Classical algorithm for simulating experimental Gaussian boson sampling

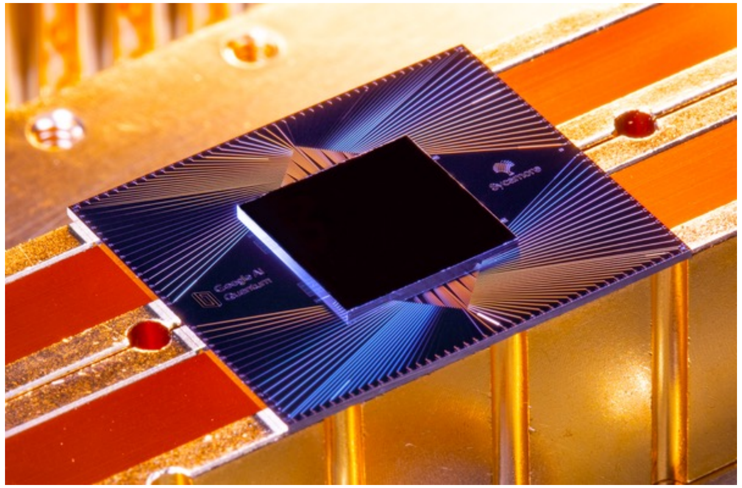
arXiv:2306.03709 (2023) (Nature Physics in press)

Changhun Oh[1,2,*], Minzhao Liu[1,3,*], Yuri Alexeev[1,3], Bill Fefferman[1], and Liang Jiang[1]

1. University of Chicago
2. Korea Institute of Advanced Science and Technology
3. Argonne National Laboratory

*These authors contributed equally to this work

The first "Quantum Advantage" claims



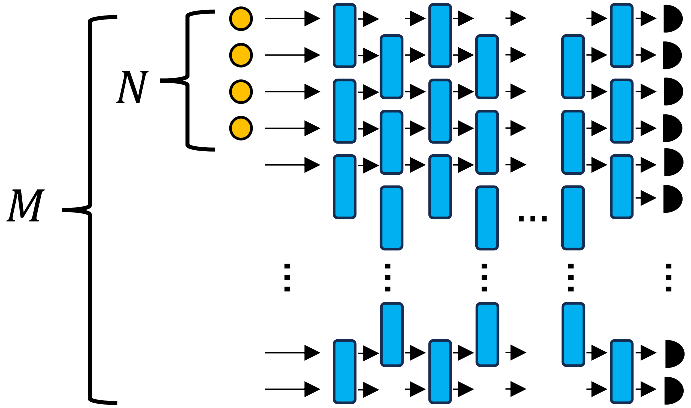
Random circuit sampling
Google (2019, 2023), USTC (2021)

Gaussian boson sampling
USTC (2020, 2021, 2023), Xanadu (2022)

Have we achieved quantum advantage?

For Gaussian boson sampling, our result says probably not yet.

Single-photon boson sampling



$$\vec{m} = (m_1, \dots, m_M)$$

$$\mathrm{Per}(Y) = \sum_{\sigma \in S_N} \prod_{i=1}^{N} Y_{i,\sigma(i)}$$

- (Input) N single-photon input in M modes $M \gg N^2$ (birthday paradox)

$$|\psi_{in}\rangle = |\overbrace{\underbrace{1 \dots 1}_{}}^{N}\underbrace{0 \dots 0}_{}\rangle \quad (\text{length } M)$$

Basis:

$$\{|n_1 \dots n_M\rangle\} \quad \sum_{i=1}^{M} n_i = N \quad dim = \binom{N+M-1}{N}$$

- (Dynamics) M-mode beam-splitter network; Haar-random unitary matrix $U \in \mathrm{U}(M)$
- (Measurement) Photon number measurement on each mode.

$$\vec{m} = (m_1, \dots, m_M)$$

- (Why hard in noiseless case) Output probability is described by the permanent of i.i.d. Gaussian matrices X:

$$p(\vec{m}) = |\langle m_1 \dots m_M|\hat{\varphi}(U)|n_1 \dots n_M\rangle|^2 = \frac{\left|\mathrm{Per}\, U_{\vec{m},\vec{n}}\right|^2}{m_1! \dots m_M!\, n_1! \dots n_M!} \propto |\mathrm{Per}\, X|^2$$

- Generating many single-photons simultaneously is extremely difficult in experiment.

