

高速な量子機械学習の 理論基盤の構築

山崎隼汰

東京大学理学系研究科物理学専攻助教

2023年8月3日

<https://hayatayamasaki.com/> hayata.yamasaki@gmail.com @hayatayamasaki

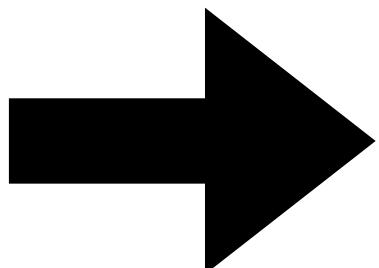
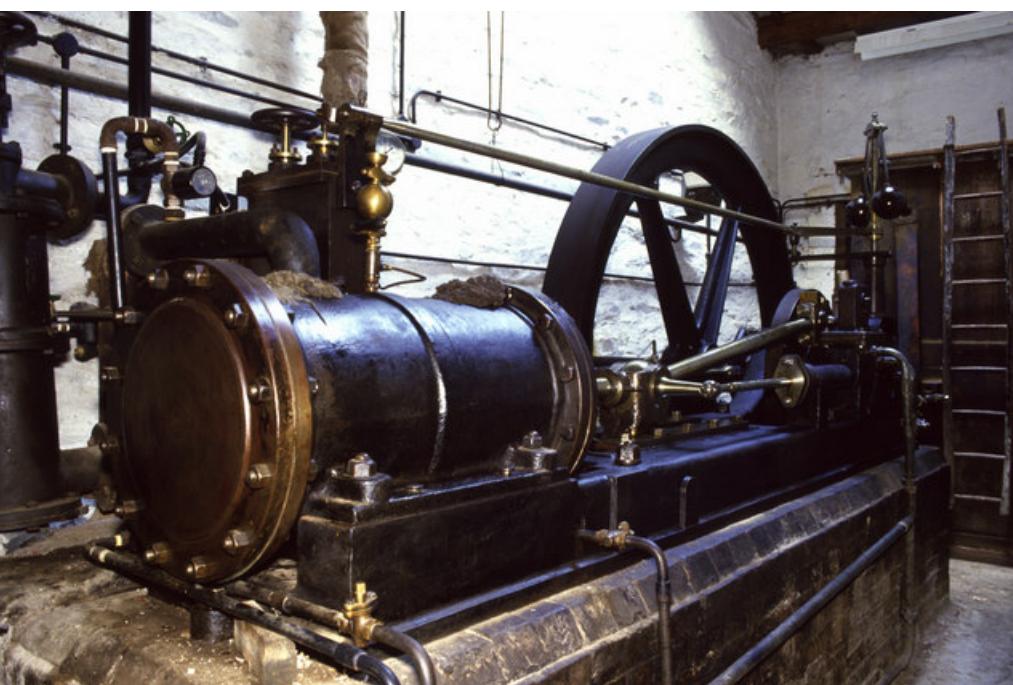
量子情報理論の研究分野

量子力学 = ミクロな自然現象を記述できる普遍的な物理法則

19世紀

蒸気機関

熱力学



技術進歩

21世紀

量子デバイス

量子情報



[© RIKEN Center for Quantum Computing]

- 量子力学に現れる特有の性質は計算に有用 = 量子コンピュータ
- 計算機科学の理論的枠組みは量子力学の定量的理解に有用

量子の世紀に我々は何を経験し得るか、何を経験し得ないか？

私の研究

社会実装

量子技術による
将来的な情報化社会の発展

有用な量子アルゴリズム

高速で適用範囲の広い
量子機械学習アルゴリズム

理論的基礎
＝私の研究

量子計算の実行方法

低オーバーヘッドでスケーラ
ブルな誤り耐性量子計算

効率的な量子操作

量子計算資源（リソース）の
定量的解析

実験の基礎

実験における量子技術の
発展

高速で適用範囲の広い量子機械学習

[arXiv:2004.10756](#),
[arXiv:2106.09028](#),
[arXiv:2301.11936](#),...

**効率的な
誤り耐性量子計算**

[arXiv:1911.11141](#),
[arXiv:2111.07952](#),
[arXiv:2207.08826](#),...

**量子計算リソースの
定量的な解析・検証**

[arXiv:2002.02458](#),
[arXiv:2201.11127](#),
[arXiv:2202.13131](#)...

<https://www.hayatayamasaki.com/>

量子計算の実行方法の基盤構築 (4-11)

高速な量子機械学習のためのアルゴリズム開発 (12-22)

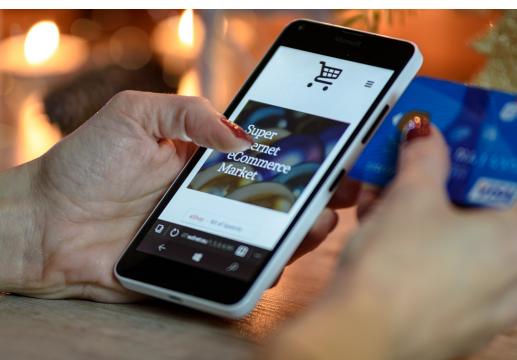
量子コンピュータを何に使うか

量子計算は以下のような計算問題を解く際に大幅な高速化を達成すると信じられている

- ・ 素因数分解 (Shorのアルゴリズム) … 数学
- ・ 量子力学の数値シミュレーション… 物理・化学

他にも様々な量子アルゴリズムがある: <https://quantumalgorithmzoo.org/>

機械学習に基づく情報化社会



日々の暮らしでのデータ処理 ↔ 背後の機械学習技術



量子機械学習 → さらなる高速化・大規模化を目指す

量子機械学習: より直接社会的にインパクトがある問題への応用可能性がある

量子アルゴリズム

量子アルゴリズムは、確率分布の代わりに**負の値（や複素数）を取りうる確率振幅**（2乗が確率分布）が活用できる乱択アルゴリズム

単位 ビット: $\{0, 1\}$ 量子ビット: $\mathbb{C}^2 = \text{span} \{ |0\rangle, |1\rangle \}$

量子状態: $|\psi\rangle = \alpha_0 |0\rangle + \alpha_1 |1\rangle \in \mathbb{C}^2$ 複素ベクトル（確率分布ベクトルの量子版）

- Normalized

$$\sum_j |\alpha_j|^2 = 1$$

- Ignoring global phase

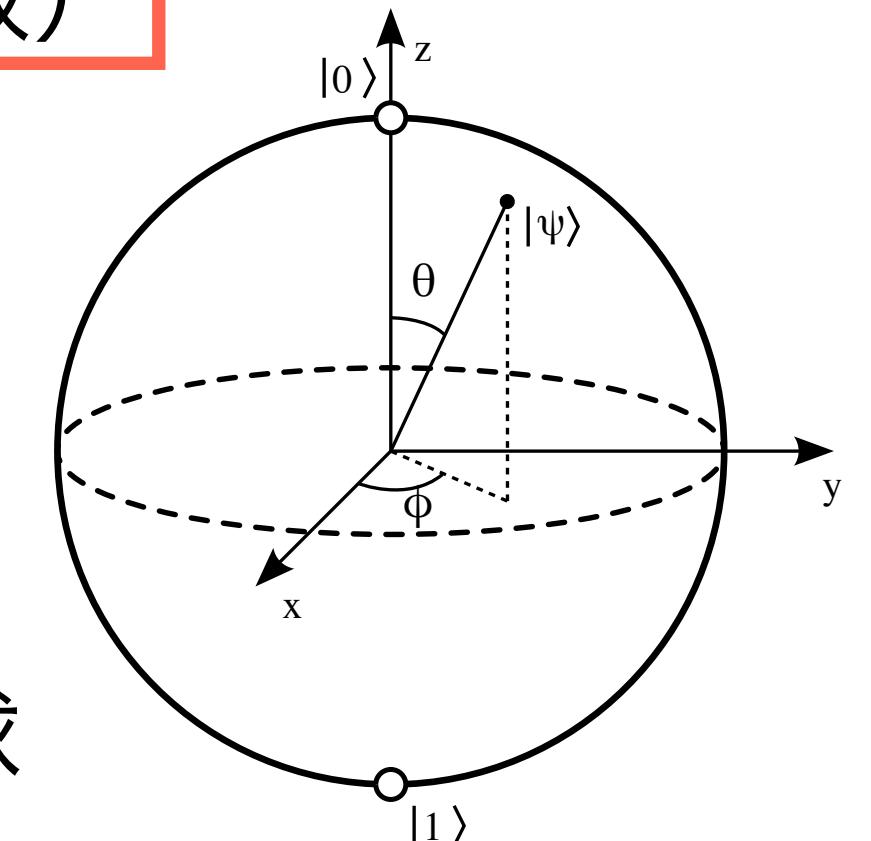
$$|\psi\rangle \sim e^{i\varphi} |\psi\rangle$$

$$|\psi\rangle = \cos\left(\frac{\theta}{2}\right) |0\rangle + e^{i\phi} \sin\left(\frac{\theta}{2}\right) |1\rangle$$

