

QNNは（準）指数加速
しない

国立情報学研究所

松本啓史

概要

1. 新しい原理の計算機と現行型を比較するためのポイント

DWAVE、QNN、NISQ

2. 上記ポイントのQNNでのチェック

3. QNNの多項式時間 (BPP)シミュレーション

QNNで多項式時間で出来ることは、
乱数を使える古典計算機で
シミュレート できる

新しい原理の計算機と現行型を比較するための
ポイント

問題の複雑さ、計算時間とは？

たった一つの例題について、
その困難さ・易しさを語ることは無理

1. 理論を作る上での理由： その例題だけには正しい答えを瞬時に吐き出す仮想的機械は定義できてしまう
2. 現実：何がくるかわからないと、単純な課題でもより困難に何がくるか制限されていれば比較的複雑な課題もより対処しやすい

例 巡回セールスマン問題 (TSP)

1. 問題のサイズ ($n=1,2,3,\dots$)
2. 様々なタイプの都市の配置

「(ある) TMがTSPを解く」という時、
これらが違っていても同じTMで解かないといけない

問題の複雑さ、計算時間とは？

物理的な機械には、問題のサイズの上限はある
しかし、それは実用に際して十分大きいか
或いはサイズの拡張が技術的には自明でないとい
けない

(シャボン玉計算、粘菌計算 etc
ソーティングを目の子で、などなど)

問題の複雑さ、計算時間とは？

Some years ago, DWave announced optimization of **a** complicated function was done ultra-fast.

This is a weird claim to make.

Have to say something about a *family* of functions.

There can be *a* physical process which goes to the minima for *a* specific function, but it is not useful, nor make any sense in theory.

問題を「解く」

1. 最悪ケースの計算量、平均ケースの計算量

最悪ケースが標準的（例題の分布が謎）

最悪ケースで困難でも、問題のサンプルの仕方次第では平均ケースの計算量は著しく簡単になる

2. 厳密解 vs 近似解 （最悪ケース）

近似解の困難さ・易しさはメジャーな問題についてはよく調べられている
理論的限界 (量子でも多分破れない)

NP完全を帰着できる場合 （PCP定理、対話証明）

付加的な予想(unique game conjectureなど)に依存する場合

限界を達成するアルゴリズムあり

著しく簡単になる場合も多

問題が「解く」

今までの話はすべて保証付き

：どんな例題が来ても正しい答えを
要求された精度で出す。

(平均を考えるときは平均で保証)

ヒューリスティック

理論的保証ゼロ、経験的に手探り

1. 個々の問題の特性を使うもの
屢々神業

2. 何にでも使える (meta heuristic)

Simulated Annealing, Genetic Algorithm

NP困難な最適化問題はなぜか人気

最適化問題

関数 $f(x)$ の最大化または最小化

拘束条件つき

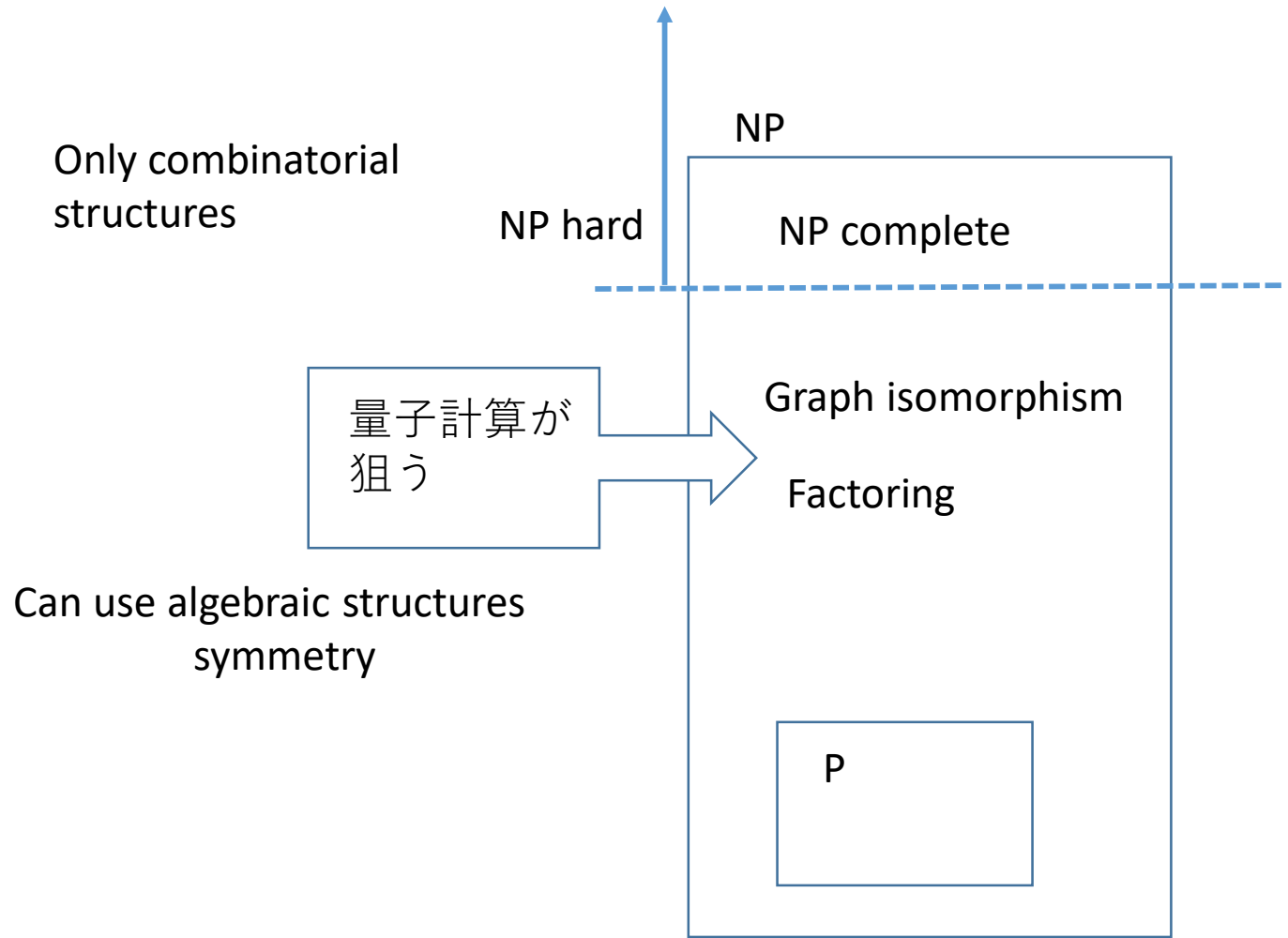
デジタルで扱うなら x は結局は離散

ぱっと見た目はわかりやすい

エネルギー最小化や適応度最大化とのアナロジー

NP困難

全てのNP問題がそれに多項式時間で
帰着されるような問題。



組合せ最適化問題で古典計算機に挑もう

最悪ケースで「保証付き」の場合：

メジャーな問題では、達成可能な精度の限界が出そろっていて、かつ古典で達成されている
しかも、量子でも破れないと思われている

特殊な例題、平均計算量：

これらは屢々古典計算機でも非常に簡単になる

ヒューリスティック

個別の問題に特化した古典のヒューリスティックが有効な場合は、非常に厳しい。

例　：　Google instance

このような状況を踏まえ、組合せ最適化で古典計算を破ることを示すにはどうしなければならないか？

組合せ最適化問題で古典計算機に挑もう

個別のヒューリスティックが効く問題は怖い

→DwaveのGoogle instance

SAに一億倍差をつけるも、
個別のヒューリスティックに完敗

よって、

作戦1. 「個別に特化は邪道」という論理展開をしよう

(古典のメタヒューリスティック研究者もよく使う。
ある意味では正しい。)

作戦2. アルゴリズムの工夫が効きにくい問題を狙う

(=部分が全体と「緩く」結合している)

昨年のQNN 完全グラフのMaxCUT

(DWaveだと結合がローカルなので難しいか)

(QNNでも大規模化したときに全結合できるか)

量子計算機の理論的優位性が示された状況との違い

- 最悪かつ厳密解の計算量という、理論的に割合とはっきりとした場合で示すことが出来た
- 最適化問題の近似解は理論がはっきりとしているところは
 簡単すぎる 又は 量子でもお手上げ

よって、理論的によく解明されていない場合へと赴かざるを得ない

実験的な比較で問題になりがちなこと

1 例題の生成方法

下手に作ると簡単な例題ばかりに (DWaveの初期のころ)
これは確立された手法がある。

(最近のDwave, QNNはここは抜かりが無い)

