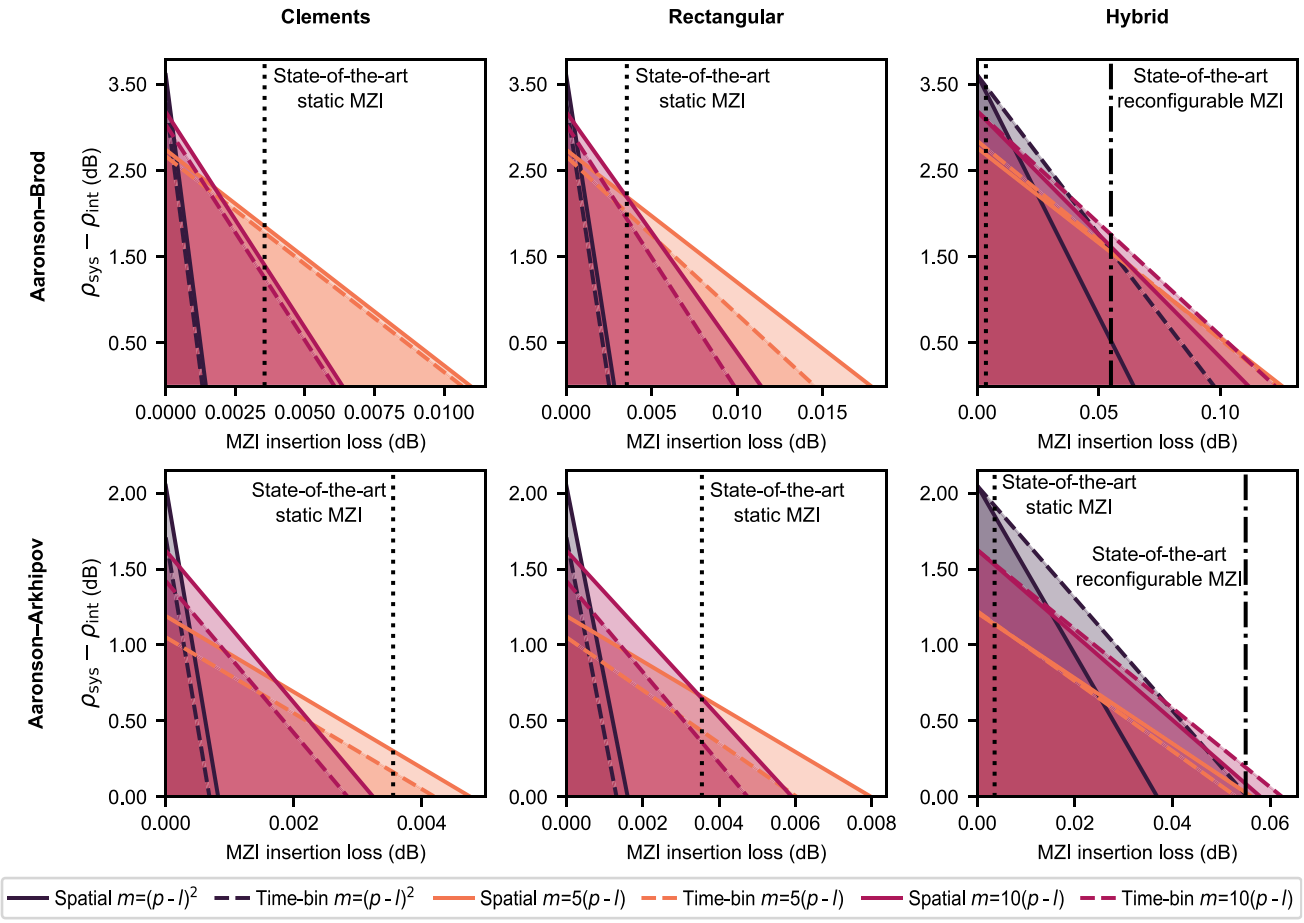

FIG. 5. Plot of the requirements on MZI insertion loss (x axis) and $\rho_{sys} - \rho_{int}$ (y axis) with photon indistinguishability set to $x^2 = 0.96$. The upper plots show the requirements for Aaronson-Brod boson sampling, where the input state consists of 59 photons with the outputs postselected to contain 50 photon detection events. The lower plots show the requirements for Aaronson-Arkhipov boson sampling, where we send in 50 photons and detect 50 photons, not allowing for photon loss or collisions. The first, second, and third columns show hardware requirements for setups where the interferometers are constructed according to the Clements architecture, the rectangular architecture, and a set of hybrid architectures, respectively. For the Clements and rectangular architectures, the solid lines correspond to the requirements for spatially encoded architectures, whereas dashed lines correspond to the requirements for time-bin architectures. For the hybrid architectures, the solid lines correspond to requirements for an interferometer encoded over two spatial-mode encodings, whereas dashed lines correspond to requirements for an interferometer encoded over time bins and two spatial-mode encodings. The dotted vertical lines mark the estimated MZI loss for a state-of-the-art experimental realization with static, nonprogrammable MZIs [25]. The dashed-dotted vertical line in the plots for hybrid interferometers marks the estimated MZI loss for a state-of-the-art experimental realization with programmable MZIs [55].

