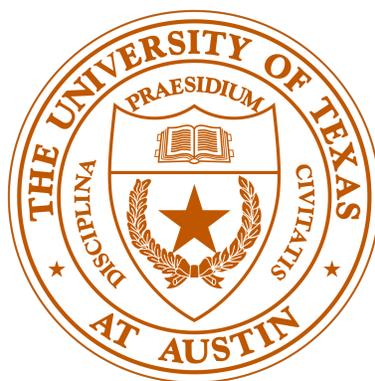


Peaked Boson Sampling: towards efficiently verifiable and NISQ-able quantum advantage

A Thesis Submitted
to the University of Texas at Austin
for the Turing Scholars program

by

MICHELLE DING



to

**Department of Computer Science
College of Natural Sciences
University of Texas at Austin**

May 2025

ACKNOWLEDGEMENT

I am sincerely grateful to Scott Aaronson and Nick Hunter-Jones for their time and advice which has been invaluable to the progression of this project. I am thankful to Yuxuan Zhang (postdoc and UT alumni) for sharing his feedback on the experimental feasibility of my simulations as well as ideas from a previously related project. I would finally like to thank the graduate students from the UT Quantum Information Center for allowing me to rehearse my ideas with them.

ABSTRACT

The use of randomness in quantum circuits is an intrinsically interesting property due to the model's convergence to the hard-to-simulate Haar measure. While random circuit sampling is both promising for quantum advantage and realistically implementable on a NISQ device, it is not yet efficiently verifiable. Previous work in studying peaked random circuit sampling models has shown optimism for a potentially viable model [2]. In this paper, we extend those observations by studying a simpler, alternative model to quantum computation involving beamsplitter networks. We present numerical and theoretical findings on the structure of peaked beamsplitter networks and evaluate their potential as a candidate for quantum advantage experiments.

1 Introduction

1.1 Overview

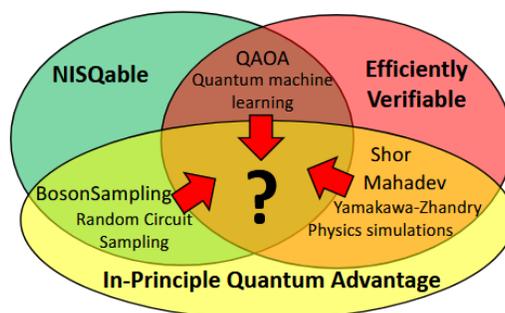


Figure 1: Landscape of quantum advantage proposals [2]

Quantum computing is a field centered on solving tasks based on the principles of quantum mechanics. Within this landscape, ideas like random circuit sampling have garnered public interest, in part due to Google’s experiment on benchmarking fidelity with linear cross-entropy (LXEB) scores on their 53-qubit Sycamore processor [5]. While this task claims to show a quantum speedup over classical supercomputers, the proposal lacks an efficient classical method for certifying the output distribution. The importance of having a clear method of verification motivates a discussion on what constitutes a convincing demonstration of quantum advantage.

The first desired property is that algorithms exhibit **in-principle quantum advantage**. At a high level, quantum advantage is the idea that quantum computers can outperform classical ones at certain tasks. Usually, the task is identified with a complexity-theoretic assumption such as the unlikely collapse of the polynomial

hierarchy [1] or the average-case hardness of highly entangled random circuits to $\#P$ -hard as they approach Haar-random states in exponential depth [7]. Consequently, discovering the existence of a classical algorithm that samples from this distribution efficiently would result in foundational revisions to complexity-theoretic assumptions. Examples of tasks in this region include BosonSampling and random circuit sampling.

It is also known that quantum hardware is inherently error-prone due to the inability of quantum systems to preserve the delicate state of quantum information (QI) such as qubits in entanglement or superposition. In order to prevent most demonstrations from breaking down at larger systems, we should look for algorithms that are feasible on **noisy, intermediate-scale quantum (NISQ) devices**. An example of such an algorithm is the linear optical model of BosonSampling, which can be implemented in optics laboratories on near-term devices.

Finally, we should specify that our algorithm is **efficiently verifiable**. This enables a classical verifier to check the outputs of the quantum computation without running the algorithm itself. More concretely, assume a classical verifier \mathcal{A} is handed a quantum algorithm \mathcal{Q} from some quantum party \mathcal{B} , for which \mathcal{B} may claim \mathcal{Q} outperforms classical computers at certain tasks. We then demand that \mathcal{A} should be able to check certain properties of the output distribution of \mathcal{Q} to ensure that \mathcal{B} is being “honest” with their description. One type of construction that falls into this category is peaked circuits.

1.2 Challenges

While these tasks are well-defined, achieving a convincing demonstration of quantum advantage remains difficult due to most demonstrations either breaking down at

larger systems or being hard to verify. In the following, we briefly describe two strong candidates for quantum advantage that fall short of a convincing demonstration due to at least one of these issues.

As mentioned previously, while Google’s experiment is efficiently implementable on a linear optical platform, the proposal lacks a clear method for classically verifying the LXEB scores of the quantum outputs.

Shor’s algorithm delivers an exponential speedup for solving the problem of factoring large numbers, which was thought to be hard to recover due to being hard to factor classically. This makes it a candidate for quantum advantage which realistically threatens to break the cryptography upon which credit card security lies. However, QI is highly fragile: quantum noise gives rise to many categorizations of errors (bit flip, phase flip, dephasing or depolarizing noise) that develop in various settings (prolonged interaction of qubits with the environment, frequent photon loss in linear optical regimes). Because of this, a physical implementation of Shor’s algorithm would require a circuit consisting of roughly 1,400 qubits and on the order of millions of gates [13]. This implementation is infeasible given current quantum systems can only sustain coherent computation across a few hundred qubits before error correction fails to reverse the noise.

1.3 Contributions

Does there exist an algorithm that lies in the intersection of all three criteria of being hard to simulate classically, efficiently verifiable, and realizable on near-term quantum devices? Proof of such an algorithm would yield the first ever quantum speedup

to be physically realized on a real near-term quantum device. While BosonSampling provides evidence for in-principle quantum advantage but lacks the property of efficient verification, peaked circuits offer the ability to verify outputs but are known to be easy to simulate classically. Combining these tasks raises the natural question of whether it is possible to generate peaked but hard-to-sample from distributions. Our model combines the linear optical setup of BosonSampling with the efficiently verifiable properties of peaked circuit sampling to address this question.

2 Background

2.1 The standard circuit model

The principle model of quantum computation is the standard circuit model, which has three components: input states called qubits, unitary gates governing the state evolution, and measurement. In classical circuits, input bits (0s and 1s) are fed into wires with gates to produce some output bit(s). Similar to classical circuits, one can imagine a quantum circuit as a framework in which input bits ($|0\rangle$ and $|1\rangle$) travel down “wires” that time-evolve the state. Importantly, the gates U that are applied along the way result in non-classical interactions.