N 量子ビット: Nビット列の状態

$\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \cdots \{ |x\rangle : x \in \{0, 1\}^N \}$ が正規直交基底となる空間

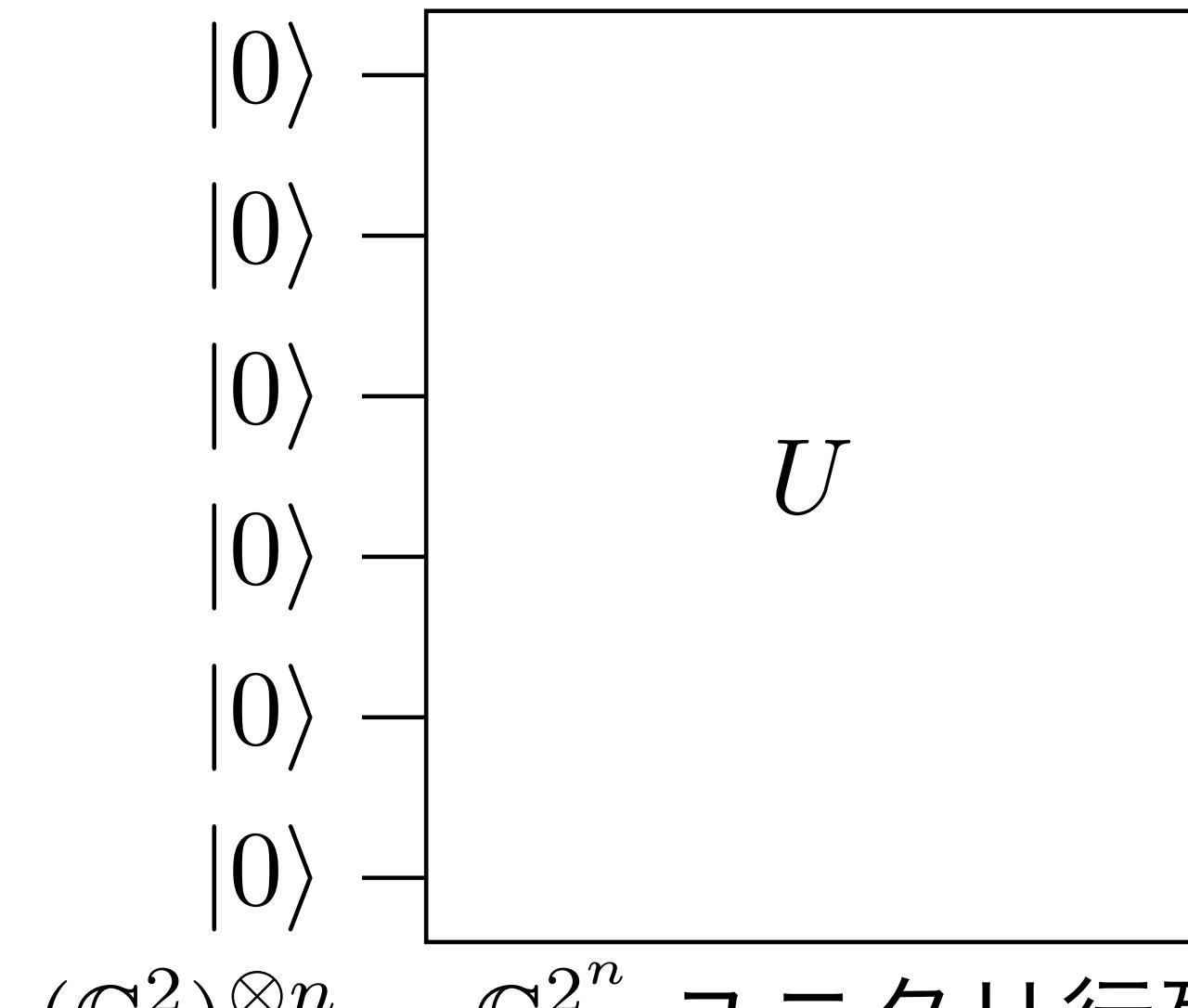
プロッホ球



例：2量子ビットの状態

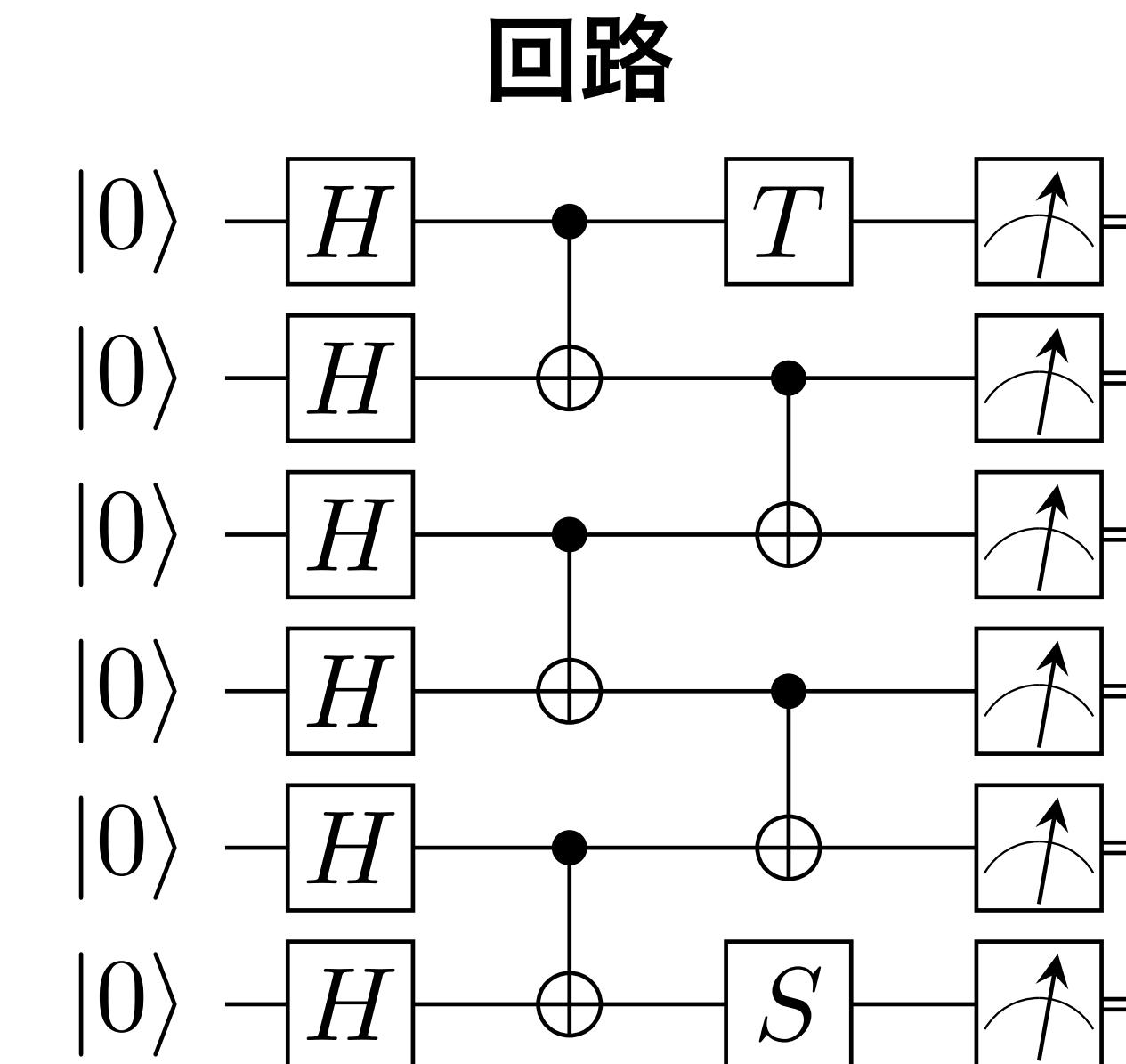
$$\alpha_{00} |00\rangle + \alpha_{01} |01\rangle + \alpha_{10} |10\rangle + \alpha_{11} |11\rangle$$
 確率振幅は複素数