In practice, coupling into the delay between MZIs will inevitably incur loss and, as such, these inequalities present a best-case scenario in favor of time-bin interferometers.

V. BENCHMARKING HARDWARE REQUIREMENTS

The requirements on component losses for a given interferometer architecture can be found by combining Eqs. (3) and (8), with either Eq. (4) or (6). To simplify the analysis, we note that only the interferometer loss scales with the number of modes, and separate the losses into the interferometer loss, ρ_{int}, and the remaining system loss, $\rho_{sys} - \rho_{int} = \rho_{sps} + \rho_{dmx} + \rho_{coupling}\rho_{det}$. We fix the degree of indistinguishability to $x^2 = 0.96$, which is readily achievable with present-day quantum dot single-photon sources [53], while routes to achieve even higher values have been laid out [54]. This allows us to fix the number of input photons to $p = 59$ with up to $l = 9$ lost photons in accordance with Eq. (5). We also consider the requirements for Aaronson-Arkhipov sampling where we postselect on detecting the same number of photons as are sent into the interferometer, i.e., fixing the number of input and output photons to $p = 50$. The hardware requirements on the interferometer can be formulated as specific requirements on the MZI insertion loss by specifying the architecture and number of modes used for the interferometer.

Single-mode encoding. The first two columns of Fig. 5 show the requirements for three different choices of mode scaling for both Clements and rectangular interferometer architectures, where the top figures show requirements for Aaronson-Brod [45] boson sampling, and the bottom figures show requirements for Aaronson-Arkhipov boson sampling [8], i.e., with postselection on the same number of detected photons as input photons. The figures include a vertical dotted line corresponding to the state-of-the-art insertion

TABLE I. Table of state-of-the-art system efficiencies. In the column for demultiplexer efficiency, the number of modes of the demultiplexer employed is indicated in parentheses. For the current best indistinguishability, we have used the measured photon indistinguishability, whereas for the near-term estimated best we have used the estimated intrinsic photon indistinguishability which has been corrected for experimental imperfections.

Reference	P_{sps}	P_{coupling}	P_{dmx} (p)	P_{det}	Total efficiency	x^2
Current best	0.658 [56]	0.902 [25]	0.83 (20) [25]	0.95 [4]	0.47	0.964 [53]
Near-term estimated best	0.78 [26]	0.902 [25]	0.92 [26]	0.95 [6]	0.615	0.985 [53,58]
Target					$\gtrsim$0.65	0.96

loss for a static MZI, estimated to be $-10\ \log_{10}(0.987)/(10+6)$ dB ≈ 0.0035 dB, which is the overall interferometer efficiency in Ref. [25] divided by the number of MZIs corresponding to the optical depth of a 10-mode interferometer followed by a 6-mode interferometer.

Comparing the requirements on MZI insertion loss with Eqs. (13), we find that we are in the regime where demultiplexing and spatial-mode-encoded interferometers are favorable in terms of loss even with access to the rapidly reprogrammable MZIs required to construct a time-bin interferometer.

As seen from the figure, a QA demonstration using single-mode-encoding architectures would be within reach if one could use interferometers with state-of-the-art efficiency in conjunction with a source and detection efficiency of around $P_{\text{sps}}P_{\text{dmx}}P_{\text{coupling}}P_{\text{det}} \geqslant 0.65$. This overall efficiency is currently beyond the state-of-the-art values reported with quantum dot sources [6,25,56,57] (see Table I for an overview of parameters already reported in the literature). Further expected near-term improvements of the approach are also listed in the table, indicating that explicit QA demonstration with single-mode encoding is not far outside reach.

Hybrid-mode encoding. Hybrid-encoding schemes make the algorithms more robust to optical loss and hence put QA demonstrations within closer reach. We consider two distinct hybrid architectures, one with two spatial-mode encodings, like the one employed in Refs. [25,30] and illustrated in Fig. 2(c), and one with two spatial-mode encodings and one time-bin encoding. For the former, the number of modes in each encoding was chosen to optimize the optical depth, as described by Eq. (8). This optimization procedure allowed for the total number of modes to be slightly increased if it led to favorable optical depth. Rectangular interferometers were employed in each mode encoding to minimize depth. For the latter encoding, we considered an architecture where the input state is partially demultiplexed, whereby sets of two spatial modes are inserted into a time-bin interferometer. By ensuring that there are no empty time bins in the time-bin interferometer, we avoid the issue of a lowered input-state generation rate for time-bin interferometers. The output modes of the time-bin interferometers are then sent into a two-spatial-mode-encoding hybrid interferometer employing rectangular interferometers across the two spatial encodings. The total optical depth of this interferometer, including the initial demultiplexer, is equal to

$$D(n,p,m_1,m_2) = n + 2\left\lceil \frac{p}{2^n} \right\rceil + \left\lceil \frac{m_1}{2} \right\rceil + \left\lceil \frac{m_2}{2} \right\rceil + \frac{2^n}{2}, \tag{14}$$

where n is the depth of the demultiplexer, such that there are 2^n spatial modes and $\lceil \frac{p}{2^n} \rceil$ time bins after the demultiplexer, and where m_1 and m_2 correspond to the number of output modes in each spatial encoding. Similarly to the case of the spatial hybrid interferometer, we allow for the number of modes to be increased if it leads to a lower optical depth, only requiring that

