Simulating quantum computers with probabilistic methods: Tokyo ImPACT lecture 3 – Boson sampling & XY machines

P. D. Drummond

Centre for Quantum and Optical Science, Swinburne University of Technology.

Outline

Idea of this talk: ‘Quantum software’

We need novel methods for high-order correlations

- Quantum software uses complex contour integrals
- Excellent results for treating boson sampling
- Leads to new exact analytic results
- Individual experiments treated with sampling methods
- Can treat: any input, decoherence, any order correlations

New verification proposals, high-N interferometer predictions

What is boson sampling?

Boson sampling is a computationally hard problem

- Send N single photons through an M-channel photonic device
- Measure the output photon number distribution
- The boson sampling device solves the problem of how to generate bits with a permanent probability distribution
- Matrix permanents are a '#P' hard problem, taking exponentially long times to compute at large N

Boson sampling is conjectured to be a '#P' hard problem (Aaronson, MIT)

Why is it worth studying this?

Boson sampling is the **simplest** quantum complexity

- Solving it may solve other problems
  - Maybe its only as useful as climbing Mt Everest?
  - Maybe it will have many applications!

Spinoffs:

- Insight into bosonic many-body complexity
- Linkages with group theory, mathematics
- Heisenberg-limited metrology
  - robust against decoherence - see later slides

High order correlations?

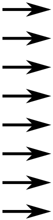
Bosn sampling emphasizes importance of very high order correlations

- Typically its very hard to simulate or measure high order correlations
  - Quantum computer outputs are often equivalent to a high-order correlation

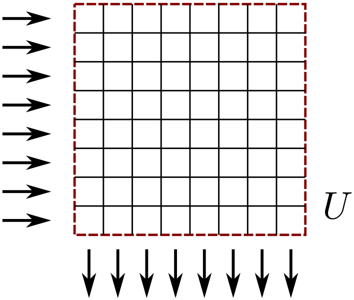
Advantages

- This type of effect greatly enhance nonclassical effects
- Quantum networks will involve high-order correlation effects
- **Is there a generic metrology** advantage?
  - Boson sampling is related to N-partite entanglement

