

3.1 データの準備と最適化

まず、比較基準となる「真値」を得るために、実データに基づく古典的な解析を行った。日本国債（1年・8年・30年）の3銘柄について、過去7年間の年間リターンデータを使用した。過去7年間の各資産の年間リターンを行列 $R \in \mathbb{R}^{7 \times 3}$ として与え、資産ごとの平均リターンベクトル μ を

$$\mu = \frac{1}{7} \sum_{t=1}^7 \mathbf{R}_t$$

により算出したものを表1に示す

表1 7年間の各資産の年間リターン

Year	1年国債	8年国債	30年国債
1	0.027	-0.046	-0.377
2	0.001	0.034	0.207
3	0.010	-0.006	0.032
4	0.049	0.333	0.787
5	0.051	0.340	0.243
6	0.470	0.246	0.527
7	0.367	0.918	1.102

次に、表1から3つの債券の共分散行列 Σ を計算した。

$$\Sigma = \begin{bmatrix} 0.037602905 & 0.038592381 & 0.057048119 \\ 0.038592381 & 0.110059476 & 0.142509524 \\ 0.057048119 & 0.142509524 & 0.24143881 \end{bmatrix}$$

その逆行列 Σ^{-1} を計算した。リスク回避度パラメータを λ とすると、平均分散最適化に基づくポートフォリオの理論的最適ウェイト \mathbf{w} は以下の式で与えられる

$$\mathbf{w} = \frac{\lambda}{2} \Sigma^{-1} \mu.$$

この時、 $\lambda = 1$, \mathbf{w} はポートフォリオの比重であり、レバレッジありとして、計算している。得られた最適ウェイトに対して、理論的な期待リターン E_{theory} および分散 V_{theory} は次式により計算される

$$E_{\text{theory}} = \mathbf{w}^\top \mu, \quad V_{\text{theory}} = \mathbf{w}^\top \Sigma \mathbf{w}.$$

モンテカルロ法によるポートフォリオリターンの推定は、平均 μ 、共分散 Σ の多変量正規分布に従う乱数 $N = 10^6$ サンプルを生成し、各サンプルに対してポートフォリオウェイト \mathbf{w} を適用することでポートフォリオリターンを計算した

$$R_{\text{portfolio}}^{(i)} = \mathbf{w}^\top \mathbf{r}^{(i)}, \quad i = 1, 2, \dots, N.$$

これらのサンプルから、モンテカルロ推定によるポートフォリオ期待値 E_{mc} および分散 V_{mc} をそれぞれ

$$E_{\text{mc}} = \frac{1}{N} \sum_{i=1}^N R_{\text{portfolio}}^{(i)}, \quad V_{\text{mc}} = \frac{1}{N} \sum_{i=1}^N (R_{\text{portfolio}}^{(i)} - E_{\text{mc}})^2$$

で計算している。ポートフォリオの VaR, 95% 信頼区間は、ポートフォリオ損失の分布に基づき、以下の式で定義される。

$$VaR_{95\%} = \mu - 1.645 \sigma \quad (1)$$

ここで、 μ はポートフォリオの期待リターン（平均）、 σ はポートフォリオ損失の標準偏差である。係数 1.645 は標準正規分布における 95% 信頼区間の下側パーセンタイルに対応している。

算出結果は以下の通りとなった。

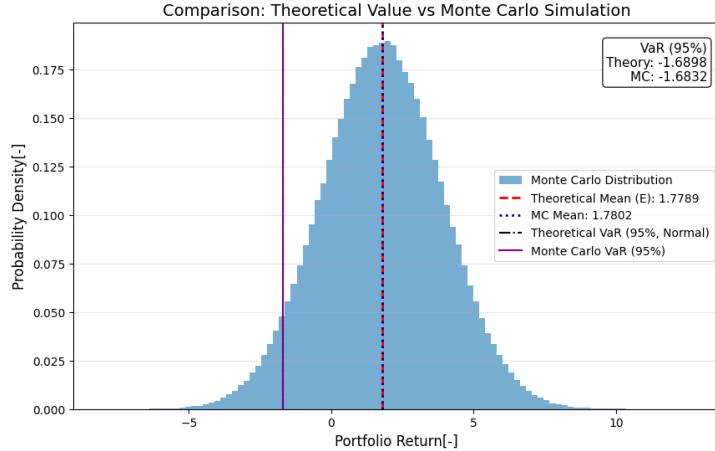


図 1 古典計算に基づく国債のポートフォリオのモンテカルロシミュレーション結果。95% 信頼水準における VaR は、理論値とモンテカルロ推定値の双方において少数第 2 位まで一致しており、本研究で用いたポートフォリオ比重に対して、理論計算とモンテカルロ法の結果が良好に一致していることが確認できる。