量子回路



0 or 1

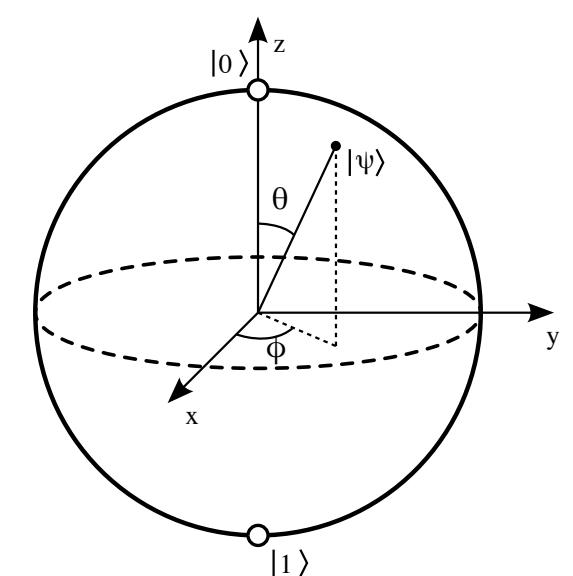
分解



測定 (measurement, next slide)

確率分布=確率振幅の2乗

エンタングルした状態
 $U |0\rangle^{\otimes n}$



量子コンピュータ=各操作を高精度で実行できるデバイス

ゲートセット

$$H = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

$$S = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}$$

$$T = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\frac{\pi}{4}} \end{pmatrix}$$

$$CNOT = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$$CZ = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

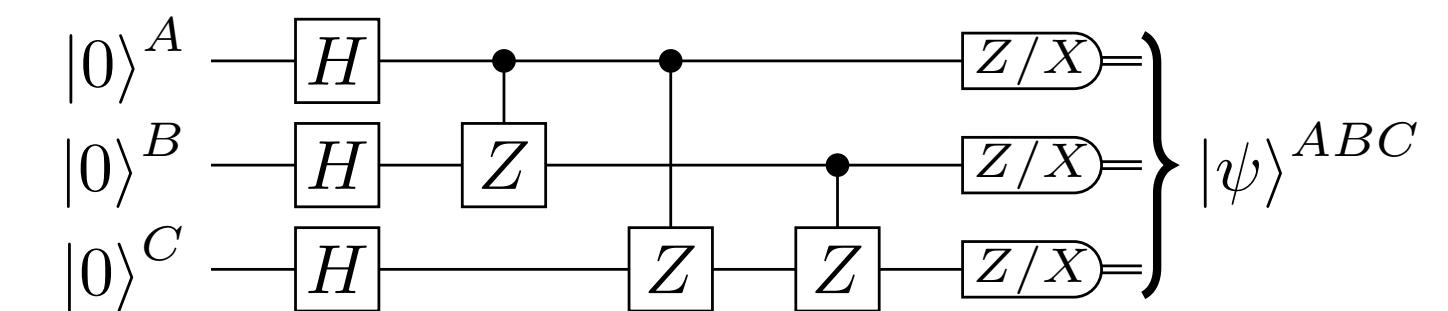
量子測定

$$\alpha_0 |0\rangle + \alpha_1 |1\rangle \xrightarrow[\text{重ね合わせ}]{\{|0\rangle, |1\rangle\}} \begin{cases} p(0) = |\alpha_0|^2 \\ p(1) = |\alpha_1|^2 \end{cases} \text{確率振幅の2乗が確率分布}$$

エンタングルメント $|\psi\rangle^{ABC} = \frac{1}{2\sqrt{2}}(|000\rangle + |001\rangle + |010\rangle + |100\rangle - |011\rangle - |101\rangle - |110\rangle + |111\rangle)$

Z 基底 $\{0,1\}: \{|0\rangle_Z, |1\rangle_Z\}$

X 基底 $\{\pm\}: \{|j\rangle_X = H|j\rangle_Z : j = 0, 1\} = H$ ゲートをかけて測定

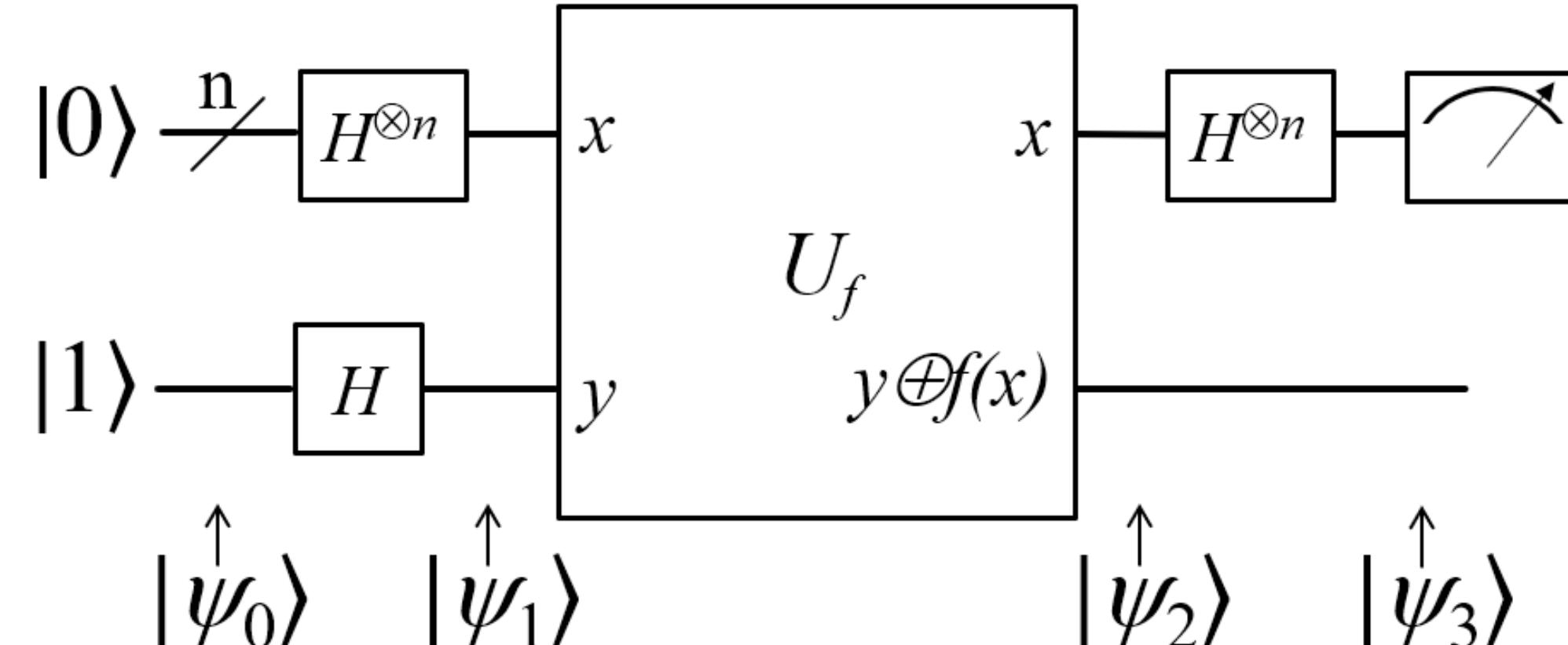


- 量子状態を測定すると確率振幅の2乗で与えられる確率分布 $p(x)$ で答えが確率的に出力される
- エンタングルした量子状態の確率振幅の干渉を古典計算で効率的にシミュレートする方法は知られていない
= 量子高速化のための計算資源（リソース）
[Bravyi, Gosset, Koenig, arXiv:1704.00690](#)