$$m_1 m_2 \left\lceil \frac{p}{2^n} \right\rceil \geqslant m,$$

where the left-hand side is the actual number of modes, and the right-hand side is the target number of modes, calculated from the mode scaling. The depth of the demultiplexer and the number of spatial modes were optimized in order to minimize Eq. (14). For the case of a quadratic number of time bins, the ideal demultiplexer depth was found to be $n = 3$, with $m_1 = m_2 = 18$, whereas for the linear-mode scalings, the ideal demultiplexer depth was found to be $n = 4$ with the number of spatial modes equal to $m_1 = 11,\ m_2 = 12$ and $m_1 = 8,\ m_2 = 9$ for the case where $m = 10(p-l)$ and $m = 5(p-l)$, respectively.

The resulting hardware requirements for Aaronson-Brod (Aaronson-Arkhipov) boson sampling are shown in the top (bottom) plot of the right column of Fig. 5, where the solid lines ("spatial") refer to the case with two spatial-mode encodings, and the dashed lines ("time bin") refer to the case with time bin and two spatial encodings. In addition to the dashed line for the state-of-the-art static MZI insertion loss, the plots include a dashed-dotted line marking the state-of-the-art insertion loss for a reconfigurable MZI, estimated from Ref. [55] to be 1.1 dB/20 = 0.055 dB, which is the interferometer insertion loss divided by the number of MZIs. We note that though this circuit was designed for light at telecom wavelengths, the silicon nitride material platform that was employed is compatible with the wavelengths of state-of-the-art quantum dot sources.

The hybrid encoding with the time-bin encoding performs better at larger mode scalings but seems to perform comparably or slightly worse for the case where $m = 5(p-l)$. The advantage at higher mode scalings comes from the fact that the number of modes is distributed over three encodings, which means that the sum of the number of modes can be smaller. This is less of an advantage for the case where $m = 5(p-l)$, where the additional demultiplexer loss included in the time-bin case gives a higher overall depth. It should be noted that the spatial hybrid interferometer requires the addition of a demultiplexer at a depth of $n = \lceil \log_2(p) \rceil$, which should be included as part of the source efficiency. As such, the time-bin

hybrid interferometer would be expected to perform advantageously even for low mode scaling.

Figure 5 clearly shows that the MZI insertion loss determines whether a quadratic mode scaling or linear mode scaling is favorable. Specifically, at the state-of-the-art static MZI insertion loss, employing an interferometer with a quadratic-mode scaling is best, whereas linear-mode scalings are favored when using MZIs with the state-of-the-art reconfigurable MZI insertion loss. This is not the case for Clements or rectangular architectures except for the case where the MZI insertion loss is vanishingly low. The discrepancy between the Clements and rectangular architectures and the hybrid architecture is attributed to how the optical depth scales with the number of modes, as shown in Eq. (8). With modes distributed across two mode encodings, the optical depth of hybrid interferometers scales with $\sim\sqrt{m}$, as opposed to linearly in m. As increasing the number of modes has a smaller impact on interferometer loss, the added effective postselection loss associated with linear-mode scalings has a proportionally higher impact on the system loss for hybrid interferometers.

VI. HARDWARE REQUIREMENTS FOR NEAR-TERM QA DEMONSTRATIONS

From Fig. 5 it is observed that an explicit QA demonstration is within reach with a deterministic quantum dot source. Indeed, using state-of-the-art static interferometers would imply that QA is reached for $P_{\text{sps}}P_{\text{dmx}}P_{\text{coupling}}P_{\text{det}} \geqslant 0.45$, where the required efficiency of each subcomponent was already realized experimentally (see Table I). As time-bin interferometers require the use of reprogrammable MZIs, one could not make use of hybrid architectures with time-bin encoding in this case. As for state-of-the-art reconfigurable interferometers, this would require a setup with combined source and detection efficiencies around 0.65 (0.7) for the time-bin (spatial) hybrid architectures. These values are reachable with the estimated near-term values of the approach (cf. Table I). It is important to note that the state-of-the-art values of MZIs hold for thermo-optic phase shifters, which are unsuitable for realizing time-bin interferometers due to their slow response time. Consequently, this would limit the present implementations of hybrid time-bin architectures. On the other hand, hybrid interferometers with spatial-mode encodings appear to be promising candidates for near-term QA demonstrations with quantum dot single-photon sources.

It is clear from Table I that the source efficiency is the main bottleneck in realizing a demonstration of QA. In the following, we restrict the focus to the exact requirements for the single-photon source by fixing other losses to state-of-the-art values. We fix the MZI insertion loss to the state-of-the-art value for static MZIs shown in Fig. 5. The demultiplexer efficiency and coupling losses are fixed to realistic parameters extrapolated from Refs. [26,30]:

$$\rho_{\text{dmx}} = \frac{0.458}{5}\lceil \log_2(p)\rceil \text{dB},$$

$$\rho_{\text{coupling}} = 0.458\,\text{dB}.$$

We examine how the requirements on the source efficiency change as a function of the photon indistinguishability. The resulting curves for quadratic- and linear-mode scalings are shown in Fig. 6. It is clear from the figure that quadratic-mode scaling is favored regardless of photon indistinguishability. For realistic photon indistinguishability, $x^2 \geqslant 0.96$, the results show that a QA demonstration is within reach for single-photon source efficiencies greater than 0.6, which has been demonstrated experimentally [56]. The challenge will be to construct a demultiplexer and an interferometer that are sufficiently large while maintaining sufficiently low loss, and connecting them to an exceedingly large number of low-loss detectors.