Boson sampling experiment

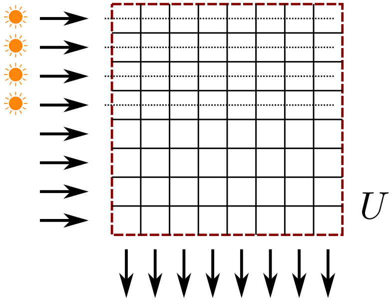


Boson sampling experiment



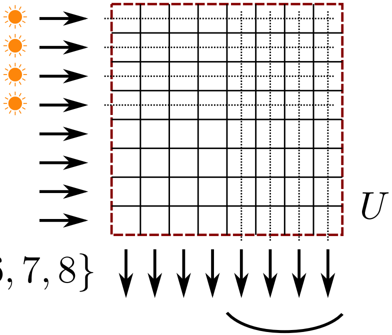
Boson sampling experiment

$$\vec{j} = \{1, 2, 3, 4\}$$



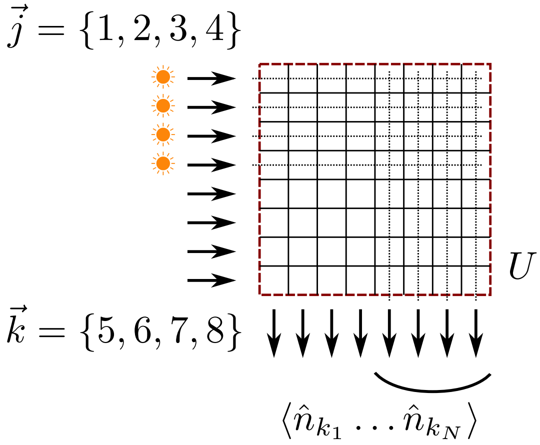
Boson sampling experiment

$\vec{j} = \{1, 2, 3, 4\}$

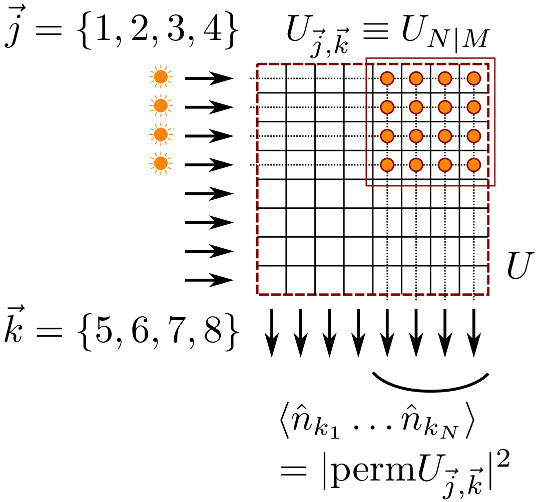


$\vec{k} = \{5, 6, 7, 8\}$

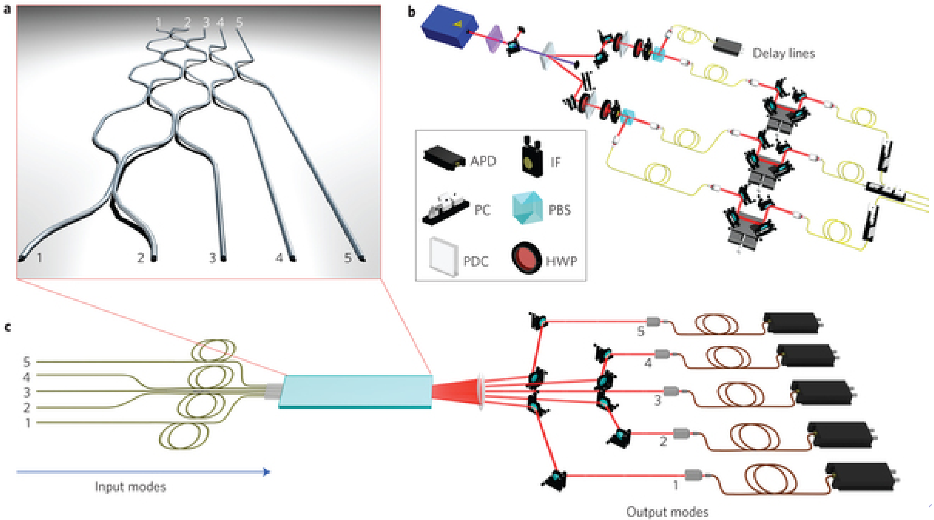
Boson sampling experiment



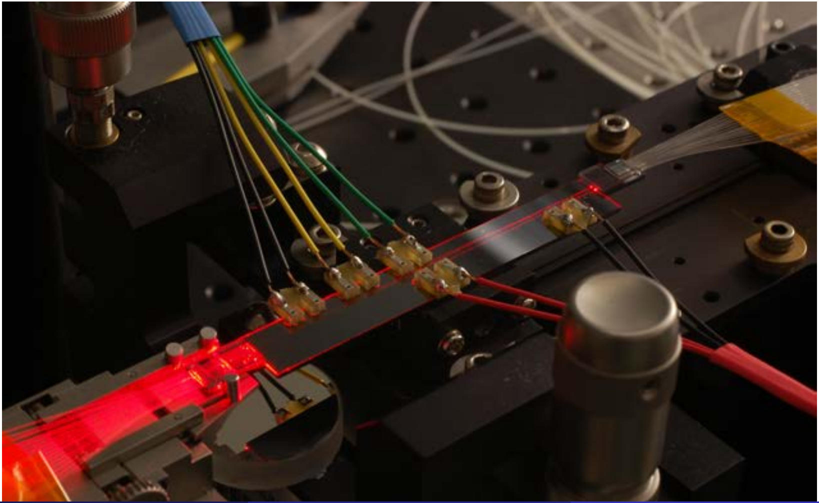
Boson sampling experiment



Experimental schematic



Experiments: Oxford, Queensland, Rome, Vienna..



Why is boson sampling computationally hard?

There are exponentially many interfering paths!

- The N-photon probability is a **matrix permanent**
- $$P = \left| \sum_{\sigma} \prod_{i} T_{i,\sigma(i)} \right|^2$$
- Here $T = \sqrt{1-\gamma} U$: U is an $N \times N$ (sub)unitary, γ a loss
- Standard methods take $N \times 2^N$ operations
- TRILLIONS of years for $N = 100$ at 1GFlop
- Impossible even on the largest supercomputers

How do we know the QC works at large N ?

Requirements for verification of boson sampling

- Must be **measurable** in finite time
- Must be **calculable** in finite time
- Must **distinguish** different unitaries
- Must be **universal** - almost all unitaries
- Must be **non-forgeable**: *no VW mode!*
- **Many suggestions: none do all this**

The complex P-representation

The contour can be a closed path

Case of an M-mode state, occupations n_j :

$$P(\boldsymbol{\alpha},\boldsymbol{\beta}) = \prod_j \left[\frac{n_j!}{2\pi i}\right]^2 \frac{e^{\alpha_j\beta_j} d\alpha_j d\beta_j}{(\alpha_j\beta_j)^{1+n_j}}$$

Any contour that encloses the origin is OK; we let $\alpha_i = r\exp\left(i\phi_i^\alpha\right)$.
Can also use a sum over $d > Max(n_j)$ discrete phases.
See: Physical Review A 94, 042339 (2016), arXiv1605.05796, 1609.05614

$$\chi(\boldsymbol{\xi}) = \oint \ldots \oint P(\boldsymbol{\alpha},\boldsymbol{\beta})\, e^{\boldsymbol{\xi}\cdot\boldsymbol{T}^*\boldsymbol{\beta} - \boldsymbol{\xi}^*\cdot\boldsymbol{T}\boldsymbol{\alpha}} d\boldsymbol{\alpha} d\boldsymbol{\beta} \,.$$

What is the transfer matrix?