量子計算 = 負（や複素数）の確率振幅による打ち消しあい（干渉）使える乱択アルゴリズム

量子計算のやり方

1 : 量子ビットの状態の初期化



2 : 入力データを量子回路のパラメータとして読み込み、対応する量子状態を生成

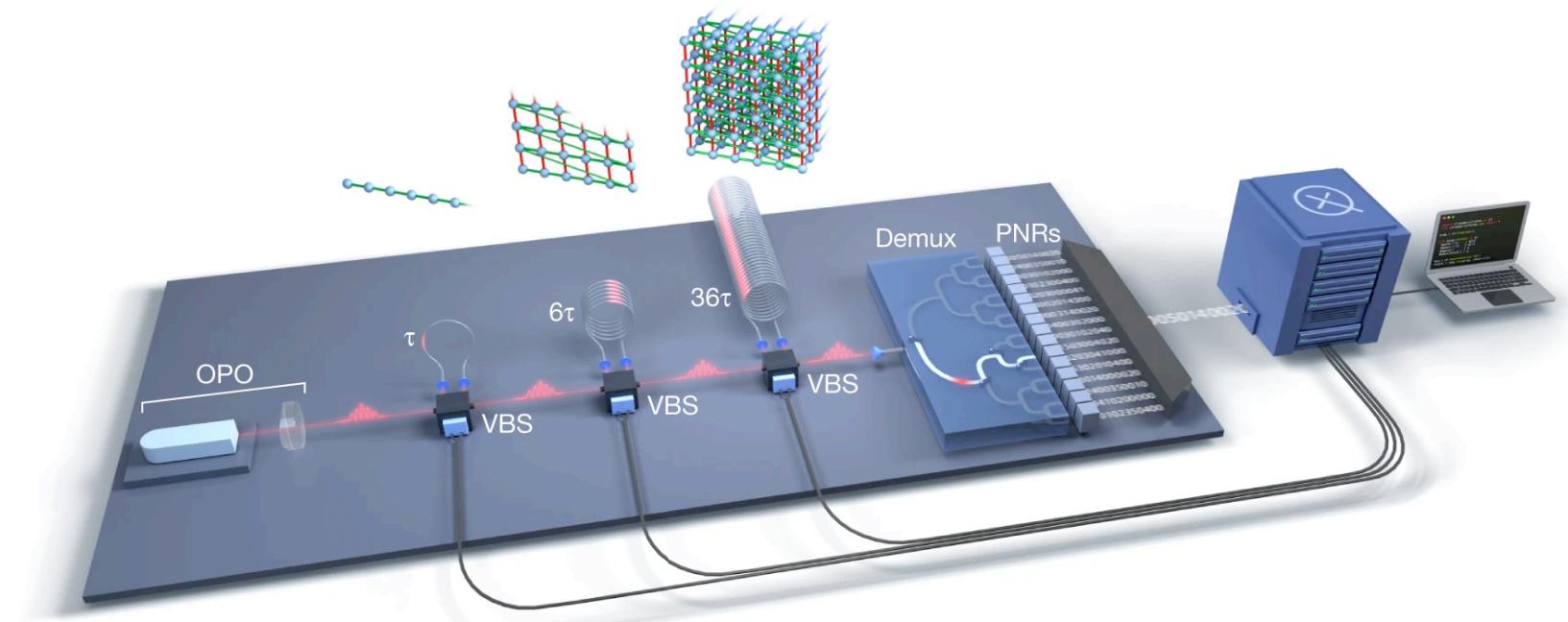
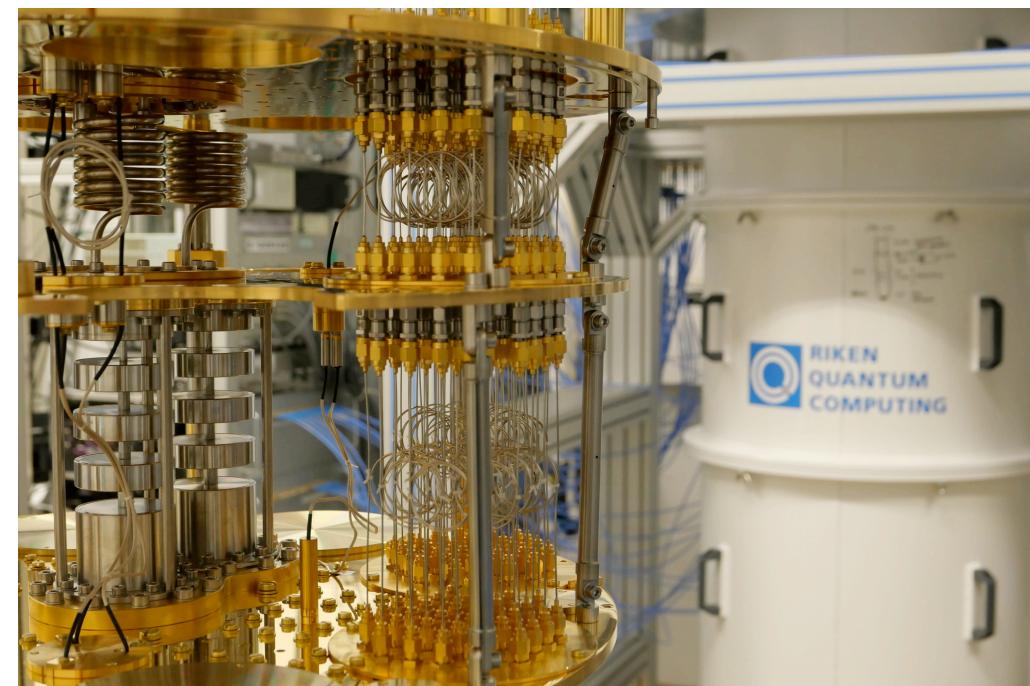
3 : 量子状態の変換 (古典確率分布の変換より大幅に速い場合がある)

- 量子フーリエ変換 $\sum_x \psi(x) |x\rangle \rightarrow \sum_v \mathcal{F}[\psi](v) |v\rangle$ 確率振幅のフーリエ変換
- 量子特異値変換 $|\psi\rangle \rightarrow \text{poly}(A)|\psi\rangle, |\psi\rangle \rightarrow A^{-1}|\psi\rangle$ [Gilyen et al., arXiv:1806.01838](#)

→測定 確率振幅の2乗の確率分布で答えが得られる乱択アルゴリズム

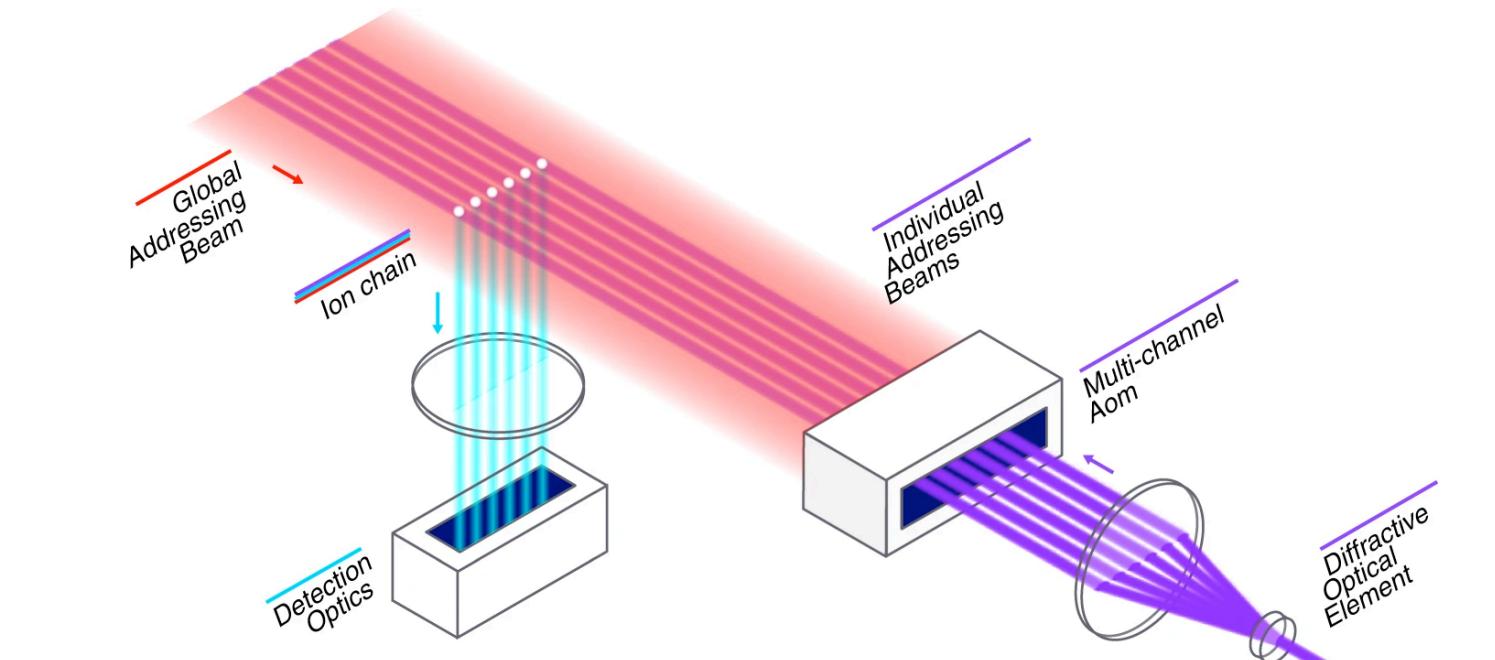
干涉を活用しほぼ正しい答えだけが効率的に得られるようにアルゴリズム設計する

量子コンピュータの現状



[© RIKEN Center for Quantum Computing]