2. 比較対象の古典アルゴリズム

3. 比較対象のハードウェア

4. メリットを保ったままスケールアップできるか？

5. 本当に「実用的」？

「理論的にはわからないが実用的」を標榜する以上、

それ相応の厳しい吟味が必要

理論計算機科学の吟味から逃れるための方便としてはいけない

実用的な問題、規模、前後のインターフェース

別の問題に帰着して解くときは、その帰着の手間

比較対象のハードウェア

非在来型のものは

最適化問題のメタヒューリスティック専用機

比較対象も似た性質のものが自然

しかし実際は汎用機と比較

スケールアップの可否

小規模でないと難しいことを用いていないか

- デジタル vs アナログ

- 並列 & 全結合

QNNの場合はどうか

Let us listen to hrk先生, a specialist of `conventional' architecture...



hrk先生

@Prof_hrk

twitterの世界では、現場踏み込みで知られています。某オケのVn弾き。本業は速いインターネットとか。非公開アカウント、TweetのないRead Onlyアカウントはブロックします。

平木 敬 (ひらき けい)



職名 教授

所属 情報理工 情報科学科

メールアドレス

hiraki@is.s.u-tokyo.ac.jp

居室 理7号館4階406号室

電話番号 03-5841-4039
24039 (内線)



h r k先生 @Prof_hrk

2017-11-21 17:16:54

このコンピュータを超短く要約すると、[4]の光量子コンピュータは量子ビット数がスケールアップできないので、[3]で相互作用を量子ビット間ではなく、外部に出す方法を具体化し、1Kmファイバに実質2048量子ビットを循環させ、FPGAで相互作用を外部的に実現した。ファイバ上には実質ビット以外も乗っている



h r k先生 @Prof_hrk

2017-11-21 18:01:05

本文と図で計算式が違うのがご愛敬だけれど、本文が正しそう。これは2値ニューラルネットの計算式そのもの。

これで、ファイバ上の量子入出力がノイズレスならば、単なるBNN実装。
(査読者しっかりしろ！)

BNN=Binary Neural Network



h r k先生 @Prof_hrk

2017-11-21 18:06:22

つまり、FPGAによるBNN回路に擾乱を与えるノイズ源として光量子リングを用いたものが、今回公開する量子コンピュータであると私は判断した。周波数はリング1周が5 μ sなので、400MHzで量子ビットが到着し、FPGAは2Kの内部メモリを同時に読み出し、パイプライン的に積和する回路と思われる。



h r k先生 @Prof_hrk

2017-11-21 18:08:18

計算機屋としての疑問は、FPGAにDRAMのシフトレジスタを接続し、途中に適当な物理乱数回路を入れれば同じものが実現するのではないかということ。([2]によれば、答えはYESだだろう)。何かD-Waveと日立のCMOS Isingマシンの関係を思い起こす。

Annealing ? ニューラルネット ? 次に解説

Simulated Annealing (simplest version)

Annealing = 焼きなまし

$$H(x^1, x^2, \dots, x^n) = \sum_{i,j} J_{ij} x^i x^j - \sum_i h_i x^i \quad x^i = \pm 1$$

Set $T > 0$ properly, and repeat the following until certain conditions are met:

1 For $i = 1$ to n do

If $H(x^1, \dots, \bar{x}^i \dots) < H(x^1, \dots, x^i \dots)$, flip the value of x^i

Otherwise, flip x^i with the probability

$$\exp\left\{-\frac{1}{k_B T} H(x^1, \dots, \bar{x}^i \dots)\right\} / \exp\left\{-\frac{1}{k_B T} H(x^1, \dots, x^i \dots)\right\}$$

2 $T \rightarrow$ smaller (Scheduled so that $T \doteq 0$ at the end)

If $T=0$, gradient method

If T is fixed, a standard algorithm to generate canonical distribution (Metropolis)

Simulated Annealing = gradient + noise =noisy neural network

If $T=0$, gradient method

$$H(x^1, x^2, \dots, x^n) = \sum_{i,j} J_{ij} x^i x^j - \sum_i h_i x^i$$

$$H(\bar{x}^1, \dots, x^n) - H(x^1, \dots, x^n) = (\bar{x}^1 - x^1) \left\{ \sum_j J_{1j} x^j - h_1 \right\}$$

$$x^1 \leftarrow 2\Theta \left(\sum_j J_{1j} x^j - h_1 \right) - 1 \quad \Theta: \text{Heaviside function}$$

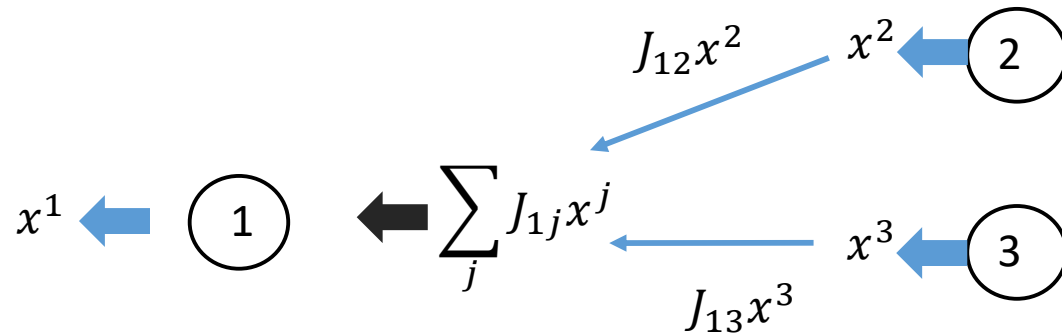
Neural network (Amari-Hopfield network)

J : connection h : threshold

If $T>0$, gradient + noise

$$x^1 \leftarrow 2\Theta \left(\sum_j J_{1j} x^j - h_1 \right) - 1 + w_{1,T,t}$$

Simulated Annealing as nosy AH-network

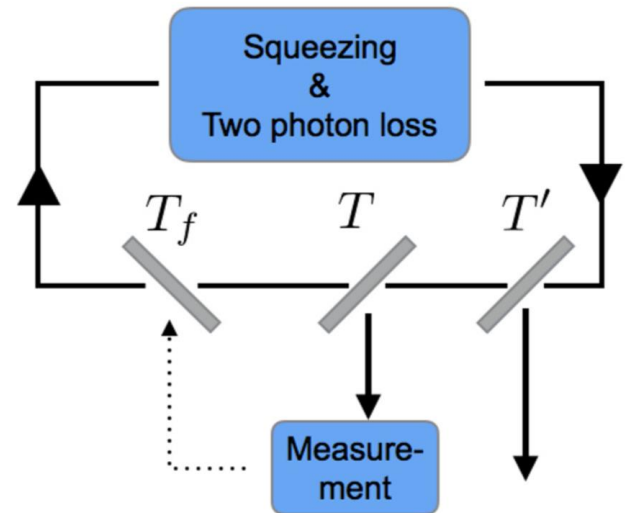
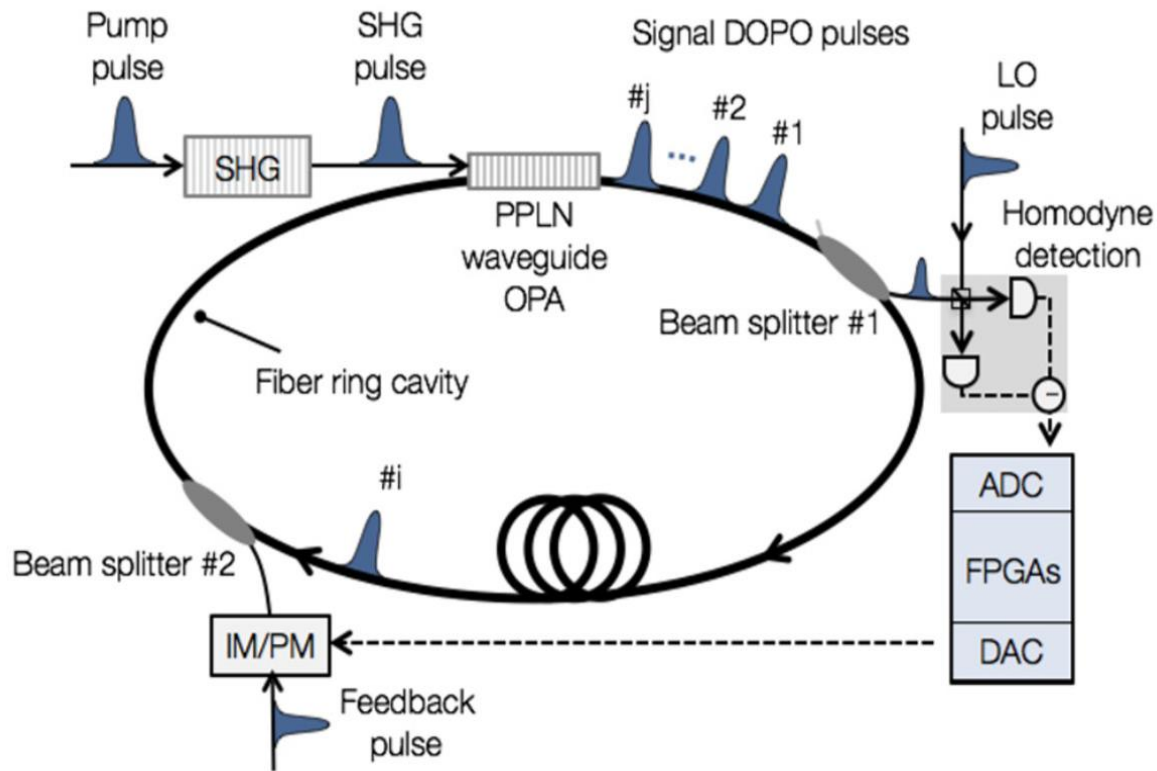