The matrix T includes linear couplings and losses

We consider the case of $T = \sqrt{t}U$, in which:

- U is the the unitary mode transformation U of the photonic network
- transmission coefficient $t = 1 - \gamma$ combines losses and detector inefficiencies
- T can be considered a submatrix of a larger unitary that includes couplings to unobserved reservoirs.

Unitary averages can be calculated exactly

We need an exact analytic theory for scaling laws

So, we take averages over all unitaries
This can be carried out EXACTLY
Can predict scaling laws for any size matrix!

Thanks to Yan Fyodorov's nice theorem!

- Unitary averaging methods integrate over a Haar measure.
- Fyodorov's theorem averages over exponentials of matrix traces
- **This allow us to average permanents analytically**

Unitary averaging

How do we interpret the physics?

- The output coherence properties ONLY depend on $\boldsymbol{\beta} \cdot \boldsymbol{\alpha}$,
- Phase-space equivalent of the total *input* photon number $\hat{N}$:

$$\left\langle \chi^{(\mathrm{out})}(\boldsymbol{\xi}) \right\rangle_U = (M-1)! \sum_{j=0}^{M} \frac{\left(-t\left|\boldsymbol{\xi}\right|^2\right)^j \left\langle : \hat{N}^j : \right\rangle}{j!\,(M-1+j)!}$$

After unitary averaging, it doesn't matter where the photons originate: any mode or combination is OK!

Next, let's calculate the photon correlations

Photon correlations are just derivatives of the photon-number generator

$$G(\boldsymbol{\gamma}) \equiv Tr\left(\hat{\rho}\prod_i (1-\gamma_i)^{\hat{n}_i}\right) = \int\int\prod_i\left[\frac{1}{\pi\gamma_i}\exp\left(\frac{-|\xi_i|^2}{\gamma_i}\right)\right]\chi(\boldsymbol{\xi})\,d^{2m}\boldsymbol{\xi}$$

A case of special interest is $P_{N|M}$, the probability of observing 1 photon in each of N channels, given a N-photon input and an M-mode network. This is found on taking n first derivatives of $G(\boldsymbol{\gamma})$, so that:

$$P_{N|M} = \frac{t^N (M-1)!N!}{(M-1+N)!} = t^N\left[C_N^{M+N-1}\right]^{-1}$$

First type of asymptotic scaling law

Scaling for count-rates of just one set of channels

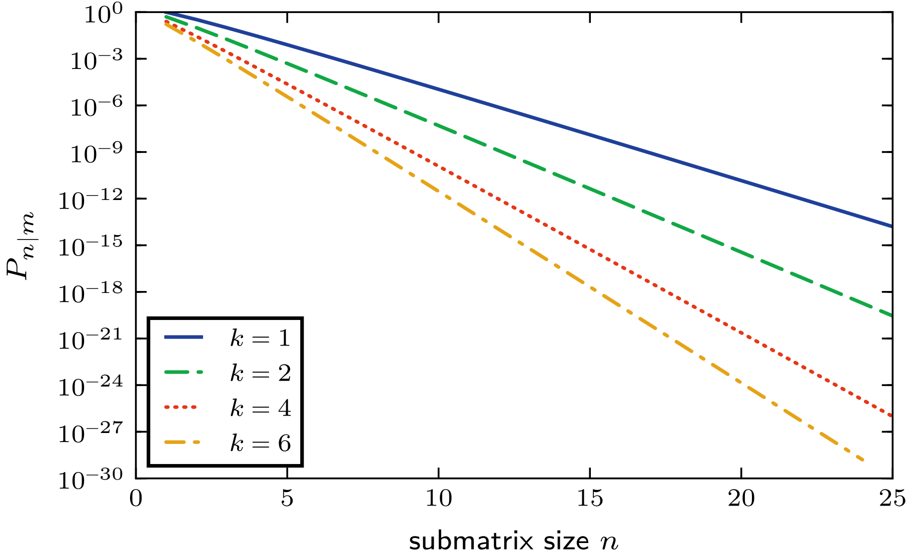
For the general case, the scaling exponent is
$\varepsilon = \log t + k \log k - (1+k)\log(1+k)$, and the asymptotic result is:

$$\log P_{N|M} \underset{n\to\infty}{\sim} N\varepsilon + \frac{1}{2}\log\left[2\pi N(1+1/k)\right] \quad (1)$$

Note, here $k = M/N$ is the channel ratio of filled to unfilled channels.