S. Aaronson and A. Arkhipov (2011)

Gaussian state and covariance matrix

- Elementary bosonic operators - photon creation and annihilation operators: $\hat{a}^\dagger$ and $\hat{a}$.

  $[\hat{a}, \hat{a}^\dagger] = \hat{a}\hat{a}^\dagger - \hat{a}^\dagger\hat{a} = 1.$ $\hat{a}|n\rangle = \sqrt{n}|n-1\rangle, \ \ \hat{a}^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle$

  $\hat{x} = \hat{a} + \hat{a}^\dagger$, $\hat{p} = i(\hat{a}^\dagger - \hat{a})$

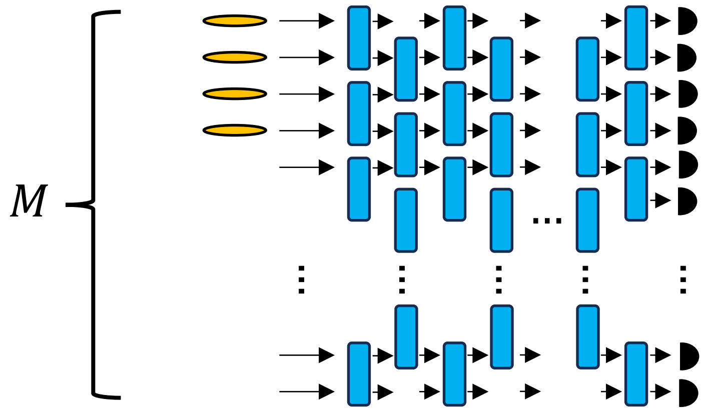
- Gaussian state: Quantum states that can be fully characterized by its mean and covariance matrix.

  $$d = \begin{bmatrix} \langle\hat{x}\rangle \\ \langle\hat{p}\rangle \end{bmatrix}, V = \begin{bmatrix} \langle\hat{x}^2\rangle - \langle\hat{x}\rangle^2 & \langle\hat{x}\hat{p} + \hat{p}\hat{x}\rangle/2 - \langle\hat{x}\rangle\langle\hat{p}\rangle \\ \langle\hat{x}\hat{p} + \hat{p}\hat{x}\rangle/2 - \langle\hat{x}\rangle\langle\hat{p}\rangle & \langle\hat{p}^2\rangle - \langle\hat{p}\rangle^2 \end{bmatrix} \qquad \langle\hat{O}\rangle = \mathrm{Tr}[\hat{\rho}\hat{O}]$$

- Examples: Vacuum state $V = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, d = 0$, Squeezed vacuum state $V = \begin{bmatrix} e^{2r} & 0 \\ 0 & e^{-2r} \end{bmatrix}, d = 0$

- What's not Gaussian? Single-photon state

Gaussian boson sampling



$\vec{m} = (m_1, \dots, m_M)$

- (Input) N **squeezed** vacuum states $d = 0.$

  M independent covariance matrix $\begin{bmatrix} e^{2r_i} & 0 \\ 0 & e^{-2r_i} \end{bmatrix}$, $r_i = r,\ i \in [N]$; $r_i = 0$, otherwise

- (Dynamics) M-mode Haar-random beam-splitter network $U \in \mathrm{U}(M)$
- It rotates the covariance matrix.

  $V \to OVO^{\mathrm{T}}$ O is a $2M\times 2M$ orthogonal (and symplectic) matrix depending on U.

- (Measurement) Photon number measurement on each mode.
- (Why hard in noiseless case)

$$p(\vec{m}) = \frac{1}{\prod_{i=1}^{M} \cosh r_i} \frac{\left|\mathrm{haf}\, A_{\vec{m},\vec{m}}\right|^2}{m_1! \dots m_M!},\ A = A(U, \vec{r}),\ \ \mathrm{haf}(Y) = \sum_{\sigma \in PMP(n)} \prod_{(i,j)\in\sigma} Y_{ij}$$

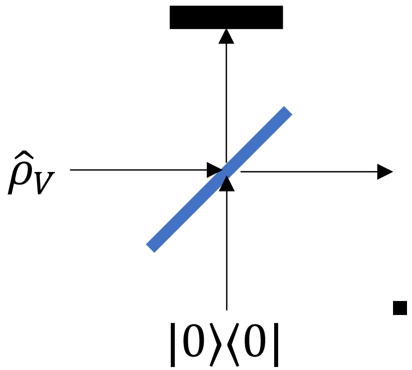
$\propto |\mathrm{haf}\, XX^T|^2$, where X is i.i.d. Gaussian matrix.

