**Q1:** In **II.B**, “However, neither does the boson-sampler know even the approximate probability distribution before enough samples are obtained.” What does that sentence mean?

**A:** It means boson-sampler generates samples without guidelines. There is not a demon to control the sampling process. Or one can say that the sampler itself does not know the distribution. Because it cannot tell other people any details about the distribution. So it’s unfair to let a classical computer to compute the whole distribution before simulating boson-sampling. Just want to lead to introduce the following smarter classical algorithms. In order to avoid misunderstanding, maybe deleting this sentence is better.

**Q2:** In **II.B**, “For example, using the acceptance-rejection method (Wikipedia, 2016)” Is there a better reference than Wikipedia?

**A:** How about this one?

@book{fishman2013,

title={Monte Carlo: concepts, algorithms, and applications},

author={Fishman, George},

year={2013},

publisher={Springer Science \& Business Media}

}

**Q3:** In **III.B,** “following the same way for original boson-sampling in Sec. II”. What does this mean?

**A:** It means the comparing method is the same as the one in Sec. II.C. That is to say, the line means , and the region to the right of each line is the regime of quantum supremacy, etc.

**Q4:** In the figure caption of **Fig5,** Why aren't the curves monotonic?! It's very strange that the curves change direction as photon number increases??

**A**: Do you refer to the lines 7 and 8 in Fig5(a)? I think these two curves are not monotonic because they locate near the critical region. In other words, their clock speeds (107Hz,108Hz) are close to the pump repetition rate 76 MHz.

**Q5:** In **III.C,** “Shouldn't this be m2?”

**A**: If a high-Q cavity, X-mon qubit and a low-Q cavity constitute a spatial mode of the unitary network, then we can see that there are  spatial modes. Every time we apply an operation to each mode, so  operations can be done simultaneously. From the view point of time consumption of a  unitary transformation, it is , where  is the time needed for a single step.

**Q6**: In **III.C,** This doesn't make sense to me – explain it better

**A:** It is showed in the Fig5(b) that microwave boson-sampler has a more demanding requirement for total system efficiency to achieve the same sampling rate as spatially- encoding optical boson sampler. We can understand it from two aspects. First, the reason of higher demand is that a single operation for microwave boson sampler takes time of the order of μs. By contrast, the processing speed of an optical network is rather fast, on the order of ns. Second, high efficiency is not such a difficult problem for superconductor platform like that in optical boson-sampler whose system efficiency is always limited by unsatisfying collection and detection efficiency. Preparation, operation and readout fidelity in superconducting system all can be larger than 0.9. It means that its low processing speed can be compensated by high system efficiency which finally makes it still a very promising platform for scalable boson-sampling.

**Q7:** In **III.D,** Need to define  and meaning of .

**A:** 

where  denotes the set of all possible no-collision configuration basis of  photons in the  initial input photons, and .

**Q8:** In the figure caption of **Fig7**: “(Why does the curve have the change in direction?)”

**A:** The curve is the result of match between classical computer and boson sampler among three parameters: the lost number, the input number, and efficiency. i) The lines order from right to left depends on the trade-off relation between lost number and efficiency. From equation-6, we can see that, more lost number, lower efficiency could reach. So, the line of nlost=5 is on the left of that of nlost=0. ii) The lines order from down to up depends on the reference number. We are sorry to forget to point out that the photon number in vertical axis is the input number rather than the output one. More lost number, more input photons should have. Thus the line of nlost=5 is above that of nlost=0.

**Q9:** In the figure caption of **Fig9**: (is that efficiency per mode, or total efficiency??)

**A:** It is the efficiency of a photon including coupling, transmitting, and detecting, but not generating.

**Q10:** In **III.D.1,** What's the funny binomial?

**A: ** is an abbreviated form of multinomial coefficient ****, where ****. In fact, the eq.18 comes from the multinomial theorem:



**Q11:** In **V.A,** How is this different to type-I? Maybe we could explain what this means? I'm a theorist and I don't know these terms.

**A**: We have rewritten this part as following:

In SPDC, a laser pumps a second-order nonlinear material (e.g. BBO, KDP) and creates two parametric single photons (the signal and the idler) under the constraints of energy conservation and momentum conservation (also called phase matching):



where  stands for the photon frequency,  for the wave vector, and their subscripts p, s and i for the pump, the signal and the idler, respectively. There are several different types of phase matching:

* Type-I: The signal and the idler share the same polarization.
* Type-II: The signal and the idler have orthogonal polarization.
* Type-0: The pump, the signal and the idler share the same polarization.

In a nonlinear crystal, there is phase mismatch due to dispersion, but which can be compensated by using birefringence. The birefringence phase matching (BPM) requires at least one of two parametric photons to be extraordinary, resulting in spatial walk-off between the two photons. As the pump propagates through the crystal, the walk-off leads to a decrease in the beam overlap, thus limiting the effective length of the crystal. Additionally, there is a so-called quasi-phase matching (QPM), which compensates the phase mismatch by periodical inversion of the crystal’s nonlinear coefficient. QPM allows for collinear interaction among the pump, the signal and the idler. This avoids spatial walk-off and permits the usage of longer crystals and efficient collection of parametric photons. As a result, the best collection efficiency of ~90% has been realized with QPM (Giustina et al., 2013; Christensen et al., 2013).

Q12: In **VI.A**, “The problem of physical size” I don't understand this sentence???

**A**: “The problem of exaggerated physical size” is the same problem in the last sentence of the above paragraph. They are too large for circuits containing hundreds of DCs.

Q13: In **IX**, “Therefore, improving repeat frequency plays a bigger role than photon number for those circuits with limited size but high transmittance, e.g., bulk optical circuits.” Don't quite understand this last sentence.

**A**: As we know, boson-sampling with 20 photons is a threshold to reach quantum supremacy. But as discussed before, the photon number thresholds can lower to less than 20 in systems with higher efficiencies and repeat frequencies. For those circuits with limited size, e.g., bulk optical circuit, maybe it is hard to provide large enough modes for 20 photons boson-sampling. However, one can also reach the threshold by improving transmittance and repeat frequency of the boson-sampler.

**Q14**: In **VIII**. A VISION FOR THE FUTURE -THE POST-CLASSICAL ERA

**A**: Do you mean something like this?