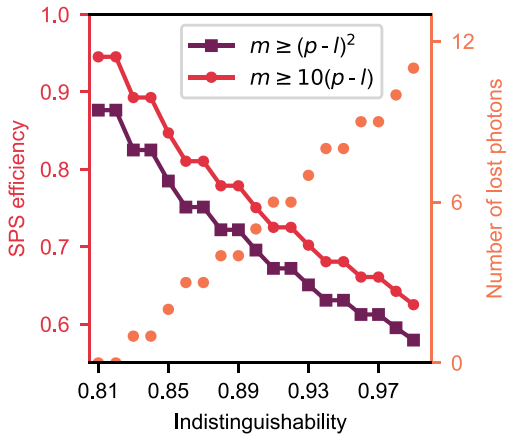

FIG. 6. Requirements on source efficiency for a given indistinguishability with a hybrid interferometer encoded in two mode encodings as per Fig. 1(d). Rectangular interferometers have been used within both mode encodings. The requirements are defined to allow for 100 samples to be obtained per day with a 1-GHz single-photon generation rate. For each value of indistinguishability, the number of photons detected and lost has been optimized to increase the loss tolerance while maintaining an error bound higher than 1% for the approximation algorithm.

VII. CONCLUSION

In conclusion, we have presented an in-depth analysis of the hardware requirements for realizing boson sampling in the QA regime with deterministic single-photon sources, notably quantum dots in nanophotonic cavities and waveguides. The estimated benchmarks provide precise requirements on optical circuits and single-photon sources that must be reached, thereby offering a road map for future engineering efforts to realize that goal. Our analysis elucidates the precise advantages and disadvantages of strategies that are commonly employed in experiments to lower hardware requirements, such as making use of specialized interferometer architectures and employing interferometers with linear-mode scaling. We have identified interferometers with hybrid-mode encoding and quadratic-mode scaling as a key strategy to demonstrating QA, an approach that has yet to see realization in experiments. Specifically, we have shown that a QA experiment based on single-photon boson sampling is within reach of current state-of-the-art hardware, provided that one can reach source efficiencies as high as 60%–70%.

In examining the requirements for time-bin-encoded interferometers, we have found time-bin interferometers utilizing the Clements and rectangular architectures to be inferior

to equivalent spatial interferometers for QA demonstrations. This is due to the disadvantages associated with empty time bins in the input state in the regime where MZI insertion loss is low enough for a QA demonstration to be feasible. For hybrid interferometer architectures, however, encoding a subset of the modes in time bins leads to improved hardware requirements on MZI insertion loss. As such, the development of phase shifters that can be reconfigured in-between time bins and that would be compatible with low-loss delays to enable low-loss time-bin-compatible MZIs is a promising direction to enable QA demonstrations.

Our analysis has focused on the requirements of the interferometers and single-photon sources, however, an underlying assumption for parts of the analysis was that the loss associated with coupling photons from the source and into the interferometer were comparable with the losses quoted in Refs. [25,30]. Achieving ultra-low-loss chip-to-fiber coupling is an important engineering challenge and an area of active research [59–61]. Ultimately, coupling losses could be further mitigated by a partial or full-scale system integration, whereby sources, demultiplexer, interferometers, and detectors would be combined in a single device [18], which constitutes an important future research direction.

ACKNOWLEDGMENTS

We acknowledge funding from the Danish National Research Foundation (Center of Excellence "Hy-Q," Grant Number DNRF139), the Novo Nordisk Foundation (Challenge project "Solid-Q"), the European Union's Horizon 2020 research and innovation program under Grant Agreement No. 820445 (project name Quantum Internet Alliance). S.P. acknowledges financial support from the European Union's Horizon 2020 Marie Sklodowska-Curie Grant No. 101063763, from the Villum Fonden research Grant No. VIL50326, and from the NNF Quantum Computing Programme.

P.L. is a founder of the company Sparrow Quantum which commercializes single-photon sources. The authors declare no other conflicts of interest.