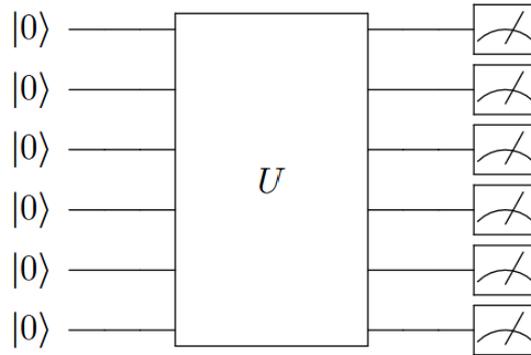


Figure 2: The standard circuit model

Gate evolution governs various quantum-specific properties of the state such as operator growth, entanglement, and superposition. Measurement collapses the state into a single classical outcome. Although we are allowed to make intermediate measurements, the principle of deferred measurement tells us that adding extra auxiliary qubits and controlled-NOT gates to our circuit and measuring at the end can simulate an equivalent circuit. While the effect of collapsing a state upon measurement is inevitable, there are ways to mitigate this information loss. For example, one might consider using partial measurements to control the spread of entanglement through the system [15] or apply the gentle measurement lemma which tells us that it is possible to minimally disturb a state with a highly likely measurement outcome upon measuring in a basis very close to it.

2.2 Sampling

The task for sampling problems is to take as input an (efficient) classical description of a circuit and generate samples from its output distribution. For instance, we may

be given a n -qubit quantum circuit U specified by $U = U_1 \cdots U_m$ in which each U_i is a 2-local gate ($U_i = I \otimes I \otimes \cdots \otimes V_i \otimes I \otimes I$) where V_i acts on two specified qubits. This may be the case if we are given a quantum circuit whose gates are 4×4 unitaries chosen randomly as in random circuit sampling. An ideal noiseless quantum computer will sample $x \in \{0, 1\}^n$ according to $\Pr_{|\psi\rangle}(x) = |\langle x | \psi \rangle|^2 = |\langle x | U | 0^n \rangle|^2$. Of course, a quantum algorithm can exactly sample from this probability distribution by simply running the circuit. A more interesting problem is examining the ability for classical algorithms to sample, approximately or exactly, from quantum distributions.

2.3 Peaked Circuit Sampling

The task of peaked circuit sampling is to sample from a peaked distribution. We say a distribution is *peaked* if $\max_{s \in \{0, 1\}^n} |\langle s | C | 0^n \rangle|^2 \geq \delta$ for $\delta \in [0, 1]$. Let the state corresponding to the maximum peaked weight be x . Since the distribution is peaked, the probability of finding a collision on outcome x by sampling from the output distribution polynomially many times is high, so with high probability one can recover the peaked string x . Thus, giving x to a classical verifier would be a way to efficiently verify the outputs of the quantum computation even if the classical verifier cannot find an efficient algorithm that produces samples approximating the quantum distribution.

2.4 BosonSampling

An alternative to the standard circuit model of qubit-based quantum computation is the noninteraction-boson model in which a problem called BosonSampling resides.

Bosons are a class of particles that can be physically realized as photons. In the laboratory, BosonSampling can be implemented on a linear-optical platform in which m modes host n identical photons which pass through a set of local optical elements called beamsplitter and phaseshifter gates, to be measured at the end of the network to determine their exact locations [1]. In comparison to the input states characterized by qubits from the standard model, BosonSampling uses m modes as input which represents different locations a boson can occupy in an m -dimensional Hilbert space.

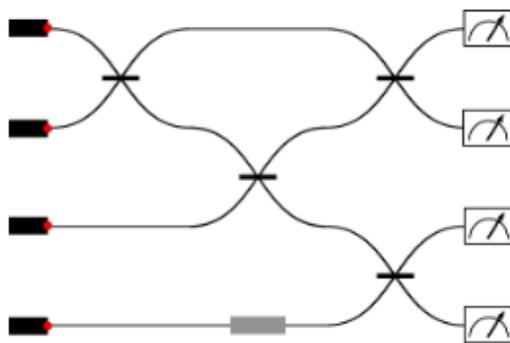


Figure 3: Illustration of a beamsplitter and phaseshifter network from [2]

Expanding off Figure 3, we can think of each mode as an initial slot labeled $i \in [m]$ with some nonnegative integer s_i denoting the number of photons in that mode. In particular, for this diagram $m = 4$. For n photons, we have that $s_1 + s_2 + \dots + s_m = n$. Applying a direct counting argument for n indistinguishable photons in m distinguishable modes gives $|\Phi_{m,n}| = \binom{m+n-1}{n}$ basis states, which is equal to the size of the Hilbert space. Any $m \times m$ unitary can be applied to this single photon state with m modes. Additionally, any unitary transformation on m modes can be decomposed into a product of beamsplitters and phaseshifters, which formally act

with the non-identity operation on two modes and one mode respectively.

$$B_{M \times M} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & & \\ \sin(\theta) & -\cos(\theta) & & \\ & & 1 & \\ & & & 1 \end{pmatrix}, \quad P_{M \times M} = \begin{pmatrix} e^{i\theta} & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix}$$

In the single-photon case, the probability of recovering the photon in some mode i is $|U_{i,i}|^2$: a function of a singular entry in the permanent of a BosonSampling matrix U . We can generalize this single-photon input state to the n -photon *standard initial state*: define $|1_n\rangle := |1, \dots, 1, 0, \dots, 0\rangle$ to be the state with one photon in each of the first n modes and the remaining $m - n$ modes unoccupied. Because only the first n columns of any unitary act on the input state, the gate complexity of this setup can be reduced from $O(m^2)$ to $O(mn)$, which can further be improved to have $O(n \log m)$ depth by parallelization (Theorem 6.1 of [2]). Correspondingly, we denote with $M := \binom{m+n-1}{n}$ the size of the larger Hilbert space.

This gives rise to a natural homomorphism that lifts from the $m \times m$ unitary to the $M \times M$ unitary [2]. Each entry of this larger unitary is a permanent of a smaller $m \times m$ unitary from the original unlifted space. Notably, while BosonSampling is suspected to be far from universal for quantum computation, the homomorphism discovered by Aaronson and Arkhipov reduces to a permanent calculation which is in principle $\#P$ -hard for classical computation by the *Permanent-of-Gaussians* conjecture.

Conjecture 1 (Permanent-of-Gaussians [2]). *Given $X \sim \mathcal{N}(0, 1)_{\mathbb{C}}^{n \times n}$ where $\mathcal{N}(0, 1)$ is the Gaussian distribution with $\mu \sim 0$ and $\sigma \sim 1$, estimating $|\text{Per}(X)|^2$ to within*

$\pm\epsilon \cdot n!$ is $\#\text{P}$ -hard.

At a high level, this says that we expect performing certain computations with beamsplitter networks can give a quantum advantage over classical computations. It is also worth noting that the Hilbert space dimension of qubit-based computation scales with tensor products while the Hilbert space dimension in the noninteracting-boson model of computation scales with the direct product. This is another property in addition to the linear optical context which makes BosonSampling an easier model to study. Thus, we suspect that investigating peaked boson models, i.e. random beamsplitter networks with specific peaked properties, may shed light on new ways to achieve quantum advantage or at least simplify the terrain for conducting numerical experiments.