本計算で得られた結果を基準（真値）とし、これを用いて PCA による次元削減手法および量子計算手法の評価を行う。

3.2 主成分分析（PCA）による次元圧縮

本研究で扱う複数銘柄の国債リターンは、銘柄間で強い相関を有している。そのため、量子計算における量子ビット数削減を目的として、主成分分析（Principal Component Analysis: PCA）を用いた次元圧縮を行った。

まず、7 年分のリターンデータを

$$\mathbf{R} = \begin{pmatrix} \mathbf{r}^{(1)} \\ \mathbf{r}^{(2)} \\ \vdots \\ \mathbf{r}^{(7)} \end{pmatrix} \in \mathbb{R}^{7 \times M}$$

と表す。ここで M は資産数であり、本研究では $M = 3$ である。各資産の平均リターンベクトル μ を

$$\mu = \frac{1}{T} \sum_{t=1}^T \mathbf{r}^{(t)}$$

として定義し、平均を差し引いたデータ $(\mathbf{R} - \mu)$ を用いて解析を行う。

次に、共分散行列 Σ に対して固有値分解を行う。

$$\Sigma \mathbf{v}_k = \lambda_k \mathbf{v}_k,$$

ただし λ_k は第 k 主成分に対応する固有値、 \mathbf{v}_k は対応する固有ベクトルである。本研究では、最大固有値に対応する固有ベクトル \mathbf{v}_1 を第 1 主成分ベクトルとして採用し、ノルムが 1 となるよう正規化を行った。

このとき、各時点のリターンデータ $\mathbf{r}^{(t)}$ は、第 1 主成分方向への射影として次式で表される主成分スコア $Z^{(t)}$ に変換される。

$$Z^{(t)} = \mathbf{v}_1^\top (\mathbf{r}^{(t)} - \boldsymbol{\mu}), \quad t = 1, 2, \dots, T.$$

第 1 主成分は全分散の大部分を説明しており、金利水準全体の上昇・下降といった共通因子を表すモードとして解釈できる。

主成分分析において、各主成分が全体の分散をどの程度説明しているかは、固有値を用いた寄与率によって評価される。第 k 主成分の寄与率 C_k は次式で定義される。

$$C_k = \frac{\lambda_k}{\sum_{j=1}^M \lambda_j},$$

ここで λ_k は共分散行列 Σ の第 k 固有値であり、 M は資産数を表す。

本研究で用いた国債リターンデータに対して主成分分析を行った結果、第 1 主成分が全分散の大部分を占めることが確認された。このことから、国債リターンの変動は主として単一の共通因子によって説明可能であると考えられる。

対応する第 1 主成分ベクトル \mathbf{v}_1 は以下のように求められた。

$$\mathbf{v}_1 = \begin{pmatrix} -0.2166 \\ -0.5273 \\ -0.8216 \end{pmatrix}.$$

各成分はいずれも同符号であり、第 1 主成分が複数銘柄の国債リターンの同方向の変動、すなわち金利水準全体の上昇・下降を表すモードとして解釈できる。

この第 1 主成分方向に平均を差し引いたリターンデータを射影することで、式 (18) に基づく主成分スコア

$$Z^{(t)} = \mathbf{v}_1^\top (\mathbf{r}^{(t)} - \boldsymbol{\mu})$$

を計算した。得られた 3 つの銘柄の年次データの主成分スコアは次の通りである。

$$Z = (0.7912, 0.2749, 0.4378, -0.3697, 0.0731, -0.2014, -1.0059).$$

これは、 7×3 の行列から第 1 主成分の 1 次元に次元削減できたことを示している。このとき、3 次元から 1 次元に削減した場合の寄与率は

$$\text{寄与率} = \frac{\text{Var}(Z)}{\sum_{j=1}^3 \text{Var}(R_j)} \approx 0.866 \tag{2}$$

3.3 量子サンプリングの実装

まず、量子サンプリングとの比較のために、古典計算から累積度数をもとめた。PCA から

$$Z = (0.7912, 0.2749, 0.4378, -0.3697, 0.0731, -0.2014, -1.0059).$$

が既知である。

PCA により簡略化したモデルを、量子ビットの重ね合わせとして表現した。PCA により得られた第 1 主成分スコア $Z = [z_1, z_2, \dots, z_7]$ は、高次元ポートフォリオのリターンの主要な変動要因を表す低次元表現である。ポートフォリオリスクの評価を目的として、損失部分の値を調べたいため、損失変数 L を第 1 主成分スコアの符号を反転することによって定義する

$$L = -Z,$$

ここで、正の値がポートフォリオの損失の大きさを示す。損失変数 L に対して正規分布による近似を行う。正規分布のパラメータとして、損失の平均 μ_L と標準偏差 σ_L をそれぞれ次式により算出する：

$$\mu_L = \frac{1}{n} \sum_{i=1}^n L_i, \quad \sigma_L = \sqrt{\frac{1}{n} \sum_{i=1}^n (L_i - \mu_L)^2},$$