If $\sum_j J_{1j}x^j \geq h_1$, $x^1 = 1 + w_{1,t,T}$

← : output

← : input

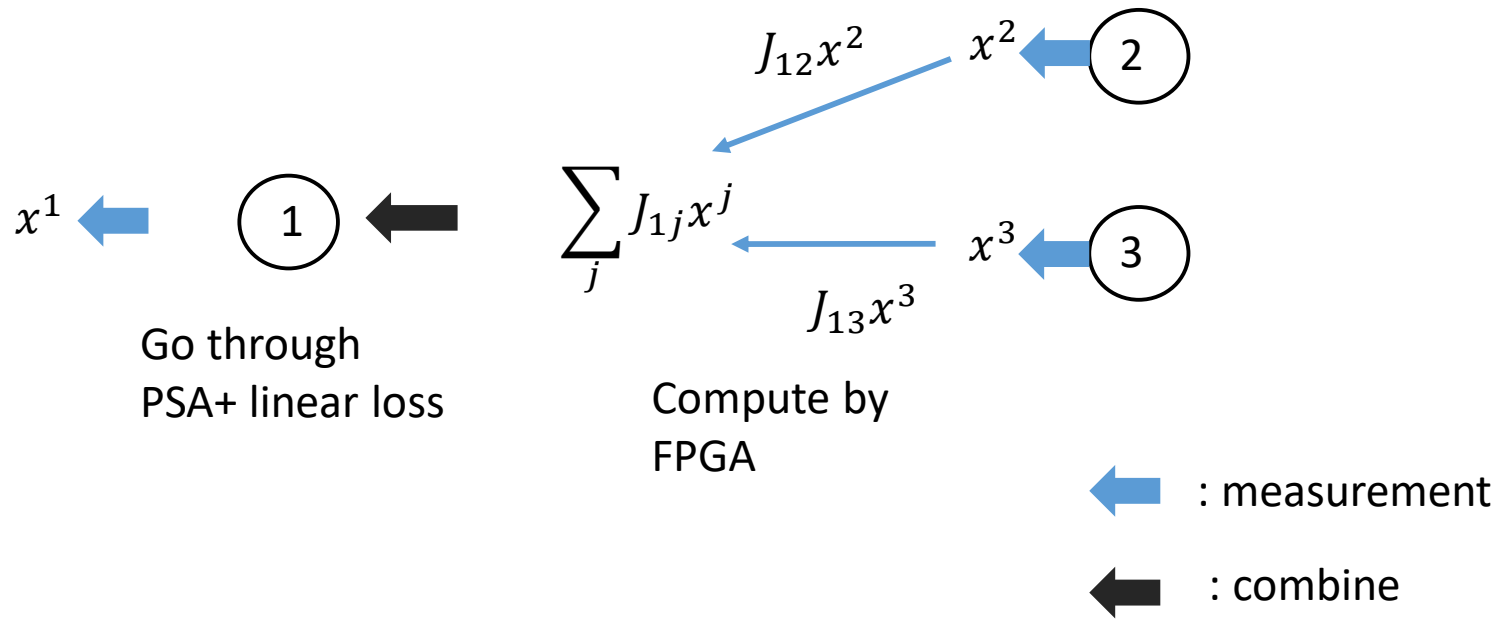
How QNN works



A. Yamamura, K. Aihara, and Y. Yamamoto,

Quantum coherent Ising machines: Discrete-time measurement feedback formulation, Phys. Rev. A 96 053834 (2017)

QNN as a neural network



1. Each pulse is functioning as an analogue register
2. No quantum interaction btw pulses (neurons)
3. Computation of the input is most essential part = done by FPGA
(other parts are independent of the problem to solve)

hrk先生 @Prof_hrk

2017-11-21 18:01:05

本文と図で計算式が違うのがご愛敬だけれど、本文が正しそう。これは2値ニューラルネットの計算式そのもの。

これで、ファイバ上の量子入出力がノイズレスならば、単なるBNN実装。

(査読者しっかりしろ！)

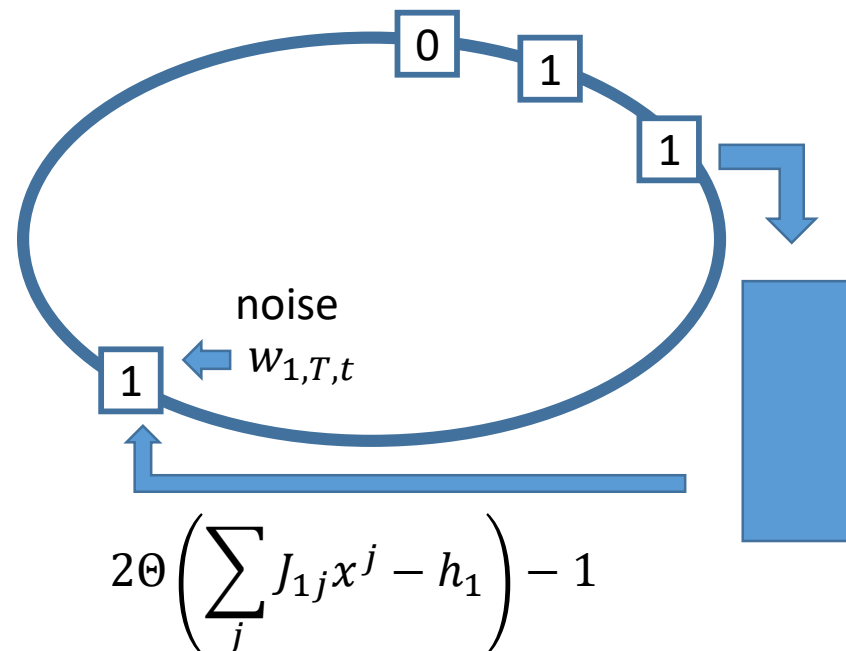
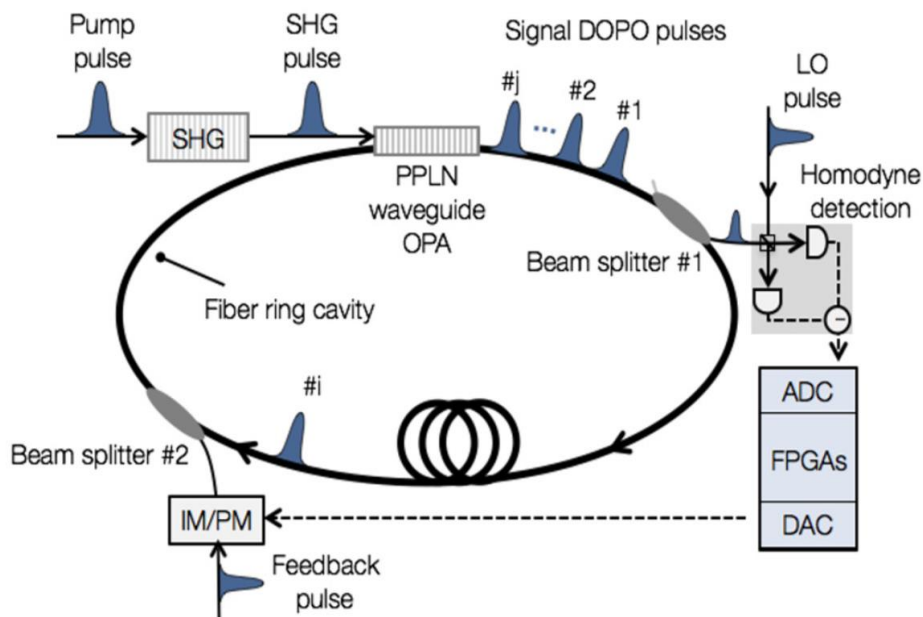
BNN=Binary Neural Network



つまり、FPGAによるBNN回路に擾乱を与えるノイズ源として光量子リングを用いたものが、今回公開する量子コンピュータであると私は判断した。周波数はリング1周が $5\mu\text{s}$ なので、 400MHz で量子ビットが到着し、FPGAは2Kの内部メモリを同時に読み出し、パイプライン的に積和する回路と思われる。



計算機屋としての疑問は、FPGAにDRAMのシフトレジスタを接続し、途中に適当な物理乱数回路を入れれば同じものが実現するのではないかということ。([2]によれば、答えはYESだだろう)。何かD-Waveと日立のCMOS Isingマシンの関係を思い起こす。



Their comparison of QNN and classical

1. Run simulated annealing on a *single* CPU.

recall FPGA was taking care of the core part

How much FPGA helps ?