APPENDIX A: ESTIMATING ERRORS INDUCED BY DARK COUNTS

In order to estimate the error induced by dark counts, we can take a similar approach as in Ref. [62] and find the likelihood that a sample has an additional photon loss replaced by a dark count. For commercially available superconducting nanowire single-photon detectors, the dark count rate r_{dc} is typically on the order of 1 Hz and, correspondingly, the probability that a detector gets a dark count p_{dc} for any given detected event will be

$$p_{\mathrm{dc}} = r_{\mathrm{dc}}/r_{\mathrm{input}}. \tag{A1}$$

The error due to a photon loss being replaced with a dark count compared to no photon loss and no dark counts can then be estimated by calculating

$$\frac{\binom{p-l}{1}(1-p_{\mathrm{sys}})\binom{m-p}{1}p_{\mathrm{dc}}(1-p_{\mathrm{dc}})^{m-p}}{p_{\mathrm{sys}}(1-p_{\mathrm{dc}})^{m-p}}, \tag{A2}$$

where the numerator is the probability for one out of $p-l$ photons being lost and one out of $m-p$ remaining detectors producing a dark count, with the denominator being the probability that the photon is not lost with no dark counts. For a dark count rate of 1 Hz, with 61 input photons (11 of which are lost) and a loss rate of 3.6 dB, i.e., the most lossy scenario with the lowest input rate, this expression evaluates to 0.0098.

In order to evaluate the impact of this effect on the computational complexity, we will assume a best-case scenario for the approximation algorithm, where every instance with a photon loss replaced by a dark count is simulable to zero error. The error the algorithm makes E_{dc} can then be estimated as

$$E_{\mathrm{dc}} = 0.0098 \times 0 + (1-0.0098)E_{\mathrm{algorithm}}, \tag{A3}$$

where $E_{\mathrm{algorithm}}$ is the error made by the algorithm in approximating the output with zero dark counts. If we set $E_{\mathrm{dc}} \leqslant 0.01$, we can find a corrected bound on the error of the algorithm

$$E_{\mathrm{algorithm}} \leqslant 0.0101, \tag{A4}$$

i.e., a negligibly small correction. Although we have only considered the case where a single photon loss is replaced by a dark count, and the probability of two photon losses or more being replaced by two or more dark counts is orders of magnitude lower, if we generously assume that the actual proportion of outputs with dark counts is a factor of 2 larger than this estimate, this would still only lead to a corrected bound of

$$E_{\mathrm{algorithm}} \leqslant 0.0102. \tag{A5}$$

APPENDIX B: SIZE OF THE HILBERT SUBSPACE FOR A GIVEN NUMBER OF PHOTON COLLISIONS

The number of basis states in the Hilbert space with a given number of collisions is equal to the product of the number of ways one can distribute d nonzero modes in m modes, and the number of ways one can place $p-l-d$ collisions into d nonzero modes. The first number is equal to the number of combinations without replacements with d choices from m possibilities:

$$\binom{m}{d}.$$

The second number is equal to the number of combinations with replacements with $p-l-d$ choices from d possibilities

$$\binom{(p-l-d)+d-1}{p-l-d} = \binom{p-l-1}{p-l-d}.$$

Consequently, the size of the Hilbert subspace with $p-d-l$ collisions, $n_{\mathrm{collisions}}(p,d,l,m)$, will be

$$n_{\mathrm{collisions}}(p,d,l,m) = \binom{m}{d}\binom{p-l-1}{p-l-d}. \tag{B1}$$

The size of the full Hilbert space n_{full} will be the number of ways one can place $p-l$ photons into m modes. This is equivalent to the number of combinations with replacements of $p-l$ choices with m possibilities:

$$n_{\mathrm{full}}(m,p,l) = \binom{m+p-l-1}{p-l}. \tag{B2}$$

We can then take the ratio between Eqs. (B1) and (B2) to find Eq. (7)

APPENDIX C: DOWNSIDES OF USING LOOP ARCHITECTURES FOR TIME-BIN INTERFEROMETER

Although it is possible to implement large multimode interferometers using only a single physical MZI connected to fiber delay loops, as in Refs. [37,38], this has two major downsides: higher propagation loss and severely reduced input-state rate. Much in the same way as in the cascaded case, a column of MZIs in the Clements or rectangular scheme can be implemented by sending the time bins through the MZI one by one and reconfiguring the MZI transformation for each time bin. In order to reuse the same physical MZI to implement additional columns, we can connect the outputs to the inputs through a delay loop, where the delay is sufficiently long that all output modes back at the input after the previous MZI column have been finished. As each column processes up to $m/2$ time bins, this requires that the loops have a delay of at least $m/2$ time bins. This is in comparison to the cascaded scheme where no such delay is necessary apart from the delay of one time bin in one of the modes, which will still be present in the loop architecture. To see the difference, we can compare the total delay t_{delay} for the worst case of the Clements scheme with a cascaded architecture and a loop architecture

$$t_{\text{delay, cascaded}} = \frac{m}{2} - 1\tau, \tag{C1}$$

$$t_{\text{delay, loop}} = (m-1)\left(\frac{m}{2} - 1\right)\tau, \tag{C2}$$

where τ corresponds to the separation between time bins. In other words, the total propagation loss scales linearly with the number of modes for cascaded time-bin interferometers, whereas it scales quadratically with the number of modes for loop time-bin interferometers.

The second downside is that one has to wait for the full output state to come out of the interferometer before processing a new input state. The time difference between the first time bin in the input state and the last time bin of the output state is equivalent to the delay in Eq. (C2). As the time-bin separation is related to the rate of the single-photon source, the input generation rate r_{input} for the loop time-bin interferometer will be given by

$$r_{\text{input}} = \frac{2r_{\text{single-photon}}}{m(m-1)}. \tag{C3}$$

This is approximately a factor $1/m$ worse than the corresponding rate for a cascaded interferometer.