[L. S. Madsen et al, Nature 606, 75 (2022)]



[K. Wright et al, Nat. Commun. 10, 5464 (2019)]

超電導量子ビット

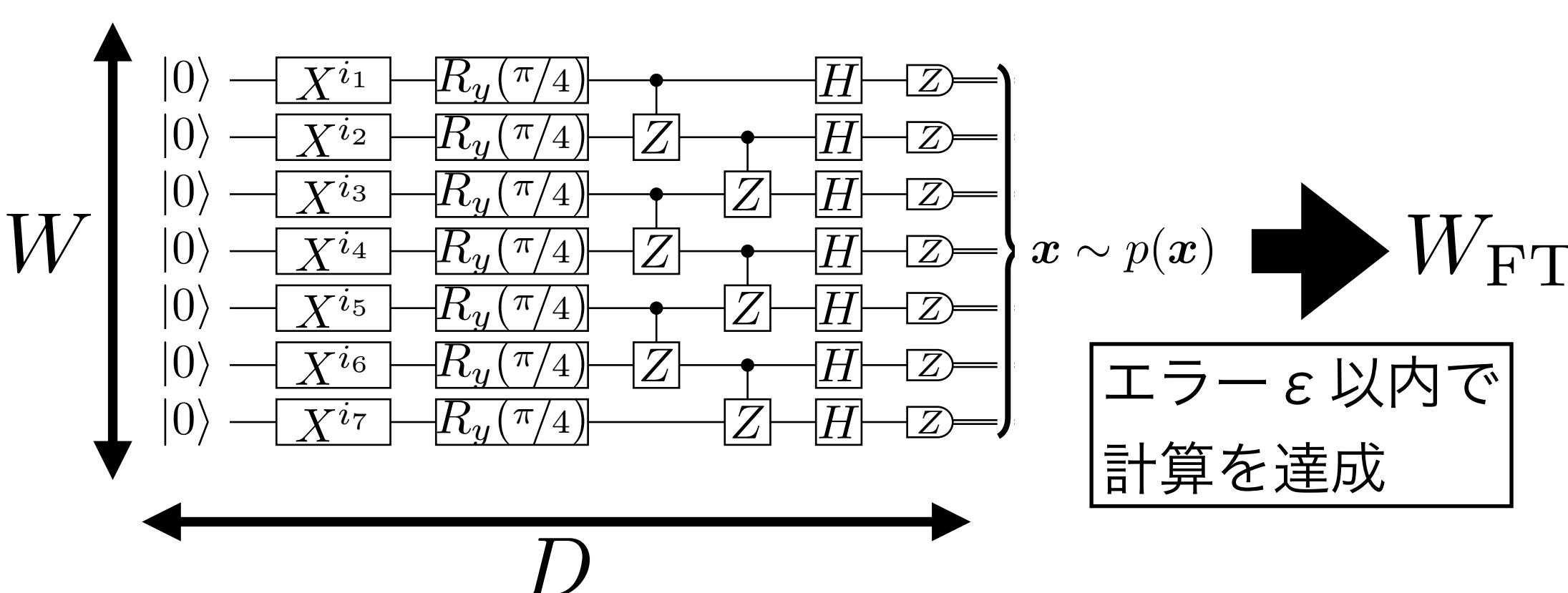
光学系

イオントラップ

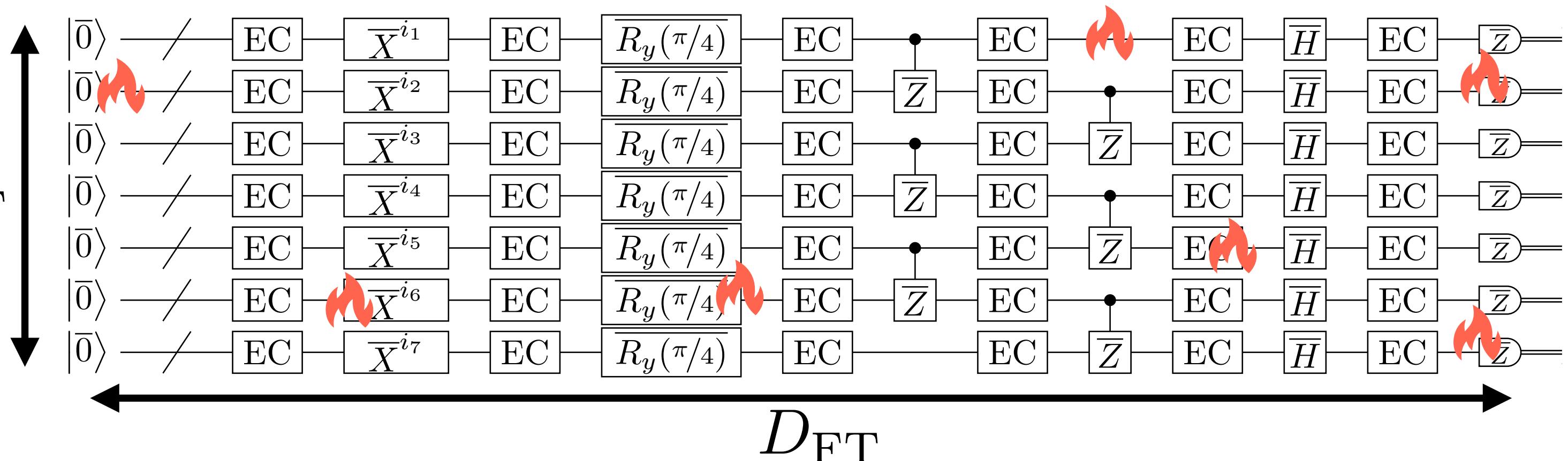
- ・ 様々な物理系で10-100量子ビットのデバイスができつつあり理論・実験とともに進展中
- ・ 課題：ノイズの影響→実用的問題を解くために大規模化すると正しい答えが出ない
(Noisy Intermediate-Scale Quantum (NISQ) devices, 量子アニーリングは原理的限界あり)
- ・ 解決方法：量子エラー訂正を用いた誤り耐性量子計算（長期プロジェクトで開発中）

誤り耐性量子計算の理論の進展

計算を記述する量子回路



ノイズのあるデバイス上で実行する量子回路



従来方式の問題点：オーバーヘッド大

$$\text{空間} : \frac{W_{\text{FT}}}{W} \approx \text{polylog}\left(\frac{WD}{\epsilon}\right)$$

$$\text{時間} : \frac{D_{\text{FT}}}{D} \approx \text{polylog}\left(\frac{WD}{\epsilon}\right)$$

(表面符号などを用いた方式)