2. compared physical time

clearly, choice of the hardware is not appropriate

Thus I think they should compare in terms of

of times of use of neurons (pulses)

= how many times $\sum_j J_{1j} x^j - h_1$ is computed

= how many times “the cycle” goes round

If we compare in terms of this measure, which wins ?

I asked these points to one of the project team, who kindly answered these points frankly.

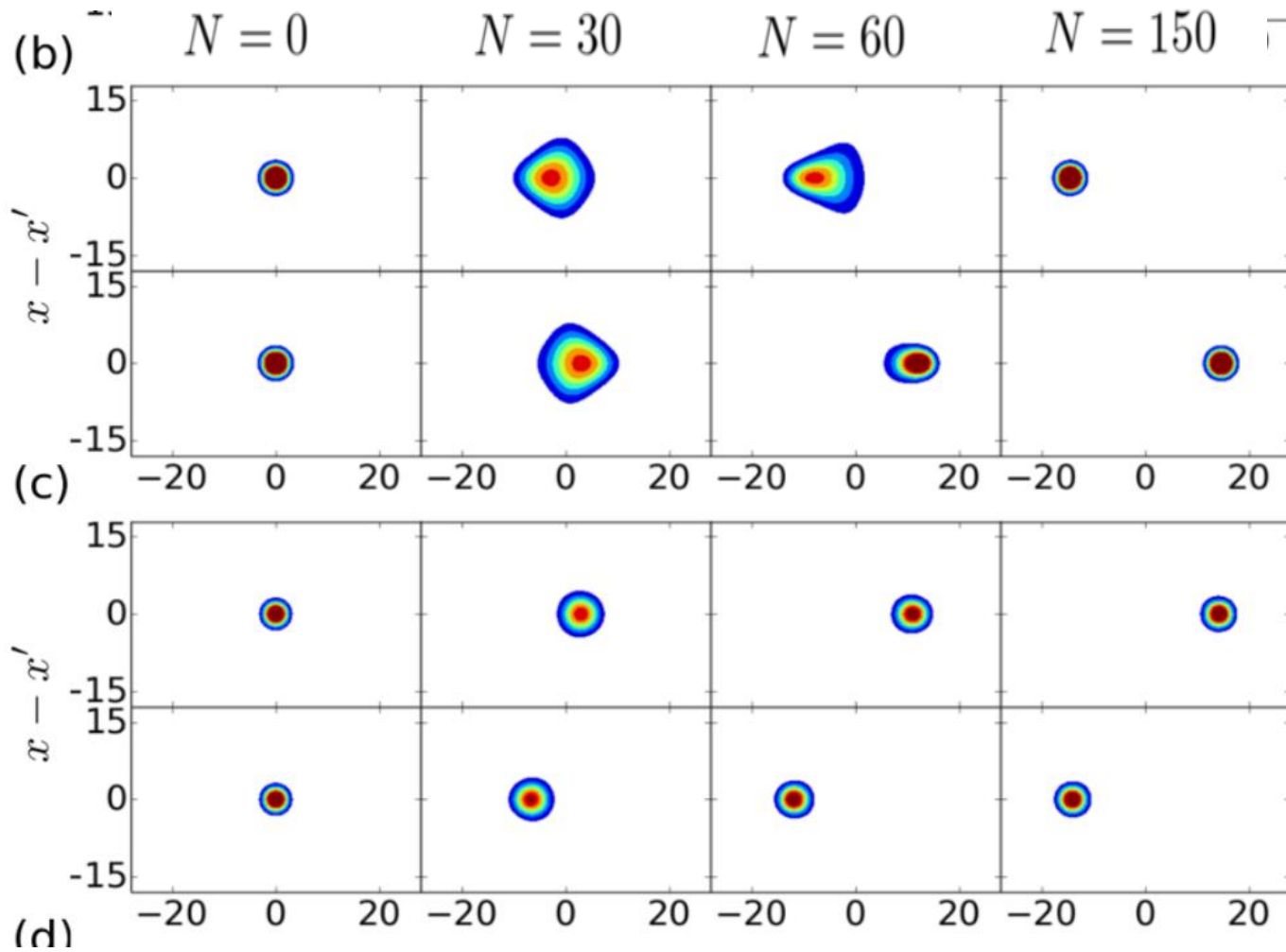
現段階で、QNNが現行型計算機より有効であるとする実験的な証拠はないと発表者は考える

QNNでは、常に状態はテンソル積状態
量子的な相互作用はない
いかなる非局所性も働かない

では、個々のパルスの中での時間発展はどうか？
仮に量子的でも、非局所性がないゆえに、あまり意味はないが、念のために見てみる

Time evolution when the 2 pulses comprises the system

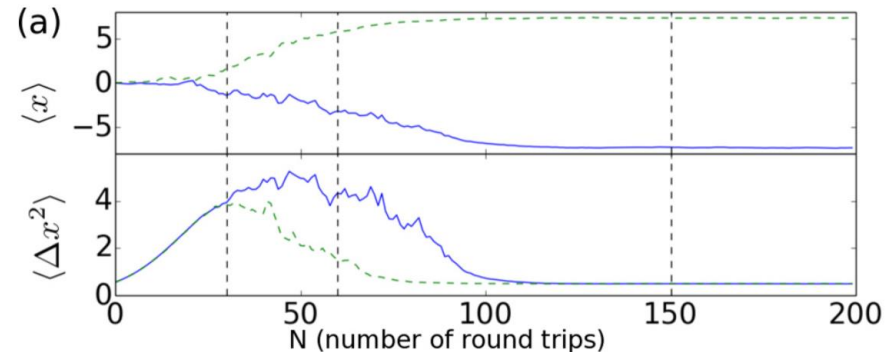
in front of the PSA



個々のパルスの時間発展の特徴

1. 実験がなされているパラメーターの値の範囲では、Wigner関数は正で丸いまま、左右に動く。非常に古典的に見える。彼らの論文も同様に述べている
2. 測定の強さを1/10に弱めても少し歪むだけ。
(光ファイバーの伝送損失は無視した計算)

3. $x=10, -10$ のところに安定点があり、波束が $x=0$ 付近にあるかなり早い段階でどちらに収束するかが決まる。



4. 2ビットの局所最適が無い場合でも、真値に収束する確率は良くて8割ほど。
(実際にはやっていない微妙なポンプレート調整をする)