2.5 Searching for Structure

We previously posited the question of whether it is possible to generate peaked but hard-to-sample from distributions. One method to build such distributions is to design circuits that encode hard-to-recover peaking structure through statistical independence. While exact Haar-random unitaries guarantee full statistical independence (i.e. all moments match that of the Haar measure), they are only efficient to generate numerically and impractical for most real experiments in requiring exponential resources (i.e. form with exponential circuit depth). A more practical alternative is using circuits that form approximate unitary t -designs, which approximate the Haar distribution up to the t th moment and are known to emerge at polynomial depth in generic random circuits [9]. However, it is not yet known if t -designs can arise

efficiently in linear optical settings, or whether such designs can coexist with output distributions exhibiting peaking behavior.

A more grounded question, then, is how postselecting from random instances affects the statistical properties of standard BosonSampling. While approximate t -designs are known to suppress peaking by closely matching the low-order moments of the Haar distribution, postselecting on rare, high probability outcomes may yield an ensemble of circuits with a modified set of moment distributions. This raises the question of whether such postselected ensembles can retain properties of random networks such as anticoncentration or k -wise indistinguishability, despite being highly structured [12]. We further discuss the overlap between indistinguishability and structure in near-term quantum advantage proposals in our Conclusion section.

Our final question is how studying explicitly peaked structures, which are structures that are easy to construct but also easy to spoof classically, can help us understand how peakedness emerges. To this end, we use tools such as stochastic gradient descent and quantum state learning to optimize over the peaking parameters of the distribution. It is worth noting that from a quantum advantage perspective, naturally postselected peaked states may be more realistic given they may still retain hardness properties from the computational complexity of BosonSampling. Here, the cost of identifying peaked circuits would likely shift the computational burden to preprocessing, rendering the overall protocol inefficient despite the potential efficiency of the final circuit. Resolving this issue is addressed in recent proposals such as Complement Sampling [6], which suggests using cryptographic constructions like S-AES to optimize their protocol.

3 Related Work

The search for quantum advantage through specialized circuit architectures has drawn significant attention in recent years. A notable contribution in this domain is the recent work by Aaronson and Zhang [2], who conduct experiments on peaked circuit sampling, which we recall as the task of sampling from output distributions with a high concentration on specific computational basis states.

A portion of our work is extending the results from Aaronson and Zhang to the linear optical regime. In particular, while the code is entirely our own, we have also replicated many of the experimental results from this paper. For instance, we can confirm the results of an explicitly-peaked output distribution for $\tau_r = 40$ random gates and $\tau_p = 10$ peaking gates,

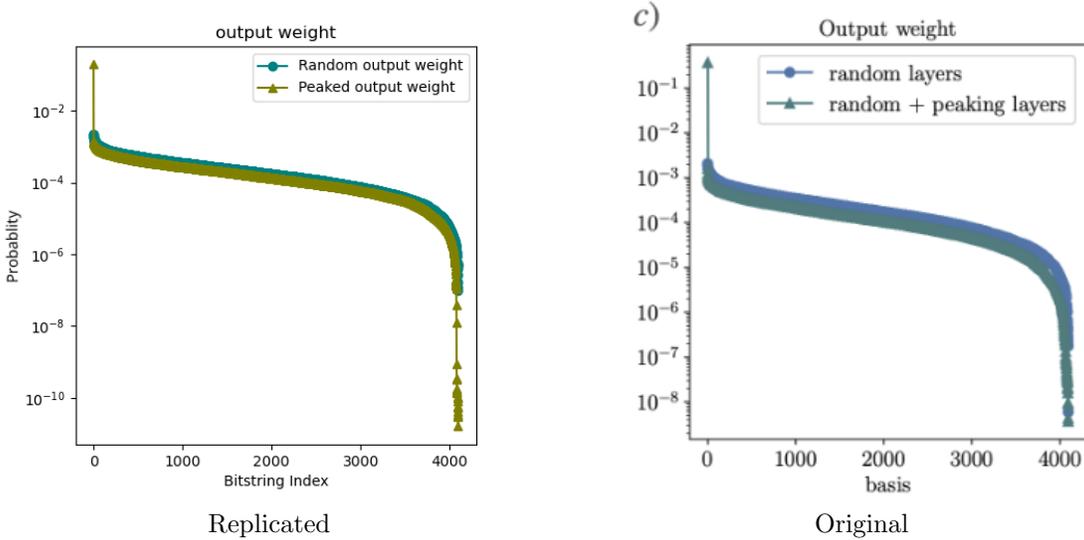


Figure 4: Peaked circuits resemble the random Porter-Thomas distribution.

Their study demonstrates that peaked circuits enable efficient verification through

the classical detection of peaked output states while still maintaining analytical and numerical evidence supporting nontrivial peaking structure such as the ability to invert beyond what is expected as compared to purely random behavior.

Complementing this work are experiments from IBM on the classical simulation of peaked shallow quantum circuits [8]. Notably, there exist classical algorithms capable of approximately sampling from output distributions of peaked circuits with constant depth. This indicates that while simulating peaked circuits may be challenging in the general case, enforcing certain structural constraints may render them more tractable for an efficient classical simulation.

In the applied setting, BlueQubit recently hosted a [Quantum Hackathon](#) providing a platform for over 400 contestants to solve instances of peaked circuits. Contestants were tasked to identify the peaked output state from circuits generated with obfuscation techniques using both classical and quantum resources. The best algorithm (i.e. highest overall accuracy and efficiency) utilized tensor network contraction techniques, a circuit optimization strategy employed by Aaronson and Zhang in studying peaked circuit sampling [2]. This competition highlighted the challenges associated with reverse-engineering peaked circuits and the potential of tensor network contraction methods to efficiently recover parts of these distributions. While exploratory in nature, the hackathon was beneficial for understanding what kinds of peaked structures may be hard to simulate classically and hence have the potential for demonstrating a quantum speedup.

Our work incorporates these ideas by attempting to better understand the naturally formed and explicitly peaked structures that emerge from linear optics. In particular, we ask whether postselection and explicit circuit design can give rise to

a hard-to-simulate peaked output distribution in the linear optical regime, and aim to recover the conditions under which such structured peaking emerges.

4 Methodology

4.1 Analytical setup

Definition 4.1 (Peaked linear optics network). Given $\delta \in (0, 1]$, we call the unitary U δ -peaked if:

$$\max_{S \in \Phi_{m,n}} |\langle S | \phi(U) | 1_n \rangle|^2 \geq \delta$$

with a corresponding peak weight $\delta_S \equiv |\langle S | \phi(U) | 1_n \rangle|^2$.

Just by defining peakedness, we already get very interesting properties. For instance,

Theorem 4.2 (Hamming weight concentration for peaked shallow circuits [8]). *Let U be a peaked shallow circuit with output distribution P . Suppose we choose a local basis so that $\mathbb{E}[x_j] \leq \frac{1}{2}$ for every qubit $j \in \{1, \dots, n\}$. Then*

$$\Pr_{x \sim P}[|x| \geq O(\log n)] \leq \frac{1}{\text{poly}(n)}$$

For comparison, these results without peakedness are

Theorem 4.3 (Hamming weight concentration for shallow circuits [4]). *Let U be a*

shallow circuit with output distribution P . Then for every $t > 0$,

$$\Pr_{x \sim P}[||x| - \text{medianHW}(P)| \geq t \cdot O(\sqrt{n})] \leq e^{-t^2}$$

The first theorem tells us that there is a concrete classical algorithm for peaked shallow quantum circuits formed from approximate sampling with noise. Taking the contrapositive gives us a useful characterization stating that the circuits used to demonstrate quantum advantage must not be shallow.