大幅に効率化

ほぼ犠牲なし

我々の研究による理論的進展

$$\text{空間} : \frac{W_{\text{FT}}}{W} = O(1)$$

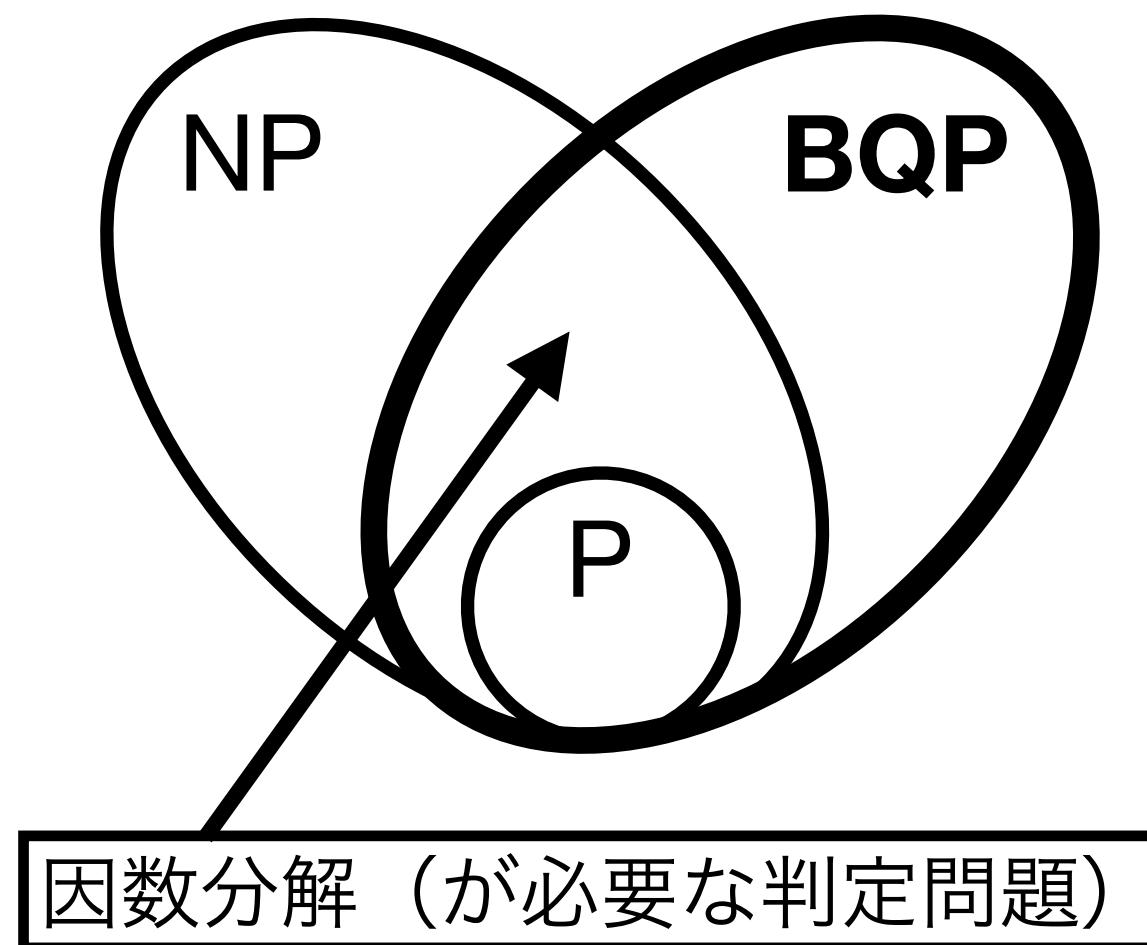
$$\text{時間} : \frac{D_{\text{FT}}}{D} \approx \text{quasi-polylog}\left(\frac{WD}{\epsilon}\right)$$

Yamasaki, Koashi arXiv:2207.08826

量子計算の実行方法の基盤構築 (4-11)

高速な量子機械学習のためのアルゴリズム開発 (12-22)

量子計算の計算能力



- P: 古典計算で多項式時間で解ける判定問題の集合
- NP: 古典計算で多項式時間で「答え合わせ」できる判定問題の集合
- BQP: 量子計算で多項式時間で解ける判定問題の集合
- 未解決問題: $P \neq NP?$ $P \neq BQP?$ 数学的な証明はまだ部分的

- 古典計算で量子計算を効率的にシミュレートするのは一般には難しい (=量子優位性)
- 量子計算は多項式時間の古典計算を常に多項式時間でシミュレート可能 (上位互換)
- NP-hardな組合せ最適化は量子計算でも多項式時間では難しい (=並列計算ではない)
- 量子力学に関わる問題でも量子計算が高速とは限らない (例: 基底エネルギーの計算)

重要な問題：量子計算が得意な範囲を活用し社会的に有用な計算問題を効率的に解けるか？

量子機械学習のアルゴリズム設計で重要なこと

量子機械学習アルゴリズムの代表例: **Harrow-Hassidim-Lloyd (HHL) アルゴリズム**

- $A \in \mathbb{C}^{N \times N}, |b\rangle \in \mathbb{C}^N$ が与えられたとき量子状態 $|x\rangle \propto A^{-1} |b\rangle$ を作る [Harrow, Hassidim, Lloyd, PRL (2009)]
- 逆行列計算は機械学習で多用される
- しかし HHL は線型方程式を解いているわけではない $Ax = b \xrightarrow{O(\text{poly}(N))} x = A^{-1}b$
- HHL は $|x\rangle$ を測定して行うサンプリングタスクに使える
s: sparsity, κ: condition number
- もし A が非常に疎なら古典計算より指数的に高速: runtime = $\tilde{O}(\log(N)s \times \kappa \text{polylog}(1/\epsilon))$
- しかし x の N 個の要素を推定しようとすると高速化が打ち消される [Aaronson, Nat. Phys. (2015)]

誤り耐性量子計算が実現されるとてもなお、量子計算の優位性を活かすには量子力学の特徴を踏まえたアルゴリズム設計が必要

量子機械学習の最近の動向

疎行列を使った高速化

- HHLアルゴリズム

$$|x\rangle \propto A^{-1} |b\rangle$$

Harrow, Hassidim, Lloyd, PRL (2009)

- A が非常に疎なとき指数的に高速 s: sparsity, κ: condition number
 $\tilde{O}(\log(N)s \times \kappa \text{polylog}(1/\epsilon))$

低ランク行列での高速化

- 同様のタスク $|x\rangle \propto A^{-1} |b\rangle$
- A が疎でなくても非常に

低ランクなら高速 r: rank

$$\tilde{O}(\log(N)r \times \kappa \text{polylog}(1/\epsilon))$$

Kerenidis, Prakash, ITCS2017

- しかしこれは多項式的高速化だと近年わかった：

Q-inspired classical algorithm

Tang, STOC2019

ヒューリスティクス

- NISQアルゴリズム,
量子アニーリング
- 高速化, 実行時間の保証なし
- 今の技術での優位性：難

量子的な問題設定

- 量子実験・特殊なデータセットでの高速化
Servedio et al., SIAM J. Comput. (2004)
Liu et al., Nat. Phys. (2021)
Heung et al. Science (2022)
- 理論として面白いが実データへの適用可能性は不明

自身の研究: 疎行列・低ランク行列の仮定なしで指数的高速化を活用する量子機械学習

= 上記の先行研究の枠を超えて、高速かつ適用範囲の広い量子機械学習の手法を構築

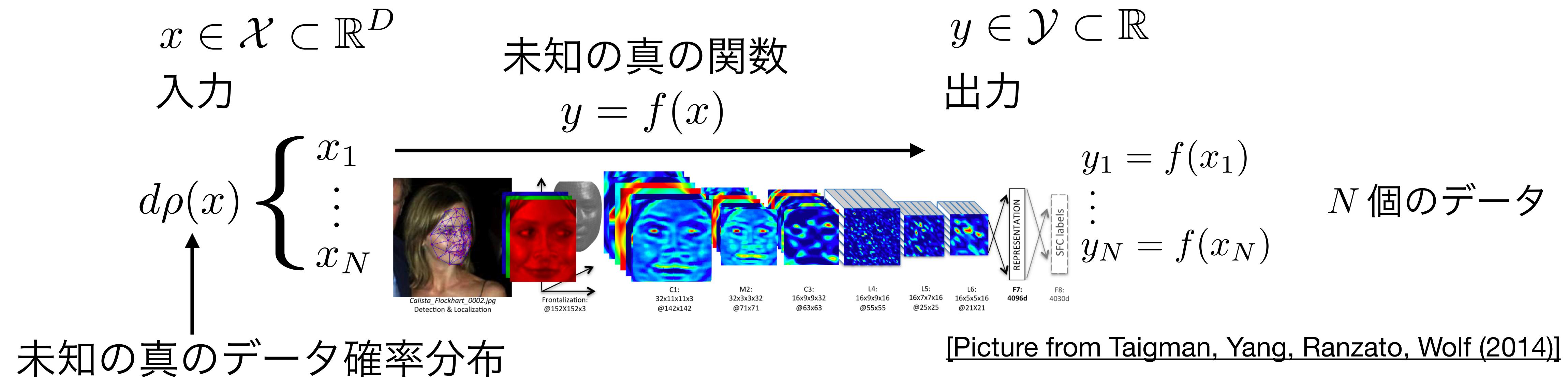
我々の研究の概要

最適なランダム特徴量をサンプリングするタスク (next slides) において、
知られている古典アルゴリズムより指数的に高速な量子アルゴリズムを開発

- ・ **目的**：学習に必要なパラメータ数を削減 = 学習や推論を高速化
- ・ **特色**：疎行列・低ランク行列の仮定なし = 幅広い学習タスクに適用可能
- ・ 高速化を損なうことなく学習タスク全体（回帰・分類）を達成