- The setup used in recent experiments.

C. S. Hamilton et al. (2017), A. Deshpande et al. (2022)

Dominant noise: Photon loss

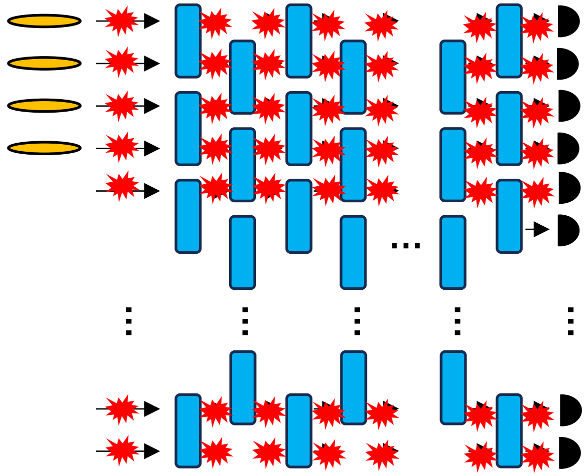
- Photon loss is the dominant noise model (local):



$V \to \eta V + (1-\eta) I_2$

η: Transmission rate
$1-\eta$: loss rate

- It transform a Gaussian state to another Gaussian state



- (Overall) loss rate : $\frac{\text{Average output photon number}}{\text{Average input photon number}}$
- (overall loss rate is between 50% to 70%.)
- Other noise types exist, but **loss is dominant in the current experiments.**

Classical-state approximation

- **Uncorrected noise** often transforms a quantum state to a **classical state** that's easy to simulate, as the system size scales.
- The output state of **random circuit sampling** with a constant level of depolarizing noise per depth converges to **maximally mixed state (easy to sample from).**

[D. Aharonov et al. (1996), A. Deshpande et al. (2022)]

- The trivial algorithm is to sample from the uniform distribution.

- Similarly, the output state of **Gaussian boson sampling** with a constant level of photon-loss per depth converges to **thermal state (easy to sample from).**

[R. Garcia-Patron et al. (2019), H. Qi et al. (2020)]

- The trivial algorithm is to sample from the distribution from the thermal state.

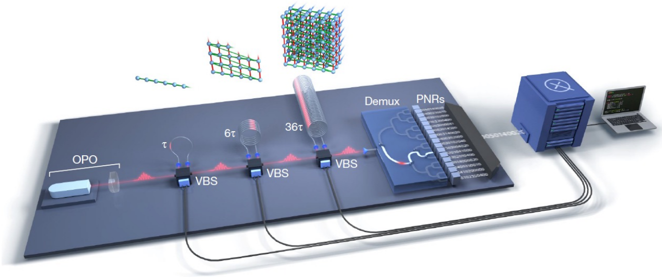
- Current GBS experiments' output state is believed to be (observably) far from **thermal state.**

[H.-S. Zhong et al. (2021), L. S. Madsen et al. (2022), Y.-H. Deng et al. (2023)]

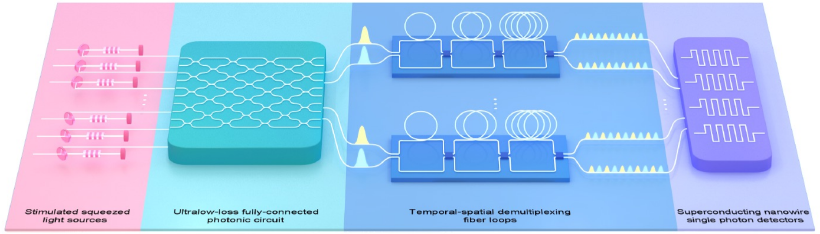
- Our new classical algorithm can simulate the largest GBS experiments.