If $k = 1$, then we get the highest count-rates possible, i.e., $\varepsilon = \log(t/4)$. This means that many counts are 'lost' because they result in two-photon (or more) outputs in a single channel - the Hong-Ou-Mandel effect.

How does this look?



Problem: count rates are very low!!

Let's estimate the typical count rate

- Take a photon number of $N = 24$ - not that large
- Assume perfect efficiency, ie $t = 1$
- Assume a small channel ratio, ie $k = 2$
- Assume a high repetition rate of 10^{12}Hz
- Average permanent-squared is $\sim 10^{-18}$
- Need 12 days to get just one count!!

Second type of asymptotic scaling law

Scaling for count-rates of **any set of combined channels**

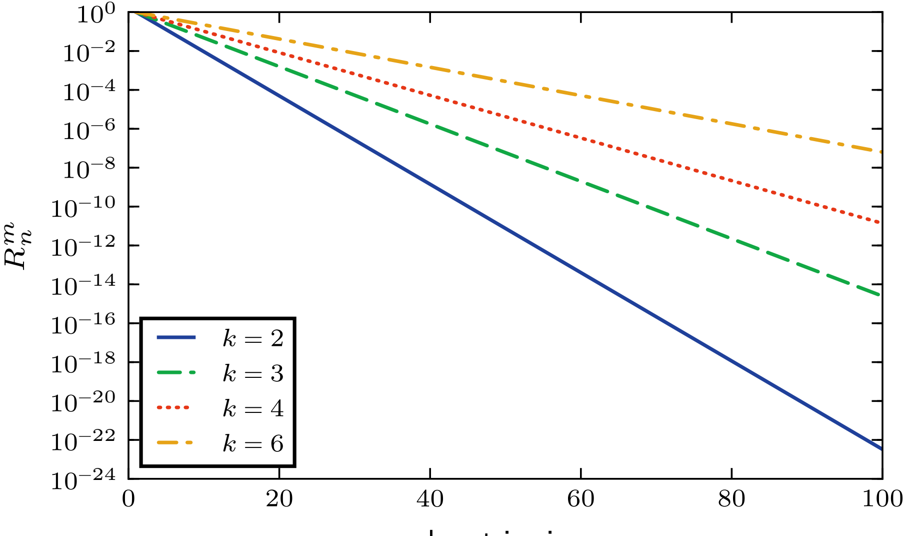
For the combined case, the scaling exponent is $\lambda = \log t + 2k\log k - (k-1)\log(k-1) - (k+1)\log(k+1)$, so that the scaling is:

$$\log R_{N|M} \underset{n\to\infty}{\sim} N\lambda + \frac{1}{2}\log\left[\frac{k+1}{k-1}\right] \tag{2}$$

Note, in this case a high channel ratio is best

If $k \to \infty$, then we get the highest count-rate possible,
i.e.,$\lambda \to \log t - 1/k$.
cf Arkhipov & Kuperberg: 'Birthday Paradox'
Carolan et al experiments, Bristol.

Verifying boson sampling with channel grouping



Physical interpretation

Let's estimate the typical count rate

Consider a number much too large for exact permanents

- Take a photon number of $N = 100$ - very large
- Assume imperfect efficiency, ie $t = 0.9$
- Assume a large channel ratio, ie $k = 6$
- Assume a high repetition rate of 10^{12}Hz
- Sum of permanent-squared is $\sim 10^{-12}$
- Get a count per second - feasible?

Also need results for individual unitaries

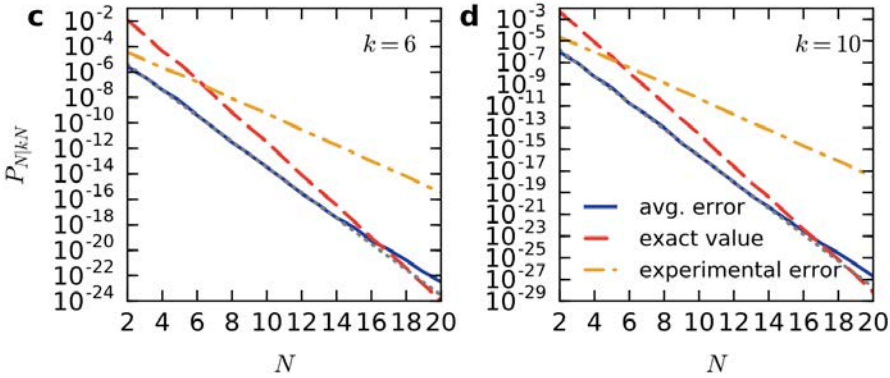
Let $\zeta_j \equiv exp(i\phi_j)$ be randomly distributed on a unit circle

$$perm(\left|\tilde{T}\,(S,S')\right|^2) = \left\langle \prod_{j\in S} (\zeta_j \zeta_j')^* \prod_{k\in S'} m_k \right\rangle$$