しかしそれでも

しかしそれでも、

- もっと改良すれば
- 一つ一つの量子性も大事かも

など言えないとも限らない。

そこで、理論的に考えて見る。

解く問題の計算機科学的考察から得ることは少ない。

そこで、機械の動作そのものを抑える。

QNNの多項式時間シミュレーション

以下のシミュレーションは何を目指してるか

経験的に効率的なシミュレーションはある

Positive P functionで微分方程式に

T. Shoji, K. Aihara Y. Yamamoto, Phys Rev A96 053833 (2017)

その計算時間を数値実験で出して、スケーリングを外挿してみせることもできる

しかし、精度と計算時間を理論的につめていないので、反論の余地を残す

精度と計算時間を理論的に詰めるのが目的

個数基底

精度が保障出来るアルゴリズムのみ使う

Motivation & implication

Sets a complexity theoretic upper bound to computational ability of QNN

Namely, **impossibility of exponential speed up**

Polynomial in

1. $1/\text{error}$
2. number of the pulses
3. number of the rounds

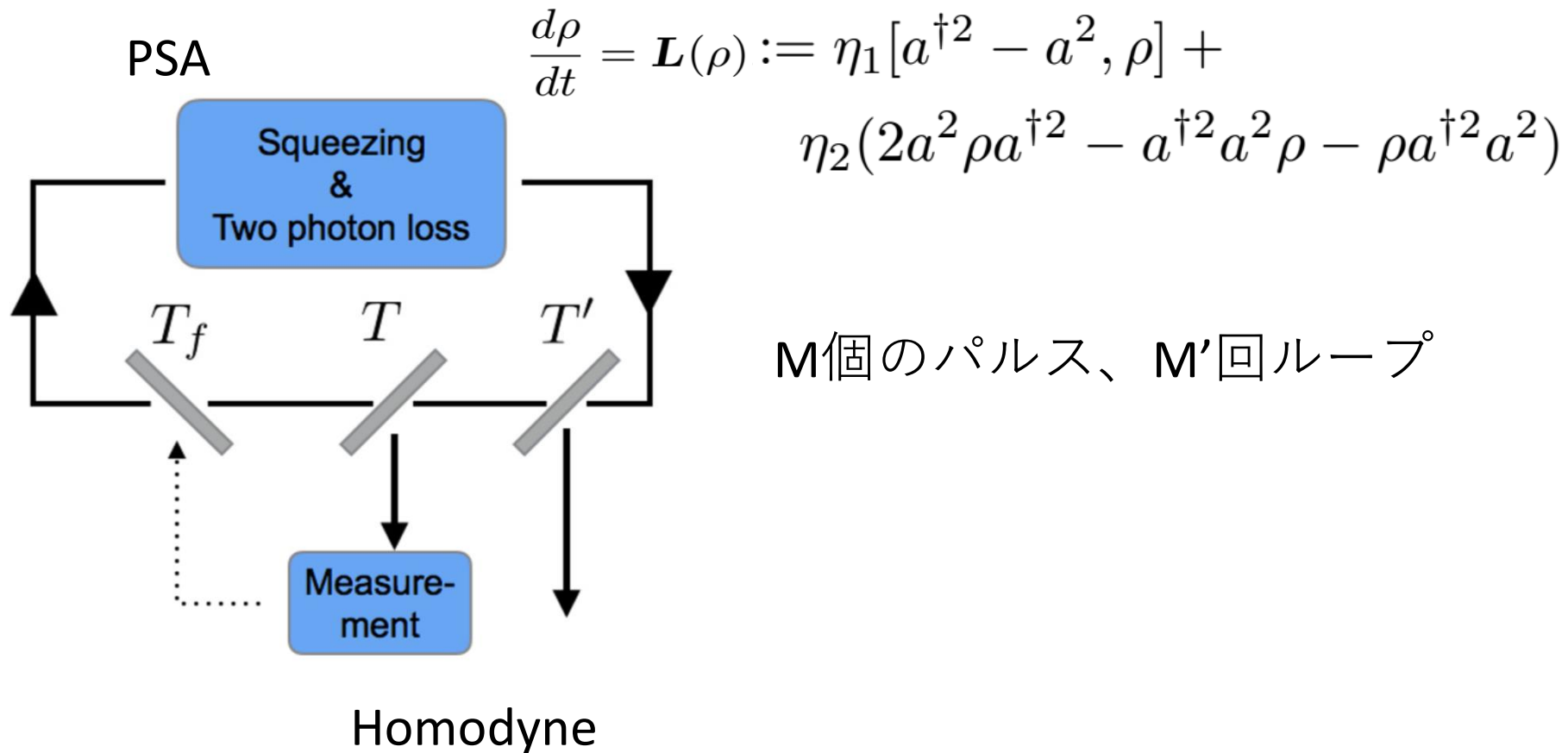
Caution :

- Not limited to worst case analysis.
- Does not deny polynomial speed up
- Does not claim QNN can be as fast as this upper bound

How QNN works

A. Yamamura, K. Aihara, and Y. Yamamoto,

Quantum coherent Ising machines: Discrete-time measurement feedback formulation, Phys. Rev. A 96 053834 (2017)



シミュレーションの概要

観測値を決めるとテンソル積状態

$H^{\otimes M}$ の上の密度演算子のかわりに、
 H 上の密度演算子を M 個と観測値の確率を計算

個数基底で展開し、 L までで止める
 L はせいぜい、 M, M' の多項式

微分方程式は差分化
観測はモンテカルロ（確率の小さな事象は捨てる）

丸め誤差は桁数を M, M' の多項式だけ保持すれば
誤差が指数的に小。多項式回溜まってもまだ十分小。

How to bound the cut-off size L

Compute an upper bound to $\langle e^{sN} \rangle$

s : arbitrary positive real

N : number operator

$\langle \rangle$: average

By Chebyshev ineq.

$$\text{tr} (I - P_L)\rho \leq \exp\{-s(L - \ln\langle e^{sN} \rangle)\}$$

P_L : projection onto the span of $\{|n\rangle\}_{n=0}^L$

We will show :

$$\ln\langle e^{sN} \rangle \leq O(M^2 + M \ln M') =: L_s$$

M : number of rounds

M' : number of pulses

By this, if $L := 2L_s$, $\text{tr} (I - P_L)\rho$ is exponentially small

How 'small' $\langle e^{sN} \rangle$ has to stay ?

$$\text{tr} (I - P_L)\rho \leq e^{-sL} \langle e^{sN} \rangle$$

P_L : projection onto the span of $\{|n\rangle\}_{n=0}^L$

L : we want this to be $\text{poly}(M, M')$

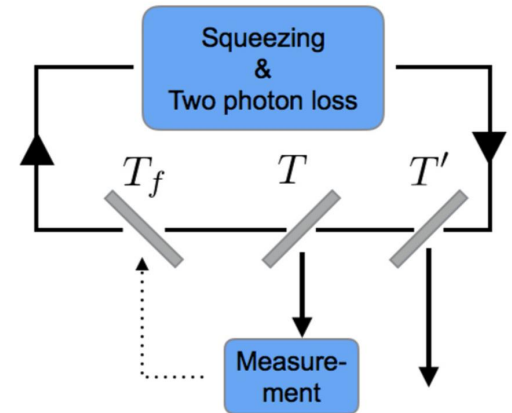
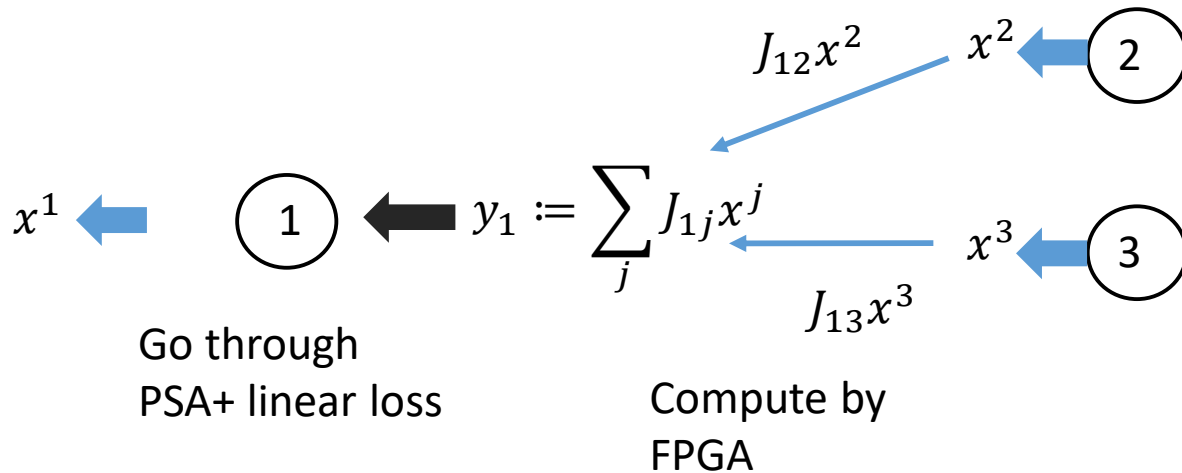
M : number of rounds

M' : number of pulses

$\therefore \ln \langle e^{sN} \rangle$ has to be bounded by $\text{poly}(M, M')$

How to bound $\langle e^{sN} \rangle$?