[1] J. Preskill, Quantum computing and the entanglement frontier, arXiv:1203.5813.

[2] J. Preskill, Quantum computing in the NISQ era and beyond, Quantum **2**, 79 (2018).

[3] S. Boixo, S. V. Isakov, V. N. Smelyanskiy, R. Babbush, N. Ding, Z. Jiang, M. J. Bremner, J. M. Martinis, and H. Neven, Characterizing quantum supremacy in near-term devices, Nat. Phys. **14**, 595 (2018).

[4] L. S. Madsen, F. Laudenbach, M. F. Askarani, F. Rortais, T. Vincent, J. F. F. Bulmer, F. M. Miatto, L. Neuhaus, L. G. Helt, M. J. Collins *et al.*, Quantum computational advantage with a programmable photonic processor, Nature (London) **606**, 75 (2022).

[5] A. Morvan, B. Villalonga, X. Mi, S. Mandrà, A. Bengtsson, P. V. Klimov, Z. Chen, S. Hong, C. Erickson, I. K. Drozdov *et al.*, Phase transition in random circuit sampling, arXiv:2304.11119.

[6] N. Maring, A. Fyrillas, M. Pont, E. Ivanov, P. Stepanov, N. Margaria, W. Hease, A. Pishchagin, T. H. Au, S. Boissier *et al.*, A versatile single-photon-based quantum computing platform, Nat. Photon. (2024), doi:10.1038/s41566-024-01403-4.

[7] G. A. Quantum, Collaborators*†, F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, S. Boixo, M. Broughton, B. B. Buckley *et al.*, Hartree-fock on a superconducting qubit quantum computer, Science **369**, 1084 (2020).

[8] S. Aaronson and A. Arkhipov, The computational complexity of linear optics, in *Proceedings of the Forty-third Annual ACM Symposium on Theory of Computing* (ACM, New York, 2011), pp. 333–342.

[9] C. S. Hamilton, R. Kruse, L. Sansoni, S. Barkhofen, C. Silberhorn, and I. Jex, Gaussian boson sampling, Phys. Rev. Lett. **119**, 170501 (2017).

[10] S. Bartolucci, P. Birchall, H. Bombin, H. Cable, C. Dawson, M. Gimeno-Segovia, E. Johnston, K. Kieling, N. Nickerson, M. Pant, F. Pastawski, T. Rudolph, and C. Sparrow, Fusion-based quantum computation, Nat. Commun. **14**, 912 (2023).

[11] S. Paesani and B. J. Brown, High-threshold quantum computing by fusing one-dimensional cluster states, Phys. Rev. Lett. **131**, 120603 (2023).

[12] L. G. Valiant, The complexity of computing the permanent, Theor. Comput. Sci. **8**, 189 (1979).

[13] S. Scheel, Permanents in linear optical networks, arXiv:quant-ph/0406127.

[14] D. J. Brod, E. F. Galvão, A. Crespi, R. Osellame, N. Spagnolo, and F. Sciarrino, Photonic implementation of boson sampling: a review, Adv. Photonics **1**, 034001 (2019).

[15] J. Renema, V. Shchesnovich, and R. Garcia-Patron, Classical simulability of noisy boson sampling, arXiv:1809.01953.

[16] J. J. Renema, A. Menssen, W. R. Clements, G. Triginer, W. S. Kolthammer, and I. A. Walmsley, Efficient classical algorithm for boson sampling with partially distinguishable photons, Phys. Rev. Lett. **120**, 220502 (2018).

[17] A. W. Elshaari, W. Pernice, K. Srinivasan, O. Benson, and V. Zwiller, Hybrid integrated quantum photonic circuits, Nat. Photon. **14**, 285 (2020).

[18] R. Uppu, L. Midolo, X. Zhou, J. Carolan, and P. Lodahl, Quantum-dot-based deterministic photon-emitter interfaces for scalable photonic quantum technology, Nat. Nanotechnol. **16**, 1308 (2021).

[19] E. Pelucchi, G. Fagas, I. Aharonovich, D. Englund, E. Figueroa, Q. Gong, H. Hannes, J. Liu, C.-Y. Lu, N. Matsuda *et al.*, The potential and global outlook of integrated photonics for quantum technologies, Nat. Rev. Phys. **4**, 194 (2022).

[20] G. Moody, V. J. Sorger, D. J. Blumenthal, P. W. Juodawlkis, W. Loh, C. Sorace-Agaskar, A. E. Jones, K. C. Balram, J. C. F. Matthews, A. Laing *et al.*, 2022 Roadmap on integrated quantum photonics, J. Phys. Photonics **4**, 012501 (2022).

[21] M. A. Broome, A. Fedrizzi, S. Rahimi-Keshari, J. Dove, S. Aaronson, T. C. Ralph, and A. G. White, Photonic boson sampling in a tunable circuit, Science **339**, 794 (2013).

[22] J. B. Spring, B. J. Metcalf, P. C. Humphreys, W. S. Kolthammer, X.-M. Jin, M. Barbieri, A. Datta, N. Thomas-Peter, N. K. Langford, D. Kundys *et al.*, Boson sampling on a photonic chip, Science **339**, 798 (2013).