$m_k = n_k/r^2 \equiv \sum_{j,j'\in S} T_{kj}\zeta_j T^*_{kj'}\zeta'_{j'}$

Like Gurvitz approximation, but **unbiased** for permanent squared

Numerical averages for $k = 5, 10$: 100 random unitaries, 10^7 samples



Average sub-unitary permanent squared $P_{N|M}$:

How do we interpret this result?

- Complex-P error in $|P|^2$ **decreases** rapidly with matrix-size N
- But, the experimental sampling error is proportional to $|P|$
- **We calculate $|P|^2$ better than experiment!**
- Don't generate a digital bitstream - doesn't solve a #P problem
- Can verify ANY possible N-th order correlation!
- Problem: correlations too small to measure at large N

Calculating ALL submatrix permanents

An efficient Fourier transform is used:

- $$\sum_{S'} perm(\left|\tilde{T}\left(S,S'\right)\right|^2) = \left\langle C_N \prod_{j\in S} \left(\zeta_j \zeta_j'\right)^* \right\rangle$$
- $$C_N = \frac{1}{M} \sum_j e^{-ijN\phi} \prod_{k=1}^{M} \left[1 + e^{ij\phi} m_k\right]$$

All $N \times N$ subpermanents computed in parallel!

- Gives enormous speed-up compared to summing permanents
- Not fast enough to reach $N = 100$
- Need analytic formulae at large N

Combining 10^{34} distinct 30×30 subpermanents of a 180×180 matrix

A large random unitary is its own ensemble?

- We average over 10^{34} 30×30 submatrices
- **Takes over 10^{10} lifetimes of the universe classically**
- Count rates: identical to the full unitary average
- Large-N unitary averages give single-matrix averages
- Provides a way to verify boson sampling at arbitrary N
- Problem: fails to distinguish large-N unitaries

Boson sampling verification by testing nulls

Measure N single counts in any of $M-p$ channels

- **Treat the null channels as losses**
- Advantages: scalable, depends on N-boson correlation
- Distinguishes between the different possible unitaries
- Leads to an analytic conjecture for all N values

Asymptotic coincidence probability is:

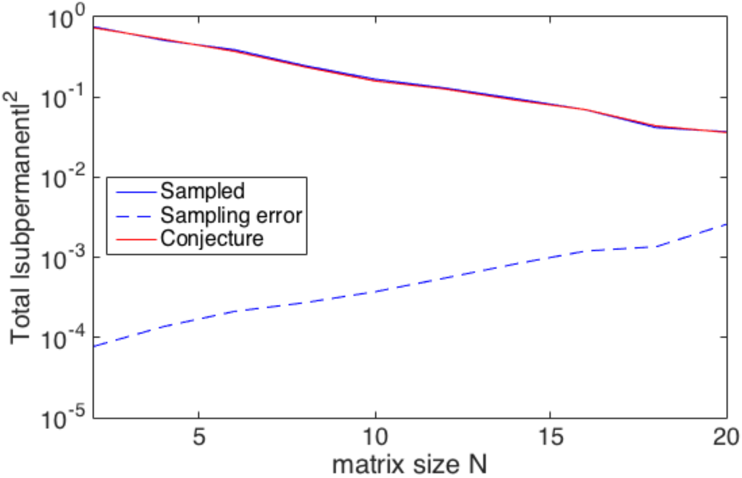
Conjecture :

$$\lim_{N\rightarrow\infty} P_{N|(kN-1)} = \frac{t_j^N (kN-1)!(kN-2)!}{(kN-N-1)!(kN+N-2)!}$$

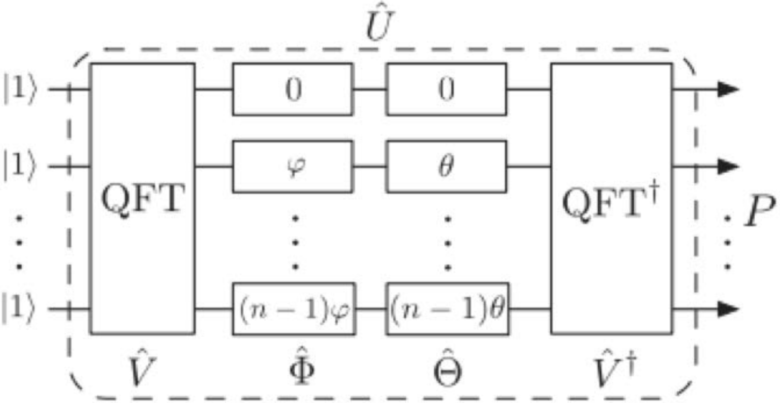
where: $t_j = \left(1 - \sum_{i=1}^{N} \left|U_{j,i}\right|^2 / N\right)$

is the effective loss for deletion of channel j.