We now transition to the linear optical regime, where we highlight specific interferometer architectures that are known to analytically produce peaked states. For these models, we also quantify the degree of peakedness by analyzing their associated peak weights.

Claim 1. *Assume input Fock states are of the form $|1_n\rangle$. Consider a linear optical network consisting of a fixed random interferometer C_r followed by a variable peaking region C_p . Then C_p must have $O(n \log m)$ -depth in order to produce an optimal peaked value of $\delta = 1$.*

Proof. Let C'_r be the decomposition of C_r into local beamsplitters and phaseshifters gates, which can be done efficiently with a Reck [14] or Clements [11] construction occurring in $O(m^2)$. Invert each gate in $O(1)$ so that $C'_r = C_r^{-1}$ and set the peaking circuit to be $C_p = C_r^{-1}$. We recover an identity circuit which is peaked by definition. The overall procedure has $O(n \log m)$ depth by Theorem 6.1 of [2]. \square

Historically, this decomposition has been done with both $O(m^2)$ optical elements and depth by constructing a network of all-to-all beamsplitters followed by a layer

of m phaseshifters [14] and can be improved to $O(mn)$ by computing only over the first n columns of the unitary matrix. It can subsequently be lowered to $O(n \log m)$ depth via a parallelization technique [1]. The unitary transformation V that acts on this circuit can be constructed as follows:

$$\underbrace{\dots}_{2^{\log_2 m - 1}} \cdots \underbrace{(B_{4,8} B_{3,7} B_{2,6} B_{1,5})}_4 \underbrace{(B_{2,4} B_{1,3})}_2 \underbrace{(B_{1,2})}_1$$

which gives $O(\log m)$ depth and $1 + 2 + 2^2 + \dots + 2^{\log_2 m - 1} = O(m)$ gates. This tells us that with at least $O(n \log m)$ peaking gates, we can always recover a fully-peaked circuit. However, we can also consider circuits producing non-optimal peakedness that have less than $O(n \log m)$ peaking gates.

Claim 2. *A network of $O(t)$ nonlocal peaking gates $BS_{\theta_i, \phi_i}(1, i)$ acting on the first mode and mode i for $i \in [2, t]$ can transfer all amplitudes from the bottom $t - 1$ modes to the first mode.*

Proof. We provide a construction. Let $|\psi\rangle$ be the state outputted from a m -mode random beamsplitter network C_r . Label the output amplitudes a_i for $i \in [m]$. Then setting the peaking circuit to be

$$C_p = \prod_{i=1}^m B_{1,i} |\psi\rangle$$

results in the aforementioned amplitude on mode 1, where

$$B(\theta, \phi) = \begin{bmatrix} \cos(\theta) & -e^{-i\phi} \sin(\theta) \\ e^{i\phi} \sin(\theta) & \cos(\theta) \end{bmatrix}$$

$$\Rightarrow B_{1,i} = \frac{|a_1|}{\sqrt{|a_1|^2 + |a_i|^2}} \begin{bmatrix} 1 & -\frac{a_2^*}{a_1^*} \\ -\frac{a_2}{a_1} & 1 \end{bmatrix}$$

is the beamsplitter transformation that transfers amplitude from i to 1: mode i now has no probability associated with it while mode 1 updates with probability $\sqrt{|a_1|^2 + |a_i|^2}$. Solving gives $\theta = \cos^{-1}\left(\frac{|a_1|}{\sqrt{|a_1|^2 + |a_i|^2}}\right)$ and $\phi = \theta_i - \theta_1 + \frac{\pi}{2}$ for amplitudes in exponential form $a_j = r_j e^{j\theta_j}$. \square

While both these constructions are efficiently verifiable (being peaked) and efficiently constructable (being a result of the decomposition described above), the distributions can be easily learned by a classical algorithm given that the structure of these circuits is exploitable. We use this model as a basis for further numerical exploration rather than as a candidate for quantum advantage. In particular, we perform stochastic gradient descent on the nonlocal peaking layers of our explicit circuit construction to find the thresholds at which we can recover mostly peaked distributions.

4.2 Experimental setup

All experiments were performed using Strawberry Fields, a Python library for simulating photonic quantum circuits. We focus on beamsplitter networks with a single photon, as results in this regime automatically generalize to the n -photon case via the homomorphism introduced by Aaronson and Arkhipov [1].

We begin by constructing interferometers initialized from one of three network structures: (1) brickwall structure, an efficiently parallelizable network of alternating vertical horizontal beamsplitter layers; (2) pollman structure, a network with

staggered gate placements; and (3) Haar-random interferometers, generated by QR decompositions of complex Gaussian matrices. These initial interferometer settings help us determine the extent to which variables such as randomness and individual gate placement contribute to an overall peaked distribution.

Each random interferometer network is augmented with a series of peaking layers, which are parameterized beamsplitter gates designed to concentrate output probability on a specific Fock state. We optimize these layers using stochastic gradient descent (SGD) across two objectives. The first objective minimizes the fidelity between the random output state and the parameterized peaking state. The second directly maximizes the squared amplitude of a chosen output mode, pulling the system toward a single-modal distribution. These methods can be used for threshold analysis with full control over peaking in linear optical settings.

In addition to explicitly peaked structures, we also examine the performance of postselected circuits using two key metrics: Shannon entropy and collision probability. Shannon entropy is defined as $H(x) = -\sum_{i=1}^n p(x_i) \log p(x_i)$. This quantifies the unpredictability of seeing a potential state in the output distribution, with lower entropy indicative of more biased or nonuniform distributions. The collision probability is defined as $\pi_C := \sum_s p_C[s]^2$ for some circuit C , where p_C is the associated probability distribution over basis states of C . This measures the likelihood that two independent samples drawn from the same distribution yield the same outcome. These metrics are applied to ensemble averages and to individual circuit realizations for both global and local statistical analysis of circuit behavior.

In summary, our experimental setup combines random and structured linear optical circuit generation, gradient-based optimization, and statistical characterization

to investigate the conditions under which peaking can emerge in linear optics. This framework also serves as a testbed for future investigations into how peaking behavior scales with the number of modes, and whether such scaling may offer a viable path toward demonstrating quantum advantage.

5 Results

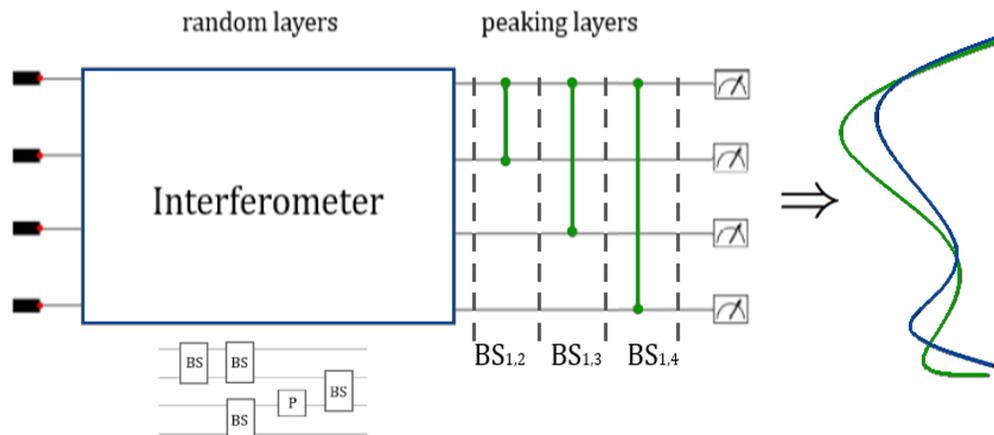


Figure 5: *Explicit constructions of peaking circuits.* An interferometer composed of beamsplitter and phaseshifter gates is followed by $O(m)$ peaked nonlocal beamsplitter gates optimized using a numerical solver.