[23] A. Crespi, R. Osellame, R. Ramponi, D. J. Brod, E. F. Galvão, N. Spagnolo, C. Vitelli, E. Maiorino, P. Mataloni, and F. Sciarrino, Integrated multimode interferometers with arbitrary designs for photonic boson sampling, Nat. Photon. **7**, 545 (2013).

[24] M. Tillmann, B. Dakić, R. Heilmann, S. Nolte, A. Szameit, and P. Walther, Experimental boson sampling, Nat. Photon. **7**, 540 (2013).

[25] H. Wang, J. Qin, X. Ding, M.-C. Chen, S. Chen, X. You, Y.-M. He, X. Jiang, L. You, Z. Wang *et al.*, Boson sampling with 20 input photons and a 60-mode interferometer in a 10^{14}-dimensional Hilbert space, Phys. Rev. Lett. **123**, 250503 (2019).

[26] R. Uppu, F. T. Pedersen, Y. Wang, C. T. Olesen, C. Papon, X. Zhou, L. Midolo, S. Scholz, A. D. Wieck, A. Ludwig, and P. Lodahl, Scalable integrated single-photon source, Sci. Adv. **6**, eabc8268 (2020).

[27] N. Tomm, A. Javadi, N. O. Antoniadis, D. Najer, M. C. Löbl, A. R. Korsch, R. Schott, S. R. Valentin, A. D. Wieck, A. Ludwig, and R. J. Warburton, A bright and fast source of coherent single photons, Nat. Nanotechnol. **16**, 399 (2021).

[28] V. Stojanović, R. J. Ram, M. Popović, S. Lin, S. Moazeni, M. Wade, C. Sun, L. Alloatti, A. Atabaki, F. Pavanello *et al.*, Monolithic silicon-photonic platforms in state-of-the-art CMOS SOI processes, Opt. Express **26**, 13106 (2018).

[29] J. Bao, Z. Fu, T. Pramanik, J. Mao, Y. Chi, Y. Cao, C. Zhai, Y. Mao, T. Dai, X. Chen *et al.*, Very-large-scale integrated quantum graph photonics, Nat. Photon. **17**, 573 (2023).

[30] H.-S. Zhong, H. Wang, Y.-H. Deng, M.-C. Chen, L.-C. Peng, Y.-H. Luo, J. Qin, D. Wu, X. Ding, Y. Hu *et al.*, Quantum computational advantage using photons, Science **370**, 1460 (2020).

[31] A. Arkhipov and G. Kuperberg, The bosonic birthday paradox, Geom. Topol. Monogr. **18**, 10 (2012).

[32] J. F. F. Bulmer, S. Paesani, R. S. Chadwick, and N. Quesada, Threshold detection statistics of bosonic states, Phys. Rev. A **106**, 043712 (2022).

[33] C.-K. Hong, Z.-Y. Ou, and L. Mandel, Measurement of subpicosecond time intervals between two photons by interference, Phys. Rev. Lett. **59**, 2044 (1987).

[34] A. L. Migdall, D. Branning, and S. Castelletto, Tailoring single-photon and multiphoton probabilities of a single-photon on-demand source, Phys. Rev. A **66**, 053805 (2002).

[35] S. Bartolucci, P. Birchall, D. Bonneau, H. Cable, M. Gimeno-Segovia, K. Kieling, N. Nickerson, T. Rudolph, and C. Sparrow, Switch networks for photonic fusion-based quantum computing, arXiv:2109.13760.

[36] P. Lodahl, A. Ludwig, and R. J. Warburton, A deterministic source of single photons, Phys. Today **75**(3), 44 (2022).

[37] K. R. Motes, A. Gilchrist, J. P. Dowling, and P. P. Rohde, Scalable boson sampling with time-bin encoding using a loop-based architecture, Phys. Rev. Lett. **113**, 120501 (2014).

[38] Y. He, X. Ding, Z.-E. Su, H.-L. Huang, J. Qin, C. Wang, S. Unsleber, C. Chen, H. Wang, Y.-M. He *et al.*, Time-bin-encoded boson sampling with a single-photon device, Phys. Rev. Lett. **118**, 190501 (2017).

[39] T. Hummel, C. Ouellet-Plamondon, E. Ugur, I. Kulkova, T. Lund-Hansen, M. A. Broome, R. Uppu, and P. Lodahl, Efficient demultiplexed single-photon source with a quantum dot coupled to a nanophotonic waveguide, Appl. Phys. Lett. **115**, 021102 (2019).

[40] S. Castelletto, Silicon carbide single-photon sources: challenges and prospects, Mater. Quantum Technol. **1**, 023001 (2021).

[41] G. Zhang, Y. Cheng, J.-P. Chou, and A. Gali, Material platforms for defect qubits and single-photon emitters, Appl. Phys. Rev. **7** (2020).

[42] M. Kianinia, Z.-Q. Xu, M. Toth, and I. Aharonovich, Quantum emitters in 2D materials: Emitter engineering, photophysics, and integration in photonic nanostructures, Appl. Phys. Rev. **9**, 011306 (2022).