Largest GBS experiments so far

Xanadu



USTC



- Output average photons: **>100**
- Loss rate: **0.5-0.7**
- Claims **at least 100 years** for classical computer to generate a sample
  [L. S. Madsen et al. (2022), H.-S. Zhong et al. (2021), Y.-H. Deng et al. (2023)]

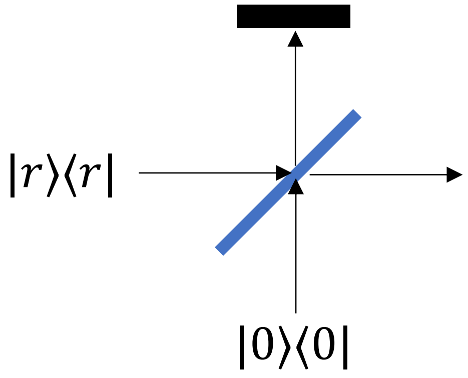
- The **best-known classical algorithm** is based on the chain-rule-based sampler.
  [N. Quesada et al. (2020, 2022), J. Bulmer et al. (2022)]

$$p(m_1, \dots, m_M) = p(m_1)p(m_2|m_1)\cdots p(m_M|m_1, \dots, m_{M-1})$$

- **Main cost = Computing the output marginal probability**, which increases as the **output photon number.**
- However, it does **not** take advantage of loss and noise.

Our contribution: decomposition of loss channel

- Photon loss on squeezed vacuum state (single-mode example)



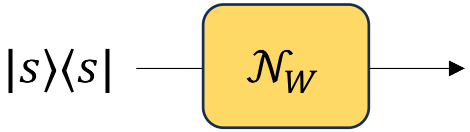
- $V = \eta \begin{bmatrix} e^{2r} & 0 \\ 0 & e^{-2r} \end{bmatrix} + (1-\eta) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

$$= \begin{bmatrix} e^{2s} & 0 \\ 0 & e^{-2s} \end{bmatrix} + \begin{bmatrix} \eta e^{2r} + 1 - \eta - e^{2s} & 0 \\ 0 & 0 \end{bmatrix}$$

$$\equiv V_p + W$$

- $e^{-2s} = \eta e^{-2r} + (1-\eta)$

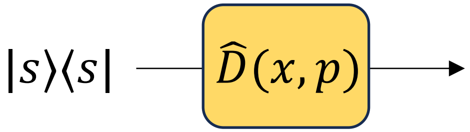
=



- Gaussian convolution

$$\hat{\rho}_{V_p+W} = \int_{-\infty}^{\infty} dxdp \; p_W(x,p)\hat{D}(x,p)\hat{\rho}_{V_p}\hat{D}^\dagger(x,p)$$

=



$\hat{D}(x,p)$: displacement operator

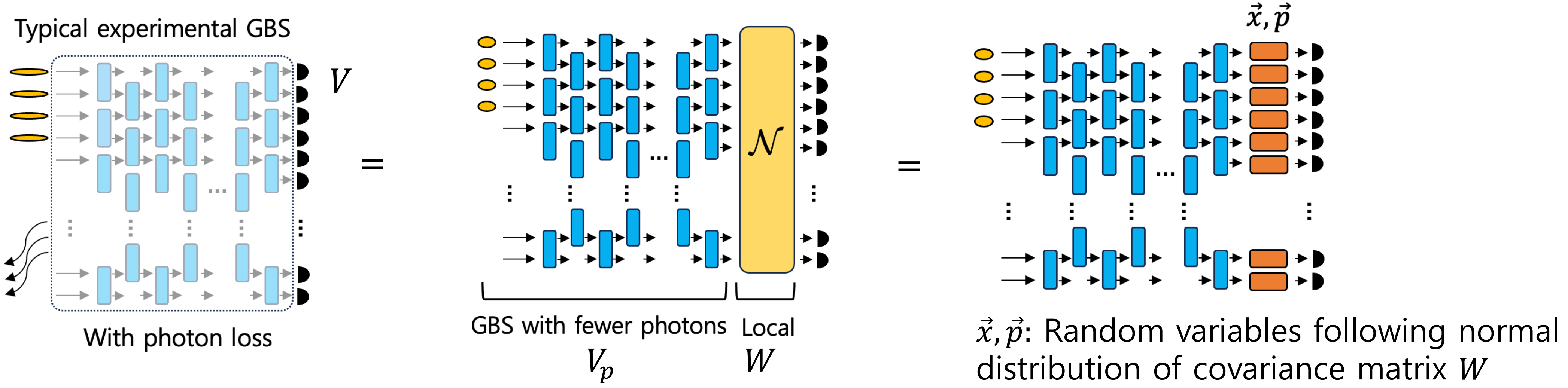
$$\hat{D}(x,p)\hat{x}\hat{D}^\dagger(x,p) = \hat{x} + x$$
$$\hat{D}(x,p)\hat{p}\hat{D}^\dagger(x,p) = \hat{p} + p$$