パルス間の相互作用があるから、全体を同時に扱わないといけない？



パルスはmeasurement-feedback だけで相互作用
Feedbackは複素振幅が y のcoherent光を合波
 $|y|$ にはレーザーのパワーからくる上限 y_{\max} あり

$$\max_{|y_1| \leq y_{\max}} \langle e^{sN} \rangle$$

を議論すれば、他のパルスからの入力は何であったか忘れてよい

Evaluation of $\langle e^{sN} \rangle$:PSA

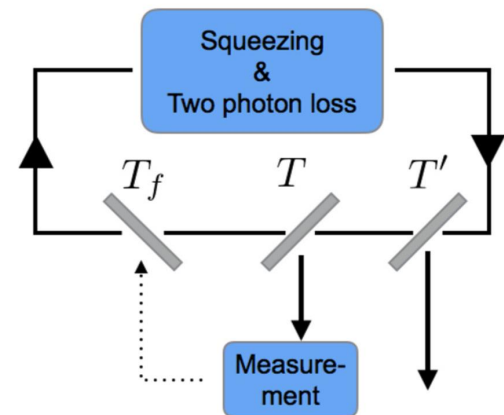
$$\frac{d\rho}{dt} = \mathbf{L}(\rho)$$

$$\mathbf{L}(\rho) := \eta_1[a^{\dagger 2} - a^2, \rho] + \eta_2(2a^2 \rho a^{\dagger 2} - a^{\dagger 2} a^2 \rho - \rho a^{\dagger 2} a^2)$$

$\langle e^{sN} \rangle$ is not increasing if it is above certain threshold

Use analogous statement about polynomial of N, which is the main result of:

R. Azouit, S. Sarlette, P.Rouchen, Convergence and adiabatic elimination for a driven dissipative quantum harmonic oscillator. arXiv:1503.06324v1



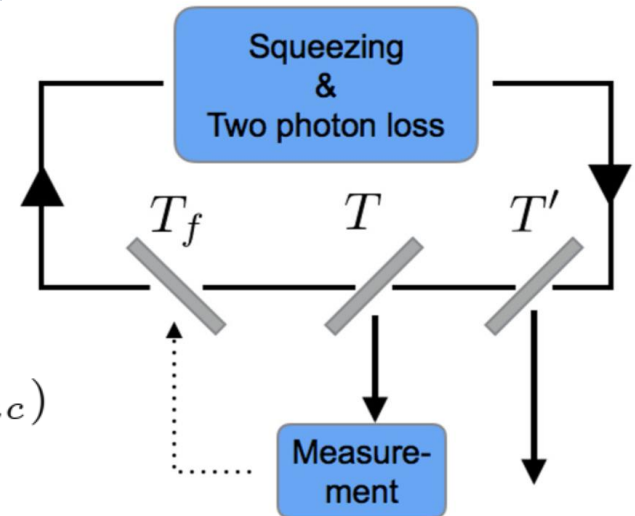
Evaluation of $\langle e^{sN} \rangle$: measurmer†

Measurement:

$N \otimes I_{vac} + I \otimes N_{vac}$ はBSの前後で保存

$$\begin{aligned} \text{tr } \rho e^{sN} &= \text{tr } U \rho \otimes |0\rangle \langle 0| U^\dagger e^{s(N \otimes I + I \otimes N_{vac})} \\ &\geq \text{tr } U \rho \otimes |0\rangle \langle 0| U^\dagger e^{sN \otimes I} \\ &= \int dx_m \text{tr } I \otimes \langle x_m | U \rho \otimes |0\rangle \langle 0| U^\dagger I \otimes |x_m\rangle e^{sN} \\ &= \int dx_m \text{tr } W_{x_m} \rho W_{x_m}^\dagger e^{sN} \\ &= \int dx_m p(x_m) \text{tr } \rho_{x_m} e^{sN} \\ &\geq \lambda \text{tr } \rho e^{sN} \Pr\{\text{tr } \rho_{x_m} e^{sN} \geq \lambda \text{tr } \rho e^{sN}\}. \end{aligned}$$

$$\Pr\{\text{tr } \rho_{x_m} e^{sN} \geq \lambda \text{tr } \rho e^{sN}\} \leq \frac{1}{\lambda}$$

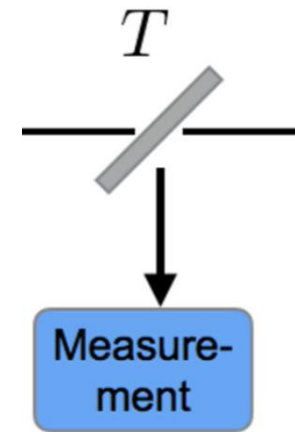


Evaluation of $\langle e^{sN} \rangle$: measurement

Measurement:

$$\Pr\{\text{tr } \rho_{x_m} e^{sN} \geq \lambda \text{tr } \rho e^{sN}\} \leq \frac{1}{\lambda}$$

この上限は確率は状態の形に依存せず、閾値 λ のみで決まる定数



M回のループの後、M'個のパルスが $\{ \}$ 内の事象を満たす確率は

$$\geq \left(1 - \frac{1}{\lambda}\right)^{MM'}$$

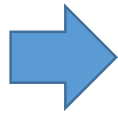
$$\lambda := (MM')^{1+\delta_1} \text{ とすると、} \geq \left(1 - \frac{1}{(MM')^2}\right)^{MM'} \approx \exp\{(MM')^{-\delta_1}\} \approx 1$$

$$\therefore \langle e^{sN} \rangle_{after} \leq \lambda \langle e^{sN} \rangle_{old} \quad \text{として扱う}$$

Evaluation of $\langle e^{sN} \rangle$: phase shift

$$\text{tr } \rho \exp(s a^\dagger a) = e^{-s} \int \frac{d^2\alpha}{\pi} \exp((1 - e^{-s}) |\alpha|^2) \langle \alpha | \rho | \alpha \rangle$$

入力 = x



$$e^{-s} \int \frac{d^2\alpha}{\pi} \exp((1 - e^{-s}) |\alpha + x|^2) \langle \alpha | \rho | \alpha \rangle$$

$|x| \leq y_{\max}$

$$\leq g(s, M) \text{tr } \rho \exp\left(s + \frac{1}{M}\right) N$$

$$g(s, M) \leq \exp \{y_{\max}^2 + 8(e^s - 1)M + M^{-1}\}$$

再帰的にこの式を使う上で、赤字が s やその他に依存しない定数であることは重要

Evaluation of $\langle e^{sN} \rangle$: 総合

$$\ln \langle e^{sN} \rangle \leq \sum_{k=0}^{M-1} \ln g(s + kM^{-1}, M) + M \ln \lambda + \max\{\ln \langle e^{(s+1)N} \rangle_{init}, (s+1)\mu\}$$

$$\sum_{k=0}^{M-1} \ln g(s + kM^{-1}, M) \leq M y_{\max}^2 + 8M^2(e^{s+1} - e^s - 1) + 1$$

$$\begin{aligned} \ln \langle e^{sN} \rangle &\leq M y_{\max}^2 + 8M^2(e^{s+1} - e^s - 1) + 1 + 2M \ln MM' \\ &\quad + \max\{\ln \langle e^{(s+1)N} \rangle_{init}, (s+1)\mu\} \\ &= O(M^2 + M \ln M') \end{aligned}$$

$$\text{tr} (I - P_L) \rho \leq e^{-sL} \langle e^{sN} \rangle \quad P_L: \text{projection onto the span of } \{|n\rangle\}_{n=0}^L$$

$$\therefore L_S = O(M^2 + M \ln M')$$

How cut-off effects total accuracy

Key : 1. triangle inequality

2. CPTP does not amplify the error

$(\prod_{i=1, \dots, t} \Lambda_i)(\rho)$: process to be simulated

$$\mathbf{P}_L(\rho) := P_L \rho P_L$$

$$\|\mathbf{P}_L \circ \prod_{i=1, \dots, t} (\Lambda_i \circ \mathbf{P}_L)(\rho) - (\prod_{i=1, \dots, t} \Lambda_i)(\rho)\|_1 \leq (t + 1)3e^{-\frac{\epsilon}{2}(L - L_s)}$$

最後尾から一つ一つ

カットオフしたもので入れ替える

$L_s = O(M^2 + M \ln M')$, $L = 2L_s$ と取っておけば t は $\text{poly}(M, M')$ までOK.