指数的な量子高速化を学習全体のend-to-endな高速化に応用する方法を確立

学習の状況設定：教師あり学習

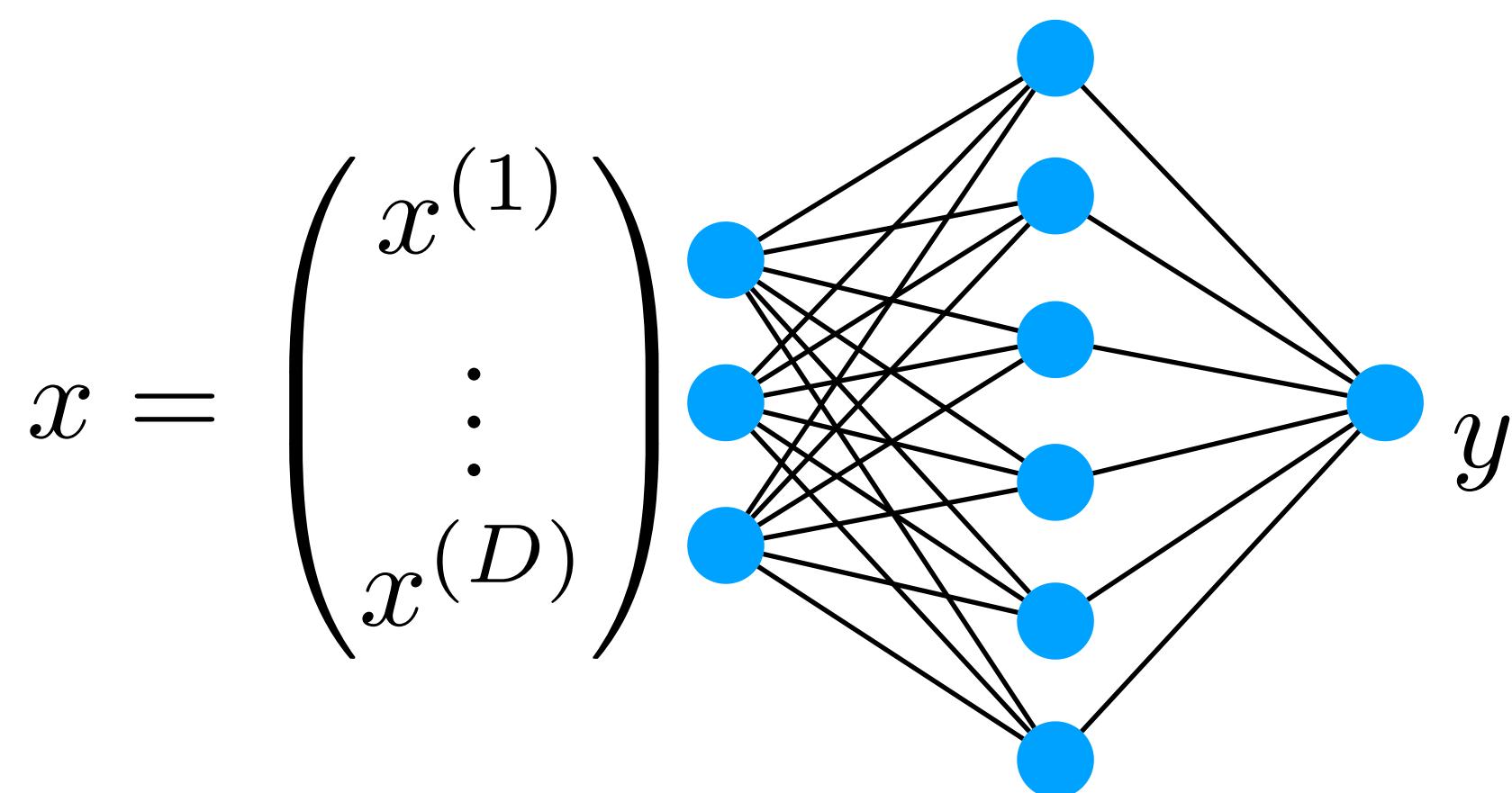


Find $\hat{f} \in \mathcal{F}$ a given space of functions = a model

error ϵ : average over true data distribution

minimizing $\int_{\mathcal{X}} d\rho(x) \left| \hat{f}(x) - f(x) \right|^2 \lesssim \epsilon$

手法：ランダム特徴量への量子機械学習の活用



ランダム特徴量

$$y = f(x) \approx_{\epsilon} \sum_{m=0}^{M-1} [\alpha_{2m} \cos(-2\pi i v_m \cdot x) + \alpha_{2m+1} \sin(-2\pi i v_m \cdot x)]$$

学習したい非線型関数

D: 入力データの次元

特徴量写像の線型結合
(e.g., sin & cos)

通常のランダム特徴量

サンプリング $v_0, \dots, v_{M-1} \sim d\tau$

data-independent

回帰 $\rightarrow \alpha_0, \dots, \alpha_{2M-1}$

特徴量の数

$$M = \tilde{O}(1/\epsilon^2)$$

[Rahimi & Recht, NIPS 2008.]

**特徴量の数
を削減**

最適化されたランダム特徴量

$v_0, \dots, v_{M-1} \sim q_{\epsilon}^*(v) d\tau(v)$

data-optimized

$\rightarrow \alpha_0, \dots, \alpha_{2M-1}$

$M = O(\text{polylog}(1/\epsilon))$ (e.g., Gaussian cases)

最適 (理論保証つき)

[F. Bach, JMLR 18, 1 (2017)]

実用上のボトルネック: $q_{\epsilon}^*(v) d\tau(v)$ からのサンプリングに指數時間 $O(\exp(D))$ かかる

研究結果: 最適化されたランダム特徴量を探す量子アルゴリズム

量子アルゴリズム

$O(D/\epsilon)$ time
per sampling from optimized distribution $q_\epsilon^*(v)d\tau(v)$

データ次元について
指数的高速化

既存の古典アルゴリズム

$O(\exp(D))$ time
for the same sampling

量子機械学習へのインパクト

疎行列・低ランク行列を仮定せずに高速化を達成 = 適用範囲の広い量子機械学習

$$q_\epsilon^*(v) \approx \langle \varphi(v, \cdot) | \hat{\mathbf{q}}^{(\rho)} (\hat{\Sigma} + \epsilon \mathbb{1})^{-1} | \varphi(v, \cdot) \rangle \quad \varphi(v, x) = \exp(-2\pi i v \cdot x)$$

[arXiv:2004.10756](#)

量子アルゴリズムの構築 + 回帰問題への応用 (NeurIPS2020)

[arXiv:2106.09028](#)

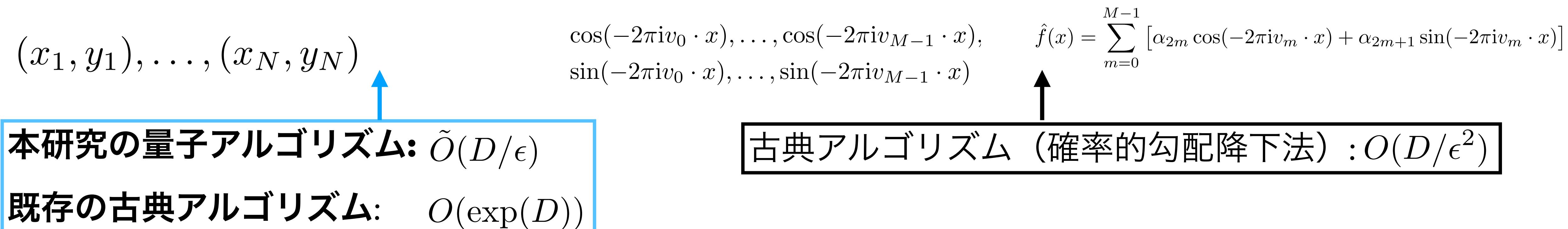
学習設定の明確化 + 分類問題への応用 (Under review)

[arXiv:2301.11936](#)

ニューラルネットワークへの拡張 (ICML2023)

応用：最適化された特徴量による学習の枠組み

データ（古典） → 最適化された特徴量（古典） → 学習したい関数（古典）



利点：学習に使う特徴量の数 M を大幅に削減

For Gaussian kernel, spherical/sub-Gaussian data distr. $M = O(\text{polylog}(1/\epsilon))$

[F. Bach, JMLR 18, 1 (2017)]

[Yamasaki, Sonoda, arXiv:2106.09028]

（通常のランダム特徴量では

$M = \tilde{O}(1/\epsilon^2)$)

[Rahimi & Recht, NIPS 2008.]