We begin by implementing explicit peaking circuits as shown in Figure 5. The circuits are composed of a standard interferometer (a sequence of beamsplitter and phase-shifter gates) followed by $O(m)$ numerically optimized nonlocal peaking gates based on the construction from 2. The goal of this construction is to maximize the

amplitude of a specific Fock state, which we can assign wlog to the first mode, thus creating a peaked output distribution.

To check the performance of our optimization, we compare the theoretical values derived from 2 against our simulation results. For example, with m peaking layers we should expect to get a peaked weight of $\delta = 1$ with all amplitude transferred to the first mode i.e. $|\langle 1_n | \phi(U) | 1_n \rangle|^2 = \sqrt{|a_1|^2 + \dots + |a_m|^2} = 1$. Getting lower than this δ -value indicates that we are not reaching the optimal peaking value, either due to stochastic gradient descent being constrained by the barren plateau issue or a non-optimal structure for our peaked circuit.

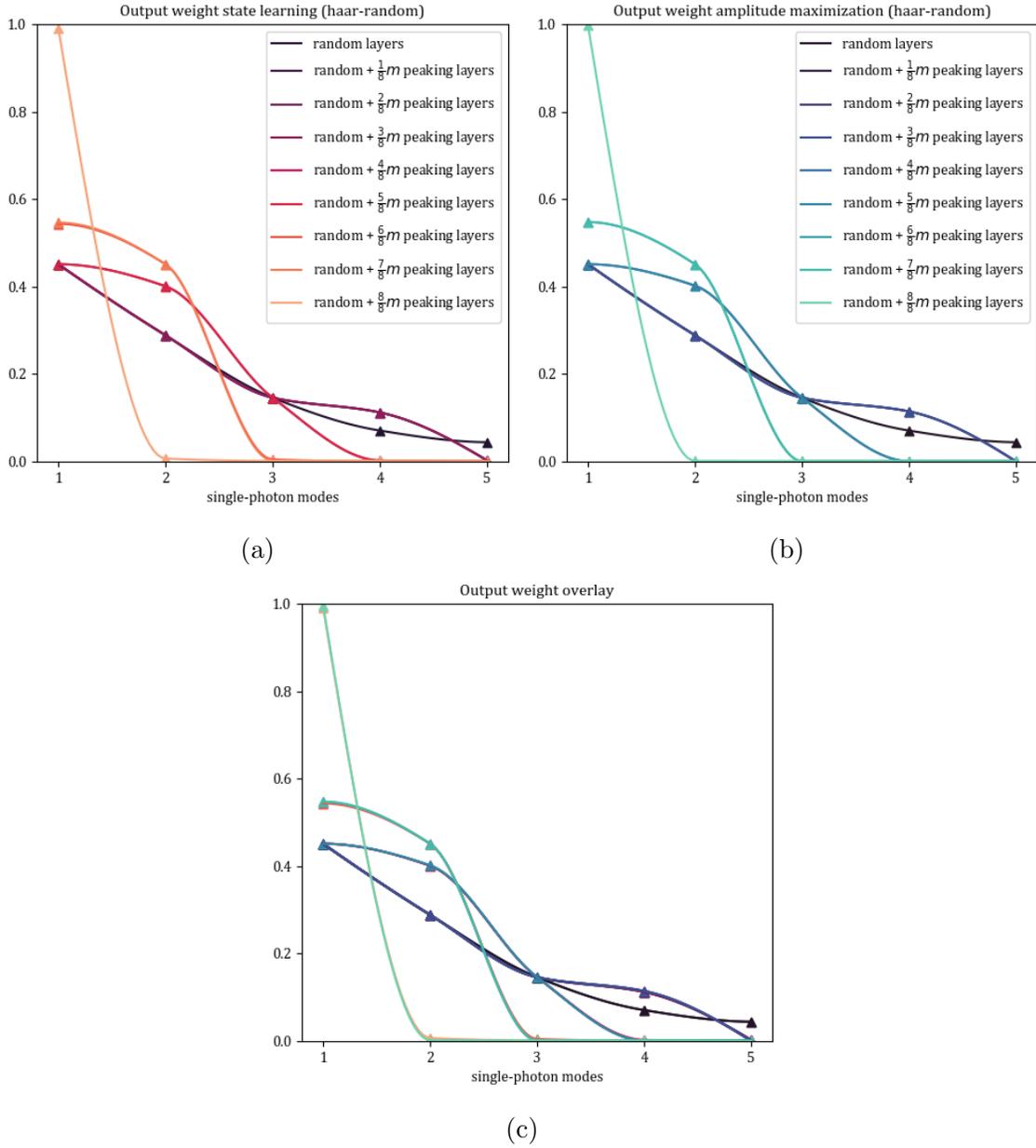


Figure 6: A numerical solver is optimized to learn the output state.

Figures 6a and 6b each show five (averaged) probability distributions of optimized

random networks over $m = 5$ modes. Our explicitly peaked linear optical networks are trained via gradient-based optimization over two distinct objectives based on the control flag `self.cft`, each corresponding to a different method of peaking. The architecture comprises an initial random layer followed by a trainable peaking layer composed of beamsplitter gates. Our parameters were defined with $m = 5$ modes, a cutoff dimension of 2 for single-photon experiments, and a random interferometer sampled according to the Haar measure using QR-based decomposition. The parameters of the peaking layer are initialized from a normal Gaussian distribution $\sim \mathcal{N}(0, 0.1)$ with mean 0 and standard deviation 0.1. They are subsequently run through a TensorFlow backend which adjusts the circuit parameters iteratively through gradient-based optimization. This is done with an Adam optimizer with learning rate 0.1, and the optimized gates are optionally inverted at the end in $O(1)$ to produce the final output distribution.

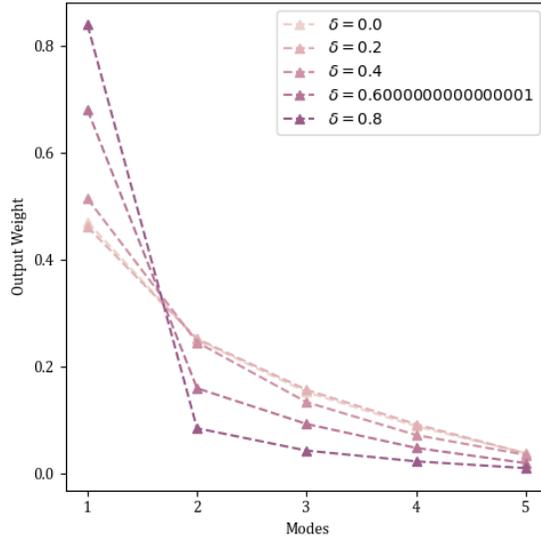
The first cost function uses quantum state learning. State learning refers to the use of variational quantum circuits (i.e. circuits with tunable gate parameters) to optimize the preparation of a state. To achieve this, we define a loss function based on fidelity, which quantifies the overlap between a pre-specified target state C_{tar} and the circuit’s final output state C_{out} . In this setting, we let C_{tar} be the state after applying random layers and C_{out} the state after applying both random and peaking layers. After converging to maximum fidelity $C_{fid} \approx 1$, we invert the layers of the peaking circuit to recover a highly peaked distribution on the first mode.