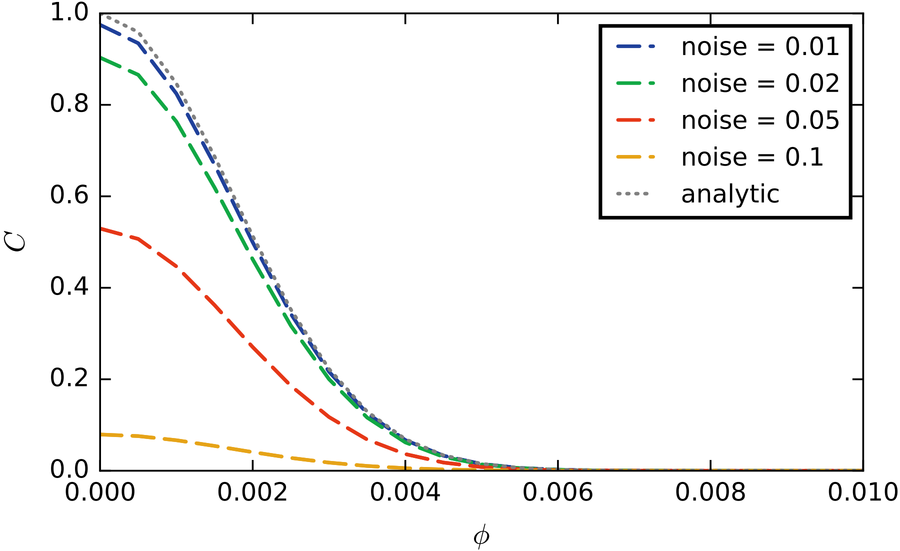
Conjecture verified numerically:
combined counts, $N = 20$, $k = 6$:



Boson-sampling interferometry



Heisenberg advantage is immune to phase-errors!



Physical interpretation

Heisenberg interferometry is a huge advantage

- Appears feasible with large boson networks
- First practical application of boson sampling
- Complex-P allows calculation up to 100x100
- Confirms analytic conjecture by Motes et al
- **Over a trillion years with classical methods**
- May be applicable to BEC as well?

XY dynamics in a planar NDPO

Consider non-degenerate paramp in a planar cavity

- Many quantum correlated, entangled quadratures
- Possible candidate for a QC hardware
- Nonclassical correlations in many modes
- Wish to predict behaviour from first principles
- See: J. Opt. Soc. Am. B **33**, 871-883 (2016).

Preliminary analysis -

- What is the universality class?
- Unsqueezed quadratures are 2D XY model
- Tricritical Lifshitz point: like magnetic system
- Squeezed quadratures - unknown universality

Hamiltonian

$$\hat{H} = \hat{H}_{free} + \hat{H}_{int} + \hat{H}_{pump} + \hat{H}_{res} \,,$$

Free evolution and interaction Hamiltonian inside the planar cavity, pump ($i = 0$) plus two downconverted fields ($i = 1, 2$) with different polarizations, plus drive and damping:

$$\hat{H}_{free} = \sum_{i=0}^{2} \hbar \int d^2\mathbf{x}\, \hat{A}_i^\dagger \left[\omega_i - \frac{v_i^2}{2\omega_i}\nabla^2\right] \hat{A}_i \,.$$

$$\hat{H}_{int} = i\hbar \int d^2\mathbf{x} \left[\chi \hat{A}_0 \hat{A}_1^\dagger \hat{A}_2^\dagger - \chi^* \hat{A}_0^\dagger \hat{A}_1 \hat{A}_2\right] .$$

Boson fields obey: $\left[\hat{A}_i(\mathbf{x}, t), \hat{A}_j^\dagger(\mathbf{x}', t)\right] = \delta_{ij}\delta\left(\mathbf{x} - \mathbf{x}'\right).$

Exact positive P-representation equation

$$\hat{\rho} = \int d^{6M}\tilde{\boldsymbol{\alpha}} d^{6M}\tilde{\boldsymbol{\alpha}}^{+}\hat{\Lambda}\left(\tilde{\boldsymbol{\alpha}},\tilde{\boldsymbol{\alpha}}^{+}\right)P\left(\tilde{\boldsymbol{\alpha}},\tilde{\boldsymbol{\alpha}}^{+}\right),$$