[43] L. Zhai, G. N. Nguyen, C. Spinnler, J. Ritzmann, M. C. Löbl, A. D. Wieck, A. Ludwig, A. Javadi, and R. J. Warburton, Quantum interference of identical photons from remote gaas quantum dots, Nat. Nanotechnol. **17**, 829 (2022).

[44] C. Papon, Y. Wang, R. Uppu, S. Scholz, A. D. Wieck, A. Ludwig, P. Lodahl, and L. Midolo, Independent operation of two waveguide-integrated quantum emitters, Phys. Rev. Appl. **19**, L061003 (2023).

[45] S. Aaronson and D. J. Brod, Bosonsampling with lost photons, Phys. Rev. A **93**, 012335 (2016).

[46] J. Carolan, C. Harrold, C. Sparrow, E. Martín-López, N. J. Russell, J. W. Silverstone, P. J. Shadbolt, N. Matsuda, M. Oguma, M. Itoh *et al.*, Universal linear optics, Science **349**, 711 (2015).

[47] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, Experimental realization of any discrete unitary operator, Phys. Rev. Lett. **73**, 58 (1994).

[48] W. R. Clements, P. C. Humphreys, B. J. Metcalf, W. S. Kolthammer, and I. A. Walmsley, Optimal design for universal multiport interferometers, Optica **3**, 1460 (2016).

[49] A. Bouland, D. Brod, I. Datta, B. Fefferman, D. Grier, F. Hernandez, and M. Oszmaniec, Complexity-theoretic foundations of bosonsampling with a linear number of modes, arXiv:2312.00286.

[50] H. Qi, L. G. Helt, D. Su, Z. Vernon, and K. Brádler, Linear multiport photonic interferometers: loss analysis of temporally-encoded architectures, arXiv:1812.07015.

[51] L. Carosini, V. Oddi, F. Giorgino, L. M. Hansen, B. Seron, S. Piacentini, T. Guggemos, I. Agresti, J. C. Loredo, and P. Walther, Programmable multi-photon quantum interference in a single spatial mode, arXiv:2305.11157.

[52] M. Ferrante and N. Frigo, A note on the coupon-collector's problem with multiple arrivals and the random sampling, arXiv:1209.2667.

[53] X. Ding, Y. He, Z.-C. Duan, N. Gregersen, M.-C. Chen, S. Unsleber, S. Maier, C. Schneider, M. Kamp, S. Höfling, C.-Y. Lu, and J.-W. Pan, On-demand single photons with high extraction efficiency and near-unity indistinguishability from a

resonantly driven quantum dot in a micropillar, Phys. Rev. Lett. **116**, 020401 (2016).

[54] C. L. Dreeßen, C. Ouellet-Plamondon, P. Tighineanu, X. Zhou, L. Midolo, A. S. Sørensen, and P. Lodahl, Suppressing phonon decoherence of high performance single-photon sources in nanophotonic waveguides, Quantum Sci. Technol. **4**, 015003 (2018).

[55] C. Taballione, M. C. Anguita, M. de Goede, P. Venderbosch, B. Kassenberg, H. Snijders, N. Kannan, W. L. Vleeshouwers, D. Smith, J. P. Epping *et al.*, 20-mode universal quantum photonic processor, Quantum **7**, 1071 (2023).

[56] S. Chen, L.-C. Peng, Y.-P. Guo, X.-M. Gu, X. Ding, R.-Z. Liu, X. You, J. Qin, Y.-F. Wang, Y.-M. He, J. J. Renema, Y.-H. Huo, H. Wang, C.-Y. Lu, and J.-W. Pan, Heralded three-photon entanglement from a single-photon source on a photonic chip, Phys. Rev. Lett. **132**, 130603 (2024).

[57] Y. Wang, C. F. D. Faurby, F. Ruf, P. I. Sund, K. Nielsen, N. Volet, M. J. R. Heck, N. Bart, A. D. Wieck, A. Ludwig *et al.*, Deterministic photon source interfaced with a programmable silicon-nitride integrated circuit, npj Quantum Inf. **9**, 94 (2023).

[58] F. T. Pedersen, Deterministic single and multi-photon sources with quantum dots in planar nanostructures, Ph.D. thesis, University of Copenhagen, 2020.

[59] T. G. Tiecke, K. P. Nayak, J. D. Thompson, T. Peyronel, N. P. de Leon, V. Vuletić, and M. D. Lukin, Efficient fiber-optical interface for nanophotonic devices, Optica **2**, 70 (2015).

[60] J. Notaros, F. Pavanello, M. T. Wade, C. M. Gentry, A. Atabaki, L. Alloatti, R. J. Ram, and M. A. Popović, Ultra-efficient cmos fiber-to-chip grating couplers, in *2016 Optical Fiber Communications Conference and Exhibition (OFC)* (IEEE, Piscataway, NJ, 2016), pp. 1–3.

[61] R. Marchetti, C. Lacava, L. Carroll, K. Gradkowski, and P. Minzioni, Coupling strategies for silicon photonics integrated chips, Photon. Res. **7**, 201 (2019).

[62] V. S. Shchesnovich, Noise in boson sampling and the threshold of efficient classical simulatability, Phys. Rev. A **100**, 012340 (2019).