The second cost function aims to maximize the probability of a specific Fock basis state in wlog the first mode, greedily peaking the distribution for a single mode. We choose the following cost function, $\min ||\langle s|\phi(U)|1_n\rangle|^2 - 1|$ where $s = |1_n\rangle$ is the basis state with a single photon occupation on the first mode. By minimizing this cost

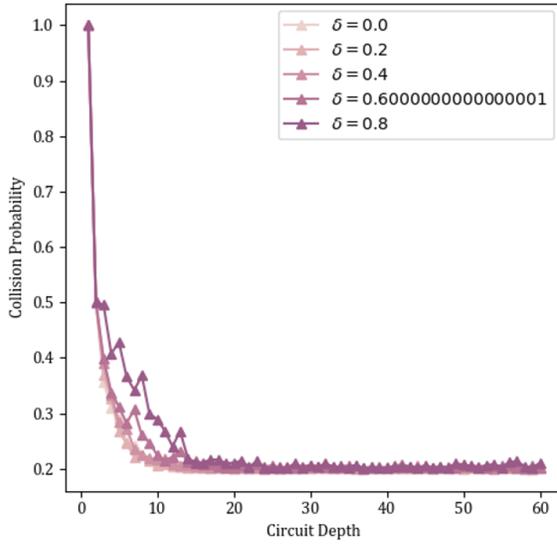
function, the quantum circuit will prepare a state with high probability on the first mode. Unlike the state-learning objective which matches the entire wavefunction to a target state, this approach focuses solely on concentrating amplitude in a chosen basis element of the Fock space. Note that no inversion is needed given we are directly optimizing over the probability of the output distribution.

For both cases, the optimization proceeds layer-by-layer, and the learned parameters are stored and later re-applied using the `run_circuit` subroutine. This execution includes post-optimization circuit reconstruction and optional gate sequence recording for analysis and reproducibility. Notably, final circuit realizations are evaluated for peakedness by checking whether the maximum output probability surpasses a defined threshold, at which point the gate sequence and output state are saved.

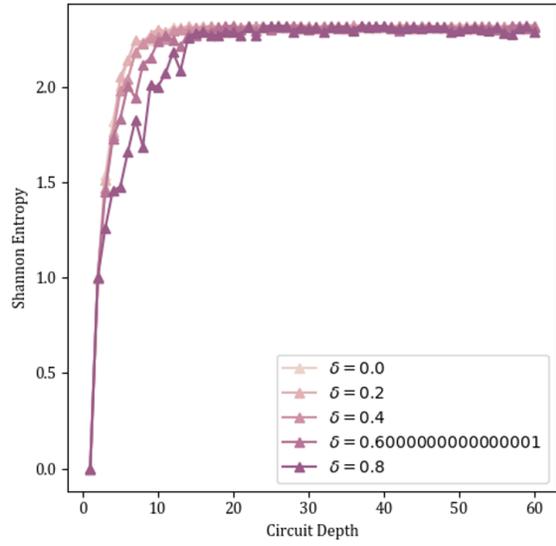
In Figure 6c, we compare both training objectives by overlaying the graphs. It is clear that both optimization techniques converge to nearly identical outcomes. At around $\frac{1}{2}m$ -depth, inversions begin to take effect. This shows that state learning is effective for single-photon modes. The high degree of overlap between both cost functions suggests that at least for small systems, convergence to the global optimum is possible in linear optics. Furthermore, we confirm these results support our theoretical findings. For example, we confirm that with m peaking gates we are able to obtain a maximally peaked weight on the first mode, as predicted by 2.



(a)



(b)



(c)

Figure 7: Peaked circuits do not scramble as quickly as Haar-random circuits.

We now transition to comparing Haar-random linear optical circuits with postse-

lected peaked circuits by tracking the evolution of collision probability and Shannon entropy over circuit depth. These metrics allow us to quantify scrambling and concentration of the output distribution.

In 7a, as higher thresholds of peakedness are enforced, we see the natural emergence of a single peak with all other probabilities decaying over time. Interestingly, the emergence of a single dominant peak (as thresholded peaking is enforced) occurs naturally even without explicit optimization, suggesting that unimodal peaking is dominant in structured random networks.

In 7b, we plot the collision probability of Haar-random linear optical circuits against various degrees of postselected peaked circuits as a function of circuit depth over 100 trials. These metrics show that peaked linear optical networks eventually converge to the same behavior as random optical ones. In the case of random circuits, the collision probability decreases with depth, reflecting the natural tendency of random circuits to spread amplitude across many outcomes. The behavior shown here is consistent with anticoncentration bounds proved by [12], which suggest that sufficiently deep random circuits produce nearly uniform output distributions with a constant factor with high probability.

Also shown in 7c is a plot of Shannon entropy as a function of circuit depth over 100 trials. We see that Shannon entropy converges to a value of $\log_2 5 \approx 2.3$, which is the maximum allowed entropy for a state. Given that peaked interferometers saturate to bounds reached by random interferometers as a function of depth, we can further ask what depth t -designs arise naturally in linear optical settings [10].

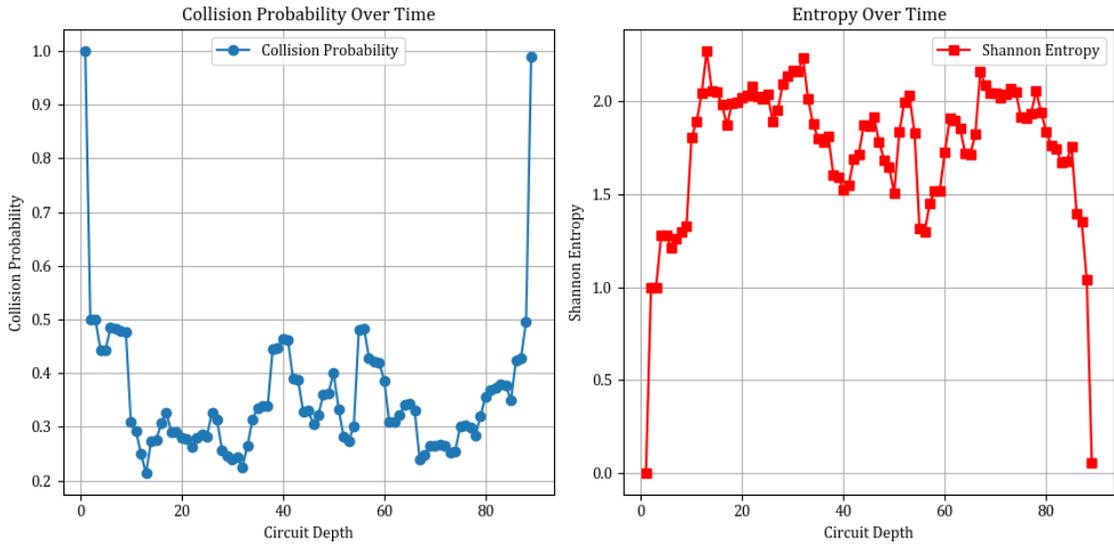


Figure 8: Single-run instance of numerically optimized peaked entropy.