Quantum problem mapped into an exact stochastic equivalent,

$$\begin{aligned}
\frac{\partial A_0}{\partial t} &= -\tilde{\gamma}_0 A_0 + \mathscr{E}(\boldsymbol{x}) - \chi^* A_1 A_2 + \frac{i v_0^2}{2\omega_0}\nabla^2 A_0\,,\\
\frac{\partial A_1}{\partial t} &= -\tilde{\gamma}_1 A_1 + \chi A_0 A_2^+ + \frac{i v_1^2}{2\omega_1}\nabla^2 A_1 + \sqrt{\chi A_0}\xi_1\,,\\
\frac{\partial A_2}{\partial t} &= -\tilde{\gamma}_2 A_2 + \chi A_0 A_1^+ + \frac{i v_2^2}{2\omega_2}\nabla^2 A_2 + \sqrt{\chi A_0}\xi_2\,,
\end{aligned}$$

Stochastic noise fields obey: $\langle \xi_1(\boldsymbol{x},t)\xi_2(\boldsymbol{x}',t')\rangle = \delta(\boldsymbol{x}-\boldsymbol{x}')\delta(t-t')$ with $\boldsymbol{A}\to\boldsymbol{A}^+$, $\boldsymbol{\xi}\to\boldsymbol{\xi}^+$ for hermitian conjugate fields. Construct from real noises using $\xi_{1,2}(\boldsymbol{x},t) = [\xi_x(\boldsymbol{x},t) \pm i\xi_y(\boldsymbol{x},t)]/\sqrt{2}$.

Critical point analysis

Steady-state: $A_1 A_2^* = A_1 A_2^* \left|\chi A_0\right|^2 / \left(\tilde{\gamma}_1 \tilde{\gamma}_2^*\right)$

Below threshold: $A_1 A_2^* = 0$

Above threshold: $\left|\chi A_0\right|^2 = \tilde{\gamma}_1 \tilde{\gamma}_2^*$. Both coincide, at critical pump intensity:

$$\left|\mathscr{E}_c\right|^2 = \bar{\gamma}^2 \left|\tilde{\gamma}_0 / \chi\right|^2 .$$

Define $\alpha_i = x_0 A_i$, $X = \sqrt{g}\left(\alpha_1 + \alpha_2^+\right), X^+ = \sqrt{g}\left(\alpha_2 + \alpha_1^+\right)$

Universal stochastic equations near threshold -

$$\frac{\partial \boldsymbol{X}}{\partial \tau} = \tilde{\mathscr{D}} \boldsymbol{X} - \left|\boldsymbol{X}\right|^2 \boldsymbol{X} + \tilde{\boldsymbol{\zeta}},$$

where $\tilde{\boldsymbol{\zeta}}$ is a noise vector and $\tilde{\mathscr{D}} = -\eta_1 + \eta_2 \nabla_{\boldsymbol{r}}^2 - \eta_3 \nabla_{\boldsymbol{r}}^4$ is a nonlocal derivative.

How does it get this simple?

Critical point adiabatic elimination

Near the critical point, the fluctuations in X are very slow, while Y responds on the fast relative time scale $1/\gamma$. We can drop terms of $\mathscr{O}(\sqrt{g})$ where $g \ll 1$, and approximate the equations as follows:

$$
\begin{aligned}
\frac{\partial X}{\partial \tau} &= -\left(\frac{1-\mu}{g}\right) X + \left(\frac{\Delta}{\sqrt{g}} - \nabla_{\boldsymbol{r}}^2\right) Y - X^2 X^+ + \zeta_+ \,, \\
\frac{\partial X^+}{\partial \tau} &= -\left(\frac{1-\mu}{g}\right) X^+ + \left(\frac{\Delta}{\sqrt{g}} - \nabla_{\boldsymbol{r}}^2\right) Y^+ - X^{+2} X + \zeta_+^* \,, \\
0 &= -(1+\mu)\, Y + \nabla_{\boldsymbol{r}}^2 X \,, \\
0 &= -(1+\mu)\, Y^+ + \nabla_{\boldsymbol{r}}^2 X^+ .
\end{aligned}
$$

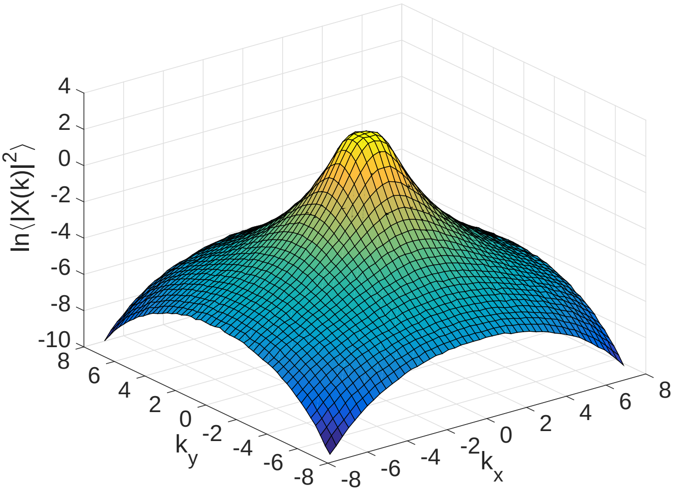
What happens to the other quadrature?