Random variables x, p following a normal distribution of covariance matrix W

- The choice of s may not be unique.

Our contribution: decomposition of lossy GBS

- More generally, the multimode output covariance matrix is decomposed as $V = V_p + W$.

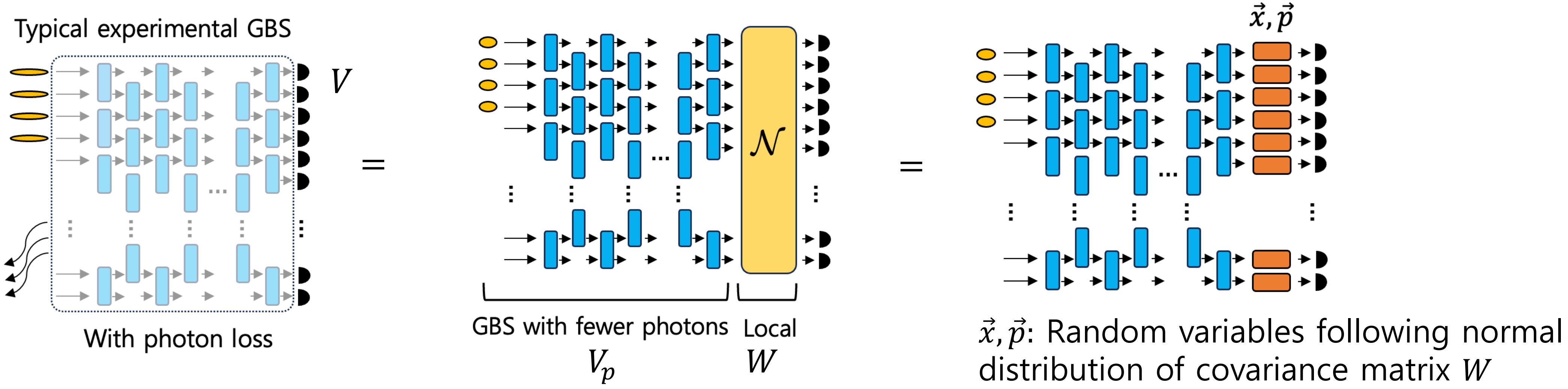


$$\hat{\rho}_{V_p+W} = \int_{-\infty}^{\infty} d\vec{x}d\vec{p}\ p_W(\vec{x},\vec{p})\widehat{D}(\vec{x},\vec{p})\hat{\rho}_{V_p}\widehat{D}^\dagger(\vec{x},\vec{p})$$

- The choice of V_p may not be unique, but we want to minimize the photon number $\propto \mathrm{Tr}[V_p]$.
- Semidefinite programming: $\min_{V_p} \mathrm{Tr}[V_p]$ with $V = V_p + W,\ W \geq 0, V_p \geq i \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \otimes I_M$.
- Ex) In recent experiments, only **3-7%** of the output photons (average ~10) are from V_p.

Our contribution: decomposition of lossy GBS

- More generally, the multimode output covariance matrix is decomposed as $V = V_p + W$.



- Thus, high-loss cases like current experiments have **very small amount of entanglement.**
- We only need to simulate **GBS with fewer photons** (V_p) for high-loss rate.
- Since V_p has a very small number of photons and the W part is **local**, we can simulate it using a tensor network method (matrix product state).
- If we approximate $V_p \approx I_{2M}$ (vacuum), the output state is a mixture of product states because displacement operators are local. This is the **thermal-state approximation** known in the literature.
- [R. Garcia-Patron et al. (2019), H. Qi et al. (2020)].

How to verify GBS in experiment?