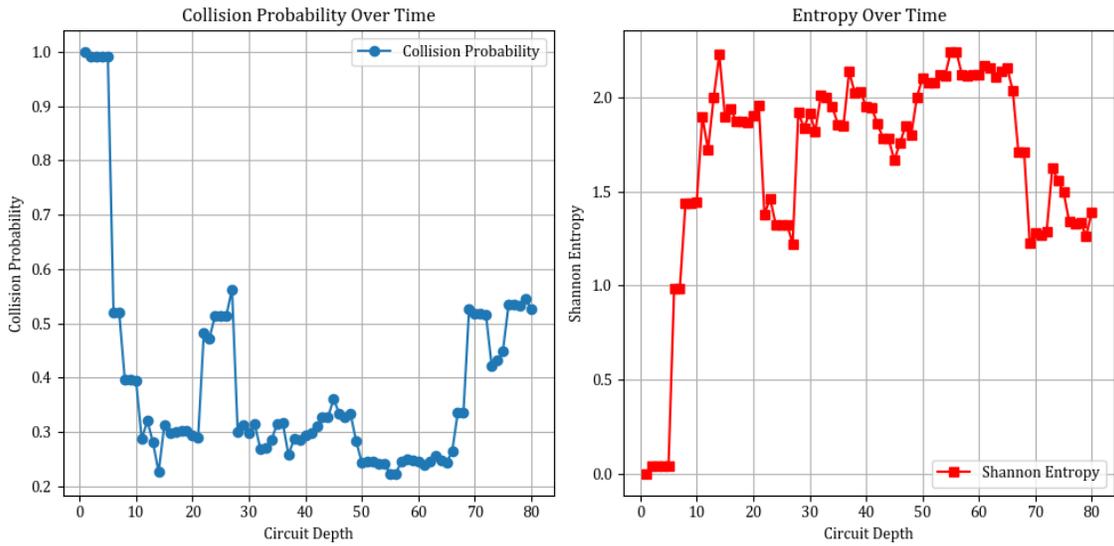


Figure 9: Single run instance of postselected peaked entropy.

In addition to examining circuit behavior when averaging over many trials, we

can also inspect individual interferometer networks. In Figures 8 and 10, we analyze single-shot instances of postselected peaked circuits, again as a function of circuit depth.

In order to analyze how structure emerges, we switch from generating QR-decomposition-based Haar-random unitaries to constructing a gate-by-gate network using pollman and brickwall network structures as shown in Figure 10. We see visually that some networks exhibit naturally occurring peaking behavior even in the absence of explicit optimization, which may suggest underlying interference structures in the gates themselves. Further exploration can be done in analyzing individual gate parameters.

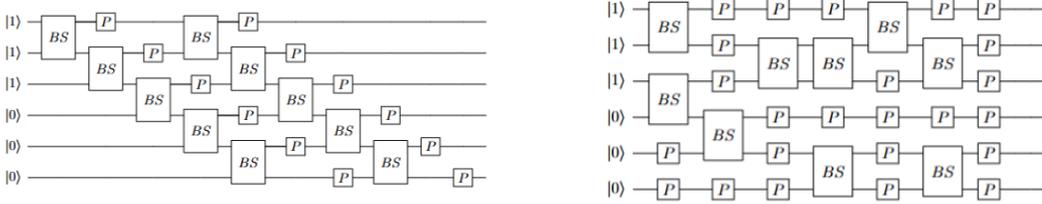


Figure 10: Pollman and brickwall networks used for single-shot experiments.

6 Conclusion

Over the duration of four months, we provide numerical and analytical findings on peaked boson sampling. Due to time constraints, there are many open questions and opportunities we leave as future work. First, we outline several of these.

While our results from section 5 show promise in peaked networks converging to the same statistical values as random linear optical networks, to make this rigorous,

we would also need to establish the hardness of simulating these circuits in a linear optical regime, in analogy to standard BosonSampling. As mentioned in 7a, we note for Haar-random states, the collision probability Z of $2^n \times 2^n$ unitary ensembles is known to approach $Z \sim \frac{2}{2^n+1}$ due to an anticoncentration result proved in 2022 [12]. Given that linear optical networks converge much faster in $O(m^2)$ time for the single-photon regime, we can naturally ask at what depth anticoncentration arises in linear optical settings.

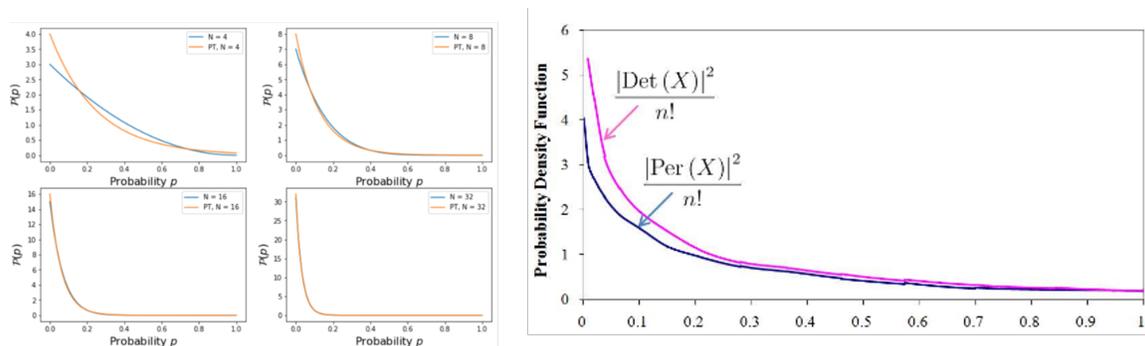


Figure 11: Output distribution for random circuits and interferometers

While the output distribution of random circuits is known to follow the Porter-Thomas distribution, i.e. $P(x) = Ne^{-Nx}$ for $N = 2^n$, we can also ask what kind of distribution is followed by the BosonSampling distribution, which we claim is given by the permanent-squared of random Gaussian matrices, see Figure 11. This is because Section 5 of Aaronson and Arkhipov’s paper [6] proves that truncations of Haar-random unitaries are close in TV distance to matrices of i.i.d Gaussians. Therefore, studying the output distribution of random linear optical networks amounts to studying the distribution of $\sim |\text{Per}(X)|^2$ where X is a random variable denoting a matrix with independently chosen random Gaussian entries.

Another potential experiment is to choose different gate sets to measure the ef-

fects on peaking. Orthogonal gates, which are real-valued unitary matrices with orthonormal columns, may for instance change the peaking landscape by creating potentially sharper-looking distributions due to their constrained geometry. Reflection matrices, similar to those used in the diffusion operator of Grover’s algorithm, perform amplitude amplification to increase the probability of measuring a target state. We can ask whether the generated circuits exhibit characteristics of Grover-like search algorithms and whether these structures arise more frequently when using constrained gate families.