To lowest order in g, we can eliminate the fast or non-critical quadrature Y, Y^+-

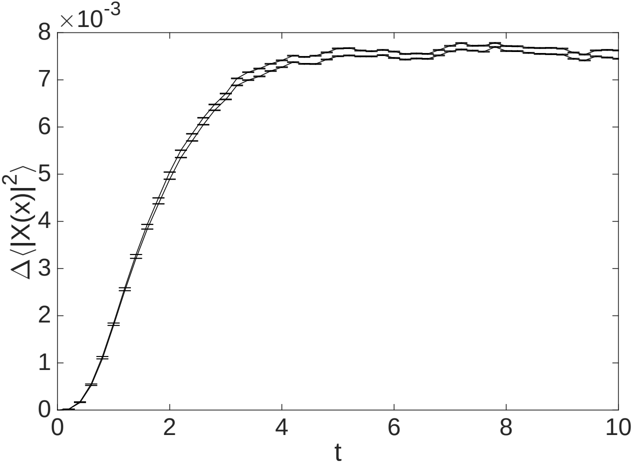
$$Y^{(+)} = \frac{\nabla^2 X^{(+)}}{1+\mu},$$

- Near threshold, the squeezed quadrature is slaved to the Laplacian of the the unsqueezed quadrature.
- Far from threshold this is not the case, and the squeezed quadrature can explore the phase-space in a way that influences the dynamics
- Quantum –> Classical transition as threshold point is approached

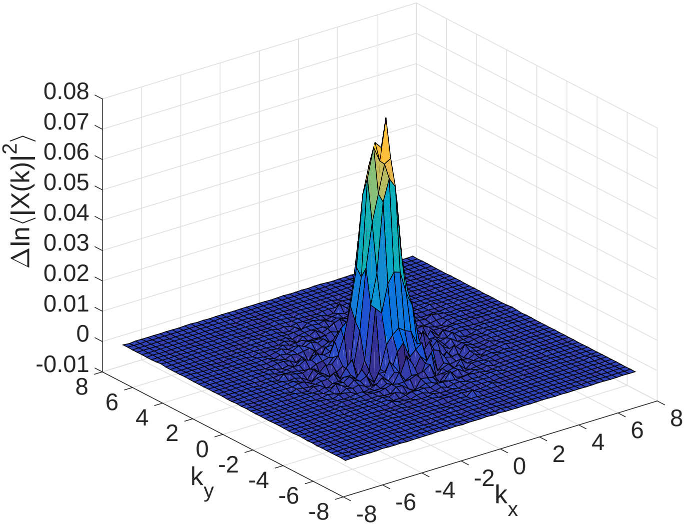
Power-law tails in k-space



Approach to equilibrium: non-Gaussianity



Non-Gaussian behaviour only for low k



Universality class

University class is a 2D nonlocal X-Y model:

- This is the free energy of a nonlocal planar magnetic interaction, with $\boldsymbol{X}$ playing the role of a two component vector order parameter. The planar type-II parametric system is a superb platform for investigating fluctuations and universal behavior.
- Like the Stanford Ising machine, it is a controllable nonequilibrium system, with potential QC applications.

Open question-

The Ising machine using a DPO, solves NP hard graph problems

Does the X-Y model solve an NP hard problem?

Kagome lattice ground state?

Does the quantum noise enhance solutions?

Unanswered questions

What is the optimal critical path?

- Near the critical point, we may be able to solve problems like a Kagome lattice.
- But this just looks like a classical equation.
- Unlikely to have quantum speedup
- Can we do better by starting far from critical, so quantum fluctuations are strong?

Hybrids: DPO vs NDPO

What is the difference to an OPO?

- Professor Mabuchi suggests one could combine DPO and NDPO.
- Can we change the detuning so this becomes like a DPO?
- Can this avoid being stuck in a non-optimal Ising ground state?

Conclusions

Promising new, highly diversified QC generation

- **Relativistic QFT simulation using ultra-cold atoms**
- Analysed using truncated Wigner simulations
- **Verification of Boson Sampling, novel metrology**
- Analysed using 'Quantum Software' complex P-functions
- **New idea: NDPO as a quantum computer?**
- Analysed with exact positive-P simulations