上記の式から正規分布に従った古典計算による真値は、

PCA により 1 次元化された確率分布に対し、量子計算を用いた確率推定の基礎的検証として、1 量子ビット回路による累積分布関数（CDF）の推定を行った。

まず、古典計算によって得られた真の累積確率を p_{true} とする。この確率値を量子状態の振幅として符号化するため、以下の回転角 θ を定義する。

$$\theta = 2 \arcsin (\sqrt{p_{\text{true}}}).$$

このとき、初期状態 $|0\rangle$ に対して y 軸回転ゲート $R_y(\theta)$ を適用すると、量子状態は

$$R_y(\theta)|0\rangle = \sqrt{1 - p_{\text{true}}} |0\rangle + \sqrt{p_{\text{true}}} |1\rangle$$

となる。

この状態に対して測定を行うことで、 $|1\rangle$ が観測される確率は理論的に p_{true} に一致する。したがって、多数回の測定結果から $|1\rangle$ の出現頻度を求ることで、累積分布関数の量子推定値を得ることができる。

本研究では、測定回数を $\text{shots} = 10,000$ とし、量子回路をシミュレータ上で実行した。測定結果における $|1\rangle$ の出現割合を

$$p_{\text{quantum}} = \frac{N_1}{\text{shots}}$$

として定義し、これを量子計算による CDF 推定値とした。ここで N_1 は $|1\rangle$ が観測された回数である。

得られた量子推定値と古典計算による真値との差を、以下の絶対誤差によって評価した。

$$\varepsilon = |p_{\text{true}} - p_{\text{quantum}}|.$$

この誤差は測定回数に依存し、理論的には二項分布に基づく統計誤差 $\mathcal{O}(1/\sqrt{\text{shots}})$ に従う。

以上より、本手法は 1 量子ビットという最小構成の量子回路を用いて、古典的に与えられた確率値を振幅として正しく符号化し、測定を通じてその値を再現できることを確認した。まず、今回用いた量子回路の図をいかに示す。

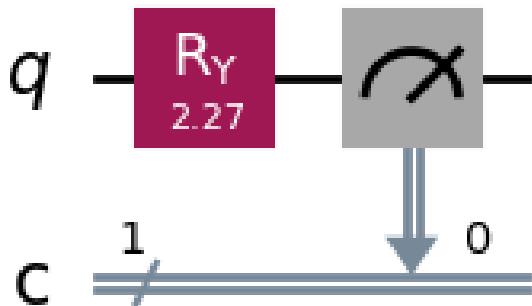


図 2 実験で用いた量子回路

以下は古典計算と量子計算、量子実機 (ibm torino) の比較である。

表 2 量子実機 (ibm torino) による 95% VaR の測定結果 (5 回の試行)

試行回数	測定値 (QPU)
1 回目	-1.0282
2 回目	-1.0914
3 回目	-1.1187
4 回目	-1.0282
5 回目	-1.0509
平均 ± 不確かさ	-1.0635 ± 0.0180

表 3 95% VaR 推定の評価ステップ別比較

理論値 (図 1)	PCA 理論値 (古典)	量子実機 (ibm torino) 推定値 (QPU)
-1.6898	-1.2100	-1.0635 ± 0.0180

単一ポートフォリオにおける 95% VaR の評価ステップ別推定結果を示す。フル理論値は元データに基づく正規分布理論値であり、誤差要因は存在しない。PCA 理論値は次元圧縮を行った損失変数に基づく古典計算値であり、第 2・3 主成分の情報欠損による誤差が生じる。量子実機推定値は IBM QPU 上の 1 qubit 回路を用いたサンプリングにより算出され、量子ノイズおよびショット誤差が主な誤差要因である。各値は負の損失方向を示しており、絶対値の大小がリスクの大きさに対応する。

算出した誤差の大きさは以下の通りとなった。

$$\epsilon_{PCA} = |-1.6898 - (-1.2100)| = 0.4798 \quad (3)$$

$$\epsilon_{QPU} = |-1.2100 - (-1.0635)| = 0.1465 \quad (4)$$

この結果から、全誤差の大部分は PCA による次元圧縮に起因していることがわかる。これは、PCA 量子実機による推定値は、PCA 後の理論値に対して概ね良好な一致を示しており、実機上での確率分布の符号化およびサンプリングアルゴリズムが妥当に動作していることが確認された。