One can also consider the impact of measurement-induced entanglement (MIE), previously shown to give rise to a quantum advantage in shallow-depth random circuits [16], in random linear optical networks. Since both platforms share features like non-adaptive measurements and unitary evolution, understanding whether shallow-depth beamsplitter networks exhibit similar entanglement transitions could provide a new framework for demonstrating quantum advantage without requiring deep or noisy circuits.

We are lastly interested in moving to a practical laboratory to conduct scalable physical experiments. A current bottleneck in the numerical investigation using Strawberry Fields is the inability to efficiently represent the Fock space in memory, which limits our study to 5-12 modes. However, recent papers on experimental BosonSampling have shown it is possible to scale physical systems to at least 100 modes [17]. One can envision that probing an interferometer model as described in 2 in the optics laboratory at the JJ Pickle Research Campus in Austin, TX would not only help to cross-verify the numerical results listed in this paper but would allow us to generalize our findings to linear optical system of size order of magnitudes higher than numerically tested.

In summary, over the past few months we conducted and analyzed experiments on BosonSampling with peaked output distributions. In particular, we replicated known behavior from the random circuit model and extended these ideas to the linear optical domain, where our work explores the use of postselecting and explicitly peaking BosonSampling circuits to make experiments more verifiable and accessible. Ultimately, we provide the first exploration of peaked linear optical networks for potentially verifiable quantum advantage, which can be seen as an alternative model to Aaronson and Zhang’s work on random circuit sampling [2]. These studies lay the foundation for further experiments on peaked optics models and point at future directions to further explore the properties of peaked optics, both analytically and experimentally on higher modes. A copy of the classical simulations for this project can be found here: <https://github.com/michelled01/Peaked-circuits>.

References

- [1] Scott Aaronson and Alex Arkhipov. *The Computational Complexity of Linear Optics*. 2010. arXiv: [1011.3245](https://arxiv.org/abs/1011.3245) [quant-ph]. URL: <https://arxiv.org/abs/1011.3245>.
- [2] Scott Aaronson and Yuxuan Zhang. *On verifiable quantum advantage with peaked circuit sampling*. 2024. arXiv: [2404.14493](https://arxiv.org/abs/2404.14493) [quant-ph]. URL: <https://arxiv.org/abs/2404.14493>.
- [3] Andris Ambainis and Joseph Emerson. *Quantum t -designs: t -wise independence in the quantum world*. 2007. arXiv: [quant-ph/0701126](https://arxiv.org/abs/quant-ph/0701126) [quant-ph]. URL: <https://arxiv.org/abs/quant-ph/0701126>.

- [4] Anurag Anshu and Tony Metger. “Concentration bounds for quantum states and limitations on the QAOA from polynomial approximations”. In: *Quantum* 7 (May 2023), p. 999. ISSN: 2521-327X. DOI: [10.22331/q-2023-05-11-999](https://doi.org/10.22331/q-2023-05-11-999). URL: <http://dx.doi.org/10.22331/q-2023-05-11-999>.
- [5] Frank Arute et al. “Quantum supremacy using a programmable superconducting processor”. In: *Nature* 574 (2019), pp. 505–510. DOI: <https://doi.org/10.1038/s41586-019-1666-5>.
- [6] Marcello Benedetti, Harry Buhrman, and Jordi Weggemans. *Complement Sampling: Provable, Verifiable and NISQable Quantum Advantage in Sample Complexity*. 2025. arXiv: [2502.08721](https://arxiv.org/abs/2502.08721) [quant-ph]. URL: <https://arxiv.org/abs/2502.08721>.
- [7] Adam Bouland et al. “On the complexity and verification of quantum random circuit sampling”. In: *Nature Physics* 15.2 (Oct. 2018), pp. 159–163. ISSN: 1745-2481. DOI: [10.1038/s41567-018-0318-2](https://doi.org/10.1038/s41567-018-0318-2). URL: <http://dx.doi.org/10.1038/s41567-018-0318-2>.
- [8] Sergey Bravyi, David Gosset, and Yinchen Liu. *Classical simulation of peaked shallow quantum circuits*. 2023. arXiv: [2309.08405](https://arxiv.org/abs/2309.08405) [quant-ph]. URL: <https://arxiv.org/abs/2309.08405>.
- [9] Chi-Fang Chen et al. *Incompressibility and spectral gaps of random circuits*. 2024. arXiv: [2406.07478](https://arxiv.org/abs/2406.07478) [quant-ph]. URL: <https://arxiv.org/abs/2406.07478>.
- [10] Chi-Fang Chen et al. *Incompressibility and spectral gaps of random circuits*. 2024. arXiv: [2406.07478](https://arxiv.org/abs/2406.07478) [quant-ph]. URL: <https://arxiv.org/abs/2406.07478>.

- [11] William R. Clements et al. “Optimal design for universal multiport interferometers”. In: *Optica* 3.12 (Dec. 2016), pp. 1460–1465. DOI: [10.1364/OPTICA.3.001460](https://doi.org/10.1364/OPTICA.3.001460). URL: <https://opg.optica.org/optica/abstract.cfm?URI=optica-3-12-1460>.
- [12] Alexander M. Dalzell, Nicholas Hunter-Jones, and Fernando G. S. L. Brandão. “Random Quantum Circuits Anticoncentrate in Log Depth”. In: *PRX Quantum* 3.1 (Mar. 2022). ISSN: 2691-3399. DOI: [10.1103/prxquantum.3.010333](https://doi.org/10.1103/prxquantum.3.010333). URL: <http://dx.doi.org/10.1103/PRXQuantum.3.010333>.
- [13] Austin G. Fowler and Lloyd C. L. Hollenberg. “Scalability of Shor’s algorithm with a limited set of rotation gates”. In: *Physical Review A* 70.3 (Sept. 2004). ISSN: 1094-1622. DOI: [10.1103/physreva.70.032329](https://doi.org/10.1103/physreva.70.032329). URL: <http://dx.doi.org/10.1103/PhysRevA.70.032329>.
- [14] Michael Reck et al. “Experimental realization of any discrete unitary operator”. In: *Phys. Rev. Lett.* 73 (1 July 1994), pp. 58–61. DOI: [10.1103/PhysRevLett.73.58](https://doi.org/10.1103/PhysRevLett.73.58). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.73.58>.
- [15] Brian Skinner, Jonathan Ruhman, and Adam Nahum. “Measurement-Induced Phase Transitions in the Dynamics of Entanglement”. In: *Physical Review X* 9.3 (July 2019). ISSN: 2160-3308. DOI: [10.1103/physrevx.9.031009](https://doi.org/10.1103/physrevx.9.031009). URL: <http://dx.doi.org/10.1103/PhysRevX.9.031009>.
- [16] Adam Bene Watts et al. *Quantum advantage from measurement-induced entanglement in random shallow circuits*. 2024. arXiv: [2407.21203](https://arxiv.org/abs/2407.21203) [quant-ph]. URL: <https://arxiv.org/abs/2407.21203>.

- [17] Muqing Xu et al. *A neutral-atom Hubbard quantum simulator in the cryogenic regime*. 2025. arXiv: 2502.00095 [cond-mat.quant-gas]. URL: <https://arxiv.org/abs/2502.00095>.