- The best way in principle would be to compute **total variation distance** (TVD):

  $$\text{TVD} = \sum_x \left| p_{id}(x) - p_{exp}(x) \right| /2$$

  But TVD is neither sample-efficient nor computationally efficient.

- **XEB** (cross entropy benchmarking) has been used inspired by random circuit sampling.

  $$\text{XEB} = \sum_x p_{id}(x) p_{exp}(x) \approx \frac{1}{N_{samp}} \sum_{i=1}^{N_{samp}} p_{id}(x_i)$$

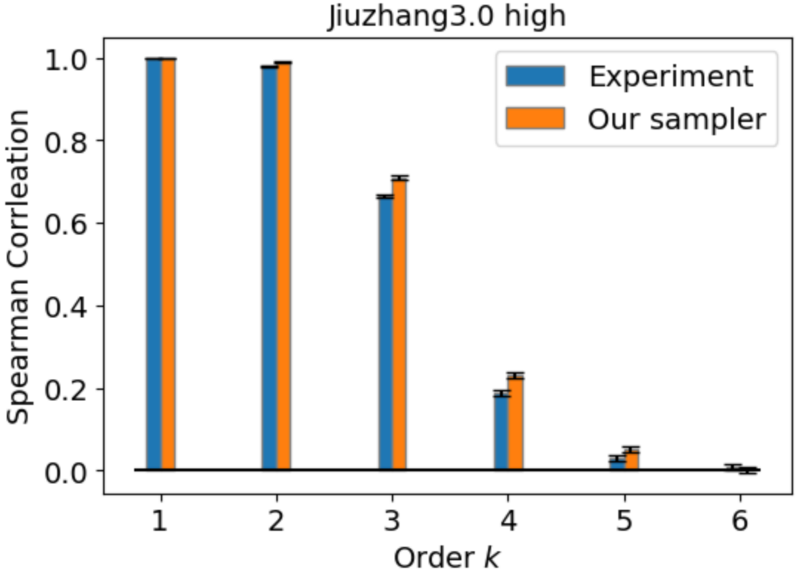
  But computing the ideal XEB is open in GBS.

- Also, Google was able to spoof all the benchmarks except **higher-order correlation** [B. Villalonga et al. (2021)].

Numerical results (largest)

Higher-order correlations $\kappa(m_1, \cdots, m_k) \equiv \mathbb{E}[m_1 m_2 \cdots m_k] - \sum_{p \in P_k} \prod_{b \in p} \kappa[(m_i)_{i \in b}]$

- Figure of merit used by experiments: Spearman correlation(q_{id}, q), a function of two distributions, q_{id}, q.
- Choose k marginal modes out of total M modes and compute $\kappa(m_1, \cdots, m_k)$ for each marginal of both q_{id} and q.



- Up to the largest order correlation that was calculated in experiments, **we do not observe any clear advantage from experiments.**
- Generating 10M samples takes around **an hour** with less than **300 GPUs.**

CO et al., arXiv:2306.03709 (2023)

Conclusion

- New classical algorithm using tensor network which can simulate the state-of-the-art Gaussian boson sampling experiments.

- For future experiments toward quantum computational advantage:
  - **Loss rate** has to be decreased significantly instead of merely increasing the **output photon number**.

  (For example, USTC's new experiment (2023) turns out to be easier than the previous experiment (2021) for our method because the loss rate increased.)

Open questions

- Better classical sampler?
  - The V_p part has such a small photon number, but our sampler still takes a long time because we need to store the full quantum state (V_p) using MPS to take care of the subsequent displacement part.
  - There might exist a clever way to simulate the system without this.
  - If one just wanted to sample from V_p without displacement, it would've been instantaneous.

- Ideal XEB score? $XEB_{id} = \sum_x p_{id}(x)^2$.
  - 2^{nd} moment of output probability: $\mathbb{E}_X[|\mathrm{haf}\, XX^T|^4] = ?$, where X is i.i.d. Gaussian matrix.
  - Note that heavy-outcome-generation in boson sampling is possible because the output probability correlates with marginal probabilities [CO et al. (2023)].

- Better verification method is needed.
  - The higher-order correlation method does not have a complexity-theoretic ground, such as the relation to TVD.

Thank you!

Estimation of larger systems