以降は、 $L \times L$ 行列だとして扱う

How to simulate PSA

$$\frac{d\rho}{dt} = \mathbf{L}(\rho) \quad \mathbf{L}(\rho) := \eta_1[a^{\dagger 2} - a^2, \rho] + \eta_2(2a^2 \rho a^{\dagger 2} - a^{\dagger 2} a^2 \rho - \rho a^{\dagger 2} a^2)$$

時間発展を刻み h ($=1/\text{poly}(M, M')$)で刻む

(刻み過ぎるとカットオフの誤差が溜まる)

$\exp h\mathbf{L}(\rho)$ を k 次のテーラー展開で近似 (k: 定数)

$$\left\| \exp(h\mathbf{L})(\rho) - \sum_{i=0}^{k-1} \frac{(h\mathbf{L})^i}{i!}(\rho) \right\|_1 \leq \frac{b^k}{k!} h^k (L + k)^{2k} \quad \begin{array}{l} b \text{は上の微分方程式から} \\ \text{きまる定数} \end{array}$$

$1/h \times M \times M'$ 回繰り返し溜まっても誤差 ϵ_1 を超えぬよう十分細かく h をとる

$$\begin{aligned} h^{-1} &:= \frac{b^{\frac{k}{k-1}}}{\epsilon_1^{1/(k-1)}} (L + k)^{\frac{2k}{k-1}} (MM')^{1/(k-1)} \\ &\leq (2b)^2 \epsilon_1^{-\delta} L^{1+\delta} (MM')^\delta \end{aligned} \quad \begin{array}{l} k \text{が十分大きい時} \end{array}$$

(気分的には $\sim L$)

How to simulate PSA

$$\mathbf{L}(\rho) := \eta_1[a^{\dagger 2} - a^2, \rho] + \eta_2(2a^2 \rho a^{\dagger 2} - a^{\dagger 2} a^2 \rho - \rho a^{\dagger 2} a^2)$$

$a \dots a a^{\dagger} a |n\rangle \langle n'| a \dots a a^{\dagger} a$ の計算を $O(L^2)$ 回やる
各項の計算は、 k の定数倍の時間かかる

また、 $\ln(1/\epsilon)$ だけ桁数をとっている

全てを併せると、1 ラウンドあたり
 $\epsilon^{-\delta} L^{3+\delta} (MM')^{\delta} \ln \epsilon^{-1}$

観測

$$\rho \rightarrow \frac{1}{\text{tr } W_{x_m} \rho W_{x_m}^\dagger} W_{x_m} \rho W_{x_m}^\dagger$$

$$\langle x | W_{x_m} | n \rangle = \pi^{-\frac{1}{4}} \langle tx + rx_m | n \rangle \exp\left(-\frac{1}{2}(rx - tx_m)^2\right)$$

$$\langle x | n \rangle: \text{エルミート多項式} \times e^{-\frac{1}{2}x^2} \quad (r, t: \text{反射率、透過率})$$

$\int \langle n' | x \rangle \langle x | W_{x_m} | n \rangle dx$ は解析的に求まり、上の行列要素、確率も求まる
エルミート多項式の項数を考えて、各項計算時間は $L^3 \ln(1/\epsilon)$
行列要素はそれの L^4 倍
確率は (tr なので) L^3 倍

$$\therefore L^8 \ln(1/\epsilon)$$

$$x_m \sim \text{tr } W_{x_m} \rho W_{x_m}^\dagger \quad \text{となる乱数を一様分布から変換}$$

累積頻度 = 確率関数の積分をリーマン和で近似
方程式を数値的にニュートン法で解いて変換

$$\text{累積頻度の勾配} = \text{密度、} \quad \text{累積頻度の計算} \quad L^7(1/\epsilon) \ln^2(1/\epsilon)$$

Phase shift

$\langle n|D(y)|m\rangle$: Laguerre 倍多項式で書ける (項の数: $m-n \leq L$)

行列の要素	L
行列の掛け算	L^3
桁数	$\ln(1/\varepsilon)$
$\not\propto$	$L^4 \ln(1/\varepsilon)$

(phase space ではずらすだけなのになぜだ)

まとめると 1 ラウンドあたり

PSA $L^{3+\delta} (MM')^\delta \epsilon^{-\delta} \ln \epsilon^{-1}$

観測 $L^8 \epsilon^{-1} \ln^2 \epsilon^{-1}$

Shift $L^4 \ln \epsilon^{-1}$



これが MM' 回

そして

$$L = O(M^2 + M \ln M')$$

合計何次でしょう . . . ?

でもまあ、多項式

しかしそもそも

上の解析では、複素振幅が一度増幅したらそれが減衰する効果は全然考慮していない。

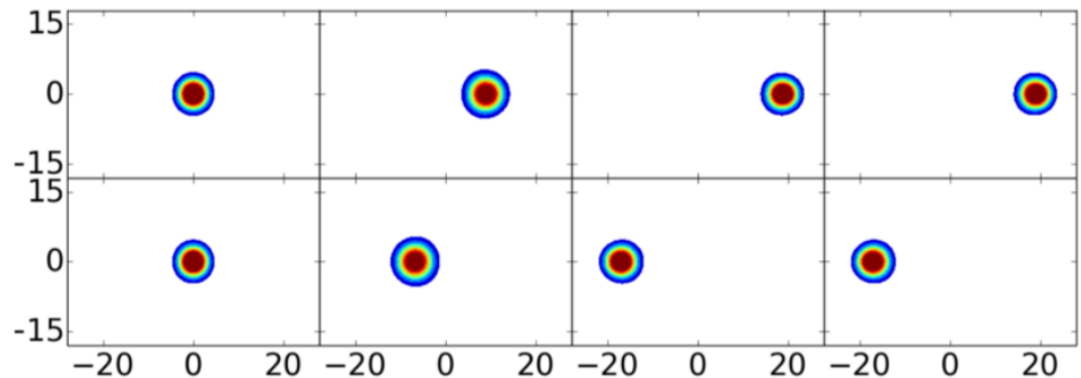
平均光子数は多項式的に増加してるかもしれない設定

でも、そんなことは多分、起こっていない

Phase space の実軸場の、二つの安定点の間を行ったり来たり

そもそも、遠くに飛んだら計算は失敗のはず

失敗した場合もフォローするために無駄に評価を損している



そこで、もしLを定数または緩やかな増加 ($\log M$ など) でいいとすれば、次数はぐっと下がる

そもそも「量子計算」の定義とは？

「量子性」「古典性」という言葉は結構曖昧

究極的には全ての物は量子なので、どういう意味での量子化をはっきりさせないといけない

意見 1

使用目的に応じて様々な「量子性」の意味がある

意見 2

物理的に明快な「量子」の定義がある

などなど。。。。

まとめ

不毛な議論で失礼しました

Simulation by classical computer: outline

1. Approximate the density op of each pulse by $L \times L$ matrix (in photon number basis)
 - possible as the amplitude does not grow too much

2. Compute time evolution

PSA : 微分方程式を差分近似

注意 : $\|d\rho/dt\|_1 \sim L^2$ L: cut-off

Lは多項式的に増加→刻みも増加

観測 :

Measurement: throw away low-prob events

- since they add up for each round, they should be proportional to $1/(\text{times of repetition})$

Simulation : how can it be polynomial time ?

Given M of L -level systems, the total system is L^M dimensional : notoriously large !!

But,

the system is always in a **tensor product** state, conditioned by the measurement results

(this is much stronger than separable)

No quantum interaction. Thus each pulse can be simulated separately.

$$L^M \times L^M \Rightarrow L \times L \times M$$

Quantumness is just inside each pulse

\Rightarrow No exponential speed up

Evaluation of $\langle e^{sN} \rangle$, phase shift

$$\text{tr } \rho \exp(s a^\dagger a) = e^{-s} \int \frac{d^2\alpha}{\pi} \exp((1 - e^{-s}) |\alpha|^2) \langle \alpha | \rho | \alpha \rangle$$

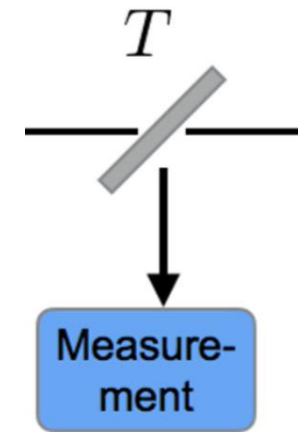
$$\begin{aligned} \text{tr } D(x) \rho D^\dagger(x) \exp(s a^\dagger a) &= e^{-s} \int \frac{d^2\alpha}{\pi} \exp((1 - e^{-s}) |\alpha + x|^2) \langle \alpha | \rho | \alpha \rangle \\ &\leq e^{s y_1} \text{tr } \rho \exp(s(1 + \frac{y}{2y_0}) a^\dagger a) \end{aligned}$$

y_0 : arbitrary positive

$$s(y^2 + 2y_0 + y/(2y_0)) =: s y_1$$

Evaluation of $\langle e^{sN} \rangle$, measurements

$$\Pr\{\text{tr } \rho_{x_m} e^{sN} \geq \lambda \text{tr } \rho e^{sN}\} \leq \frac{1}{\lambda}.$$



$$\begin{aligned} \text{tr } \rho e^{sN} &= \text{tr } U \rho \otimes |0\rangle \langle 0| U^\dagger e^{s(N \otimes I + I \otimes N_{vac})} \\ &\geq \text{tr } U \rho \otimes |0\rangle \langle 0| U^\dagger e^{sN \otimes I} \\ &= \int dx_m \text{tr } I \otimes \langle x_m | U \rho \otimes |0\rangle \langle 0| U^\dagger I \otimes |x_m\rangle e^{sN} \\ &= \int dx_m p(x_m) \text{tr } \rho_{x_m} e^{sN} \\ &\geq \lambda \text{tr } \rho e^{sN} \Pr\{\text{tr } \rho_{x_m} e^{sN} \geq \lambda \text{tr } \rho e^{sN}\} \end{aligned}$$

Chebyshev's ineq.