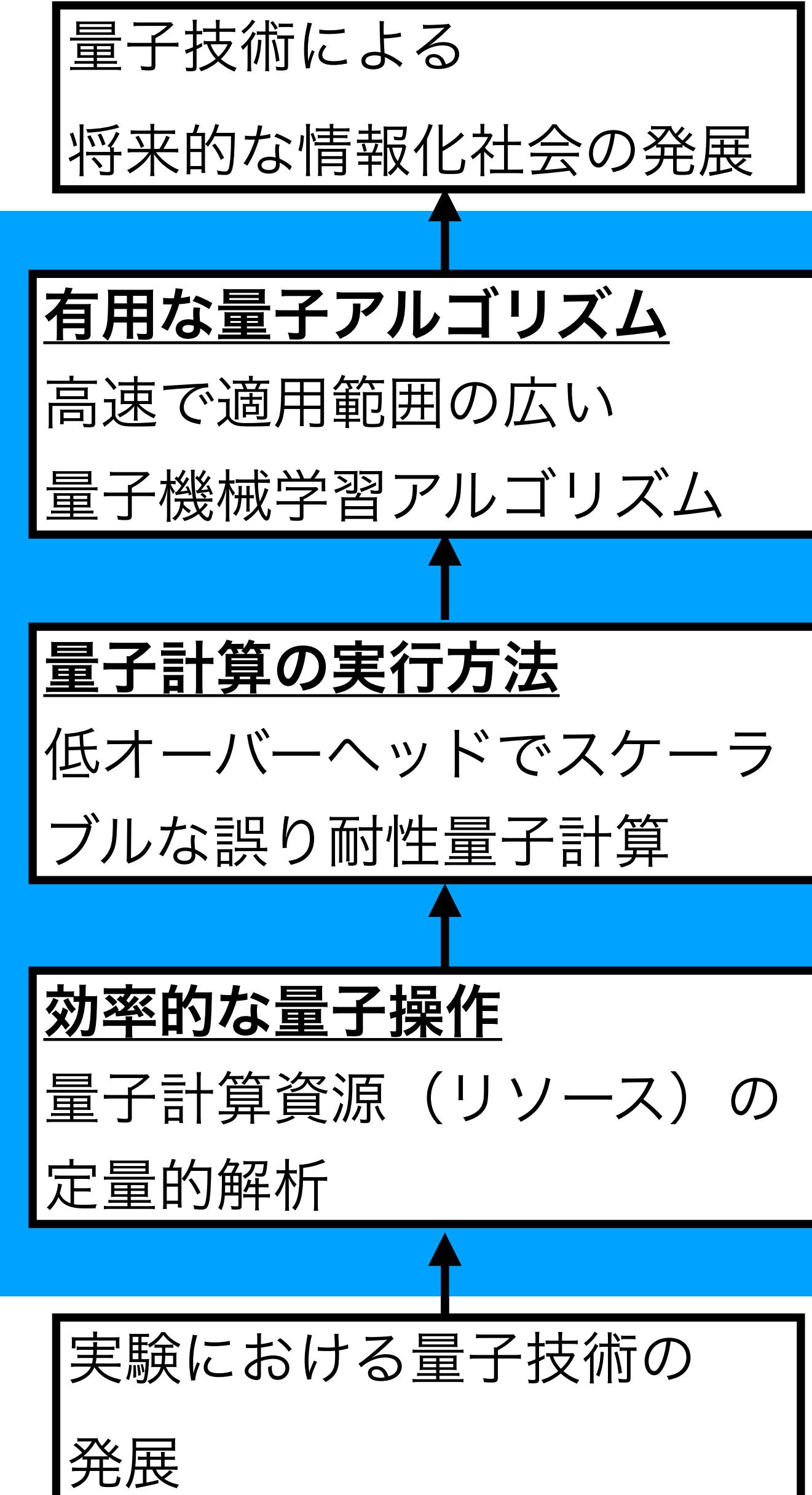
学習した関数を使う際のアドバンテージ：

$\hat{f} \rightarrow O(D \times \text{polylog}(1/\epsilon))$ runtime per use

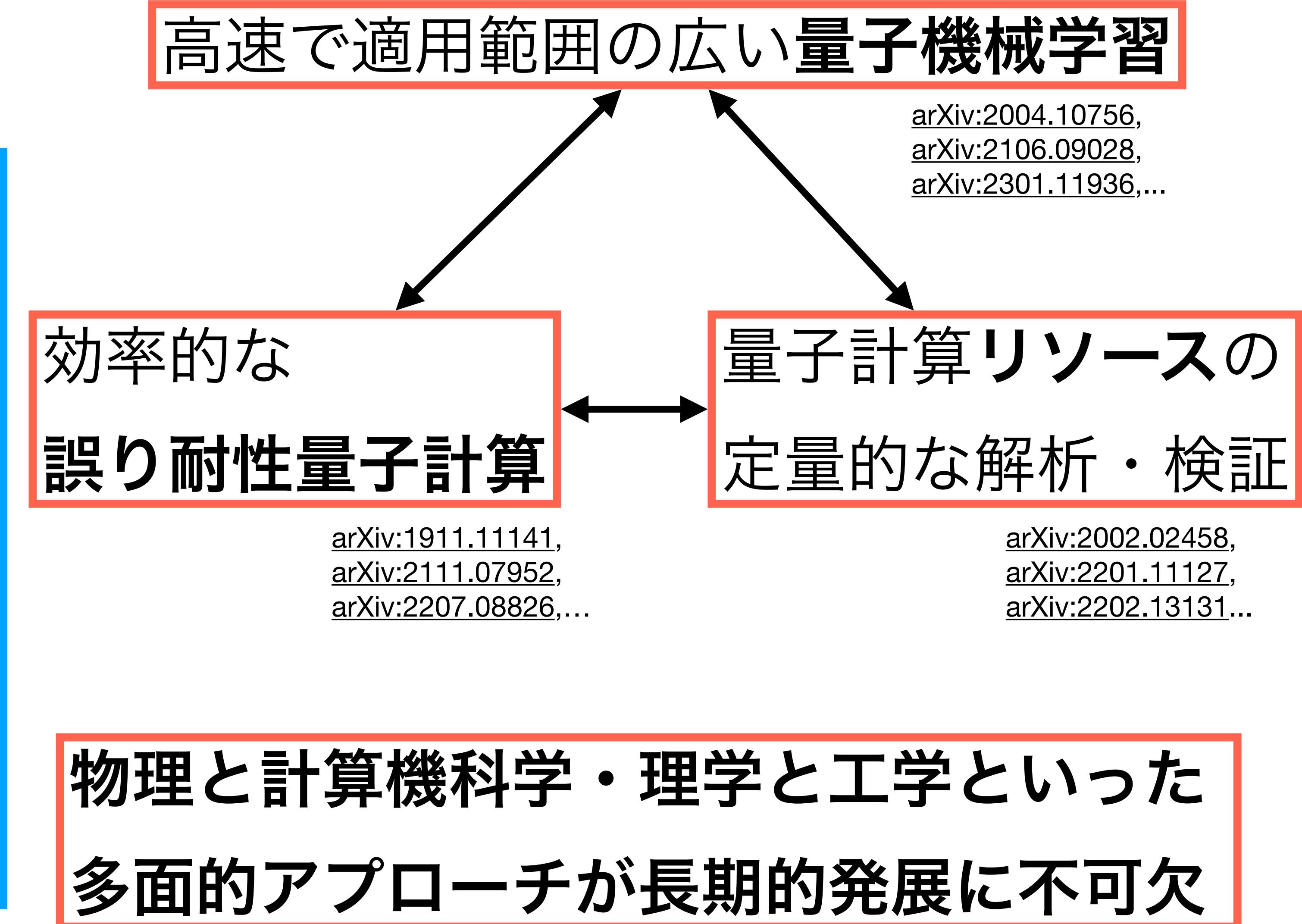
量子計算により古典計算では実行が難しい機械学習の効率化手法を開拓できる

研究の今後

社会実装



理論的基礎
=私の研究



実験の基礎

Take home message

- 量子アルゴリズムは負の値を取りうる確率振幅の干渉を活用できる乱択アルゴリズムであり、古典アルゴリズム開発では不可能な高速化を可能にするポテンシャルがある
- 機械学習などの社会的にインパクトのある問題で高速な量子アルゴリズムの応用を探すには、**量子計算が得意な守備範囲を活かすのが重要**
- 重要な特徴量のサンプリングに量子計算を活用することで、学習全体で高速化を損なく量子優位性を量子機械学習に有効活用できる

[arXiv:2004.10756](https://arxiv.org/abs/2004.10756) (NeurIPS 2020) H. Yamasaki, S. Subramanian, S. Sonoda, M. Koashi

[arXiv:2106.09028](https://arxiv.org/abs/2106.09028) (Under review) H. Yamasaki, S. Sonoda

[arXiv:2301.11936](https://arxiv.org/abs/2301.11936) (ICML2023) H. Yamasaki, S. Subramanian, S. Hayakawa, S. Sonoda