How to bound the cut-off size L

Compute an upper bound to $\langle e^{sN} \rangle$

s : arbitrary positive real

N : number operator

$\langle \rangle$: average

$$\text{tr} (I - P_L)\rho \leq e^{-sL} \langle e^{sN} \rangle$$

P_L : projection onto the span of $\{|n\rangle\}_{n=0}^L$

There is $L_s = O(M \ln MM')$

M : number of rounds

M' : number of pulses

s.t. $\langle e^{sN} \rangle \leq e^{sL_s}$ holds in the whole process
with constant probability

After the cut-off: squeezing + two photon loss

$$\left\| \rho' - \mathbf{P} \circ \left[(\mathbf{I} + h \mathbf{L} + \frac{1}{2} h^2 \mathbf{L}^2)(\rho) \right]_{i,j} \right\|_1 \leq C' h^3 (L + 2)^8$$

$$\begin{aligned} h^3 \sup_{t \in [0, h]} \left| \left[\left(\frac{d^3 \rho}{dt^3}(t) \right) \right]_{i,j} \right| &\leq h^3 \sup_{t \in [0, h]} \left| [\mathbf{L}^3(\rho(t))]_{i,j} \right| \\ &\leq h^3 \sup_{\sigma} \left| [\mathbf{L}^3(\sigma)]_{i,j} \right| \leq h^3 7^3 c_1^3 (L + 2)^6 \end{aligned}$$

1/h of them to simulate single process

$$h = \frac{\varepsilon_1^{1/2}}{C'^{1/2} (L + 2)^4 (MM')^{1/2}}$$

$$\begin{aligned} \text{Total time} & O\left(L^7 (MM')^{1/2} \varepsilon_1^{-1/2} \log \varepsilon_1^{-1}\right) \\ \text{(for the single)} & = O\left(M^{7.5} (M')^{0.5} \log MM' \varepsilon_1^{-0.5} \log \varepsilon_1^{-1}\right) \end{aligned}$$

In fact, h affects L

Since $1/h$ of infinitesimal process is necessary to simulate single one;

But the effect is log :

Since the error by cut-off is $e^{-s(L-L_s)}$ at each step,

$$\frac{1}{h} e^{-s(L-L_s)} = e^{-s(L-L_s + s^{-1} \log h)}$$

Measurement and phase shift

Each matrix components of measurement POVM and phase shift operation are computed efficiently (in terms of error).

Matrix multiplication: $O(L^3 \log \varepsilon^{-1})$

Generating random number

for given cumulative distribution function:

$$O(\varepsilon^{-2} \log \varepsilon^{-1})$$

Summing up:

Most costly part is now squeezing + two photon loss

$$O(M^{8.5}(M')^{1.5}\log MM' \varepsilon_1^{-0.5}\log\varepsilon_1^{-1})$$

Why it takes so much ?

Since the cut-off L grows in proportion to M, M'

But is it really necessary?

Numerical results seem to indicate Wigner function stays around between two fixed points.

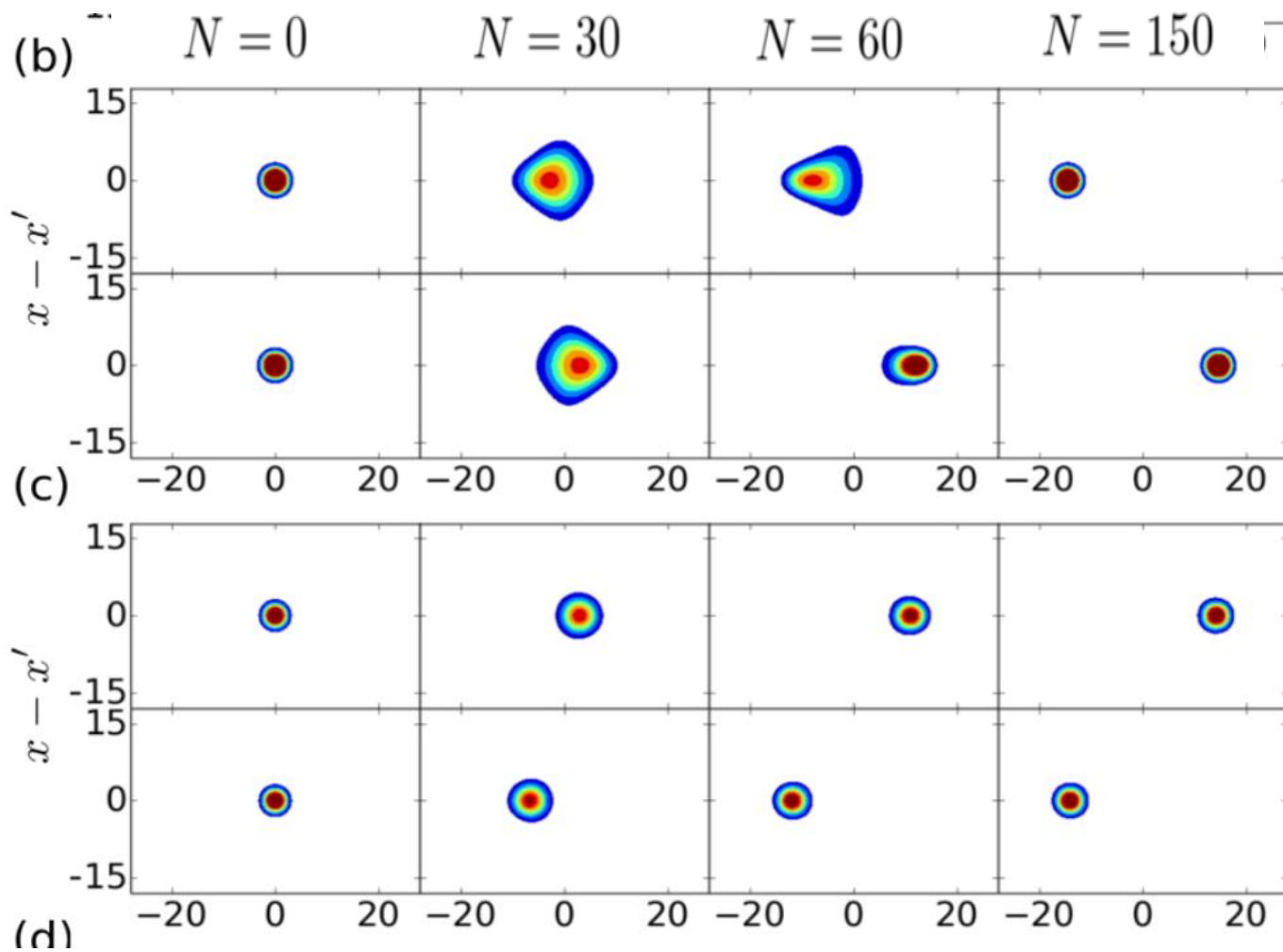
If one can assume this,

$$L = O(\log MM')$$

c.f. Time for simulating squeezing+ two photon loss

$$O\left(L^7 (MM')^{1.5} \varepsilon_1^{-1/2} \log \varepsilon_1^{-1}\right)$$

in front of the PSA



Precision might be excessive

In this simulation method, error is kept small enough that it does not accumulate throughout the run

However, CIM in real world is not free from accumulation of error

Smaller but possible improvement : squeezing + two photon loss

$$\begin{aligned} h^3 \sup_{t \in [0, h]} \left| \left[\left(\frac{d^3 \rho}{dt^3}(t) \right) \right]_{i,j} \right| &\leq h^3 \sup_{t \in [0, h]} \left| [\mathbf{L}^3(\rho(t))]_{i,j} \right| \\ &\leq h^3 \sup_{\sigma} \left| [\mathbf{L}^3(\sigma)]_{i,j} \right| \leq h^3 7^3 c_1^3 (L+2)^6 \end{aligned}$$

The last inequality is too crude :

(upper bound to each component) \times number of matrix elements