図??におけるポートフォリオの真値は、期待リターン $E_{theory} = 1.7789$ 、標準偏差 $\sigma_{total} = 2.1089$ であるため、95% 信頼水準におけるリスク距離は $1.645 \times 2.1089 \approx 3.47$ となる。したがって、真値の VaR は次式のように求められる。

$$VaR_{full} = 1.7789 - 3.4691 \approx -1.69 \quad (5)$$

次に、主成分分析 (PCA) による次元圧縮後の理論上の挙動を考察する。本ポートフォリオに対する第 1 主成分の分散寄与率は約 86.6% であった。VaR の算出には分散ではなく標準偏差を用いるため、第 1 主成分が説明するリスクの割合は $\sqrt{0.866} \approx 0.93$ 倍に縮小する。このスケーリングに基づくと、PCA を用いた場合の理想的な VaR は以下の通り予測される。

$$VaR_{PCA_ideal} = 1.7789 - (3.4691 \times 0.93) \approx -1.45 \quad (6)$$

しかしながら、表??に示した PCA 理論値は -1.2100 であり、予測される -1.45 よりもさらにリスクが小さく見積もられている。この乖離の原因は、本研究で使用した標本期間における第 1 主成分の標本分散が、理論的な共分散行列 Σ から導出される固有値 λ_1 よりも小さかったことが挙げられる。つまり、対象とした 7 年間において市場の主要な変動成分（第 1 主成分）の動きが理論予測よりも抑制されていたため、結果として VaR の絶対値がさらに減少したと考えられる。

また、以下のように PCA 後の理論値と QPU でそれが生じた。

$$\epsilon_{QPU} = | -1.2100 - (-1.0635) | = 0.1465 \quad (7)$$

これは、QPU は、離散的で今回の場合、8 個の点をプロットしてその閾値によって累積度数を決定したためである。

4 上場企業の関連性の低い銘柄のポートフォリオ分析の手法

本研究では、分散投資による、PCA の寄与率の分析、QAE の実行を目的として、互いの価格変動の関連性が低い銘柄群の中から、NISA 枠において取引量や保有者数の観点で人気の高い銘柄を選定した。また、配当金せず株価の急騰はパフォーマンスにおけるポートフォリオ分析を目的としている。

表 4 異業種 4 銘柄の年間騰落率データセット (2016–2025)

年	SBG (9984)	任天堂 (7974)	MUFG (8306)	三菱重工 (7011)
2016	35.0	46.5	-4.7	-20.0
2017	16.9	67.9	14.4	14.0
2018	-16.0	-28.9	-34.7	-5.0
2019	33.0	50.2	10.0	1.0
2020	85.6	49.7	-23.1	-25.0
2021	-38.7	-18.5	37.1	15.0
2022	-9.9	3.1	42.2	60.0
2023	4.3	33.0	36.3	55.0
2024	38.2	25.9	31.8	80.0
2025	20.5	14.4	55.2	105.0

また、相関行列は

$$R = \begin{bmatrix} 1.00 & 0.65 & -0.12 & -0.05 \\ 0.65 & 1.00 & -0.06 & -0.15 \\ -0.12 & -0.06 & 1.00 & 0.84 \\ -0.05 & -0.15 & 0.84 & 1.00 \end{bmatrix}$$

この時、1 行 1 列目は SBG で任天堂、MUFG、三菱重工の順番で並んでいる。国債の相関行列と比較すると三菱グループには相関があるがその他は比較的相関が低いといえる。また、この時のモンテカルロシミュレーションと理論値は以下の通りとなった。

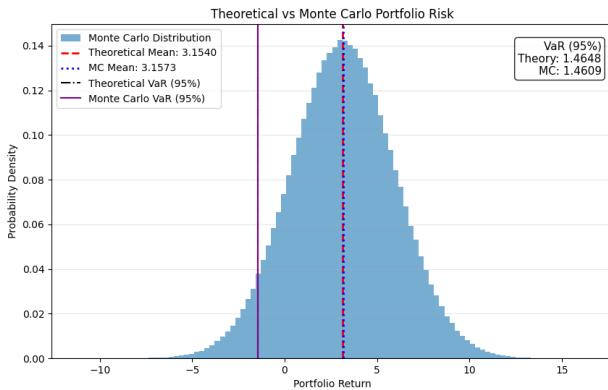


図3 古典計算に基づく選定した4銘柄ポートフォリオのモンテカルロシミュレーション結果。95%信頼水準におけるVaRは、理論値とモンテカルロ推定値の双方において少数第2位まで一致しており、本研究で用いたポートフォリオ比重に対して、理論計算とモンテカルロ法の結果が良好に一致していることが確認できる。

寄与率は以下の通りとなった。

表5 主成分分析における寄与率および累積寄与率

主成分	寄与率	累積寄与率
第1主成分 (PC1)	約 48.5%	48.5%
第2主成分 (PC2)	約 43.2%	91.7%
第3主成分 (PC3)	約 5.4%	97.1%
第4主成分 (PC4)	約 2.9%	100.0%

また累積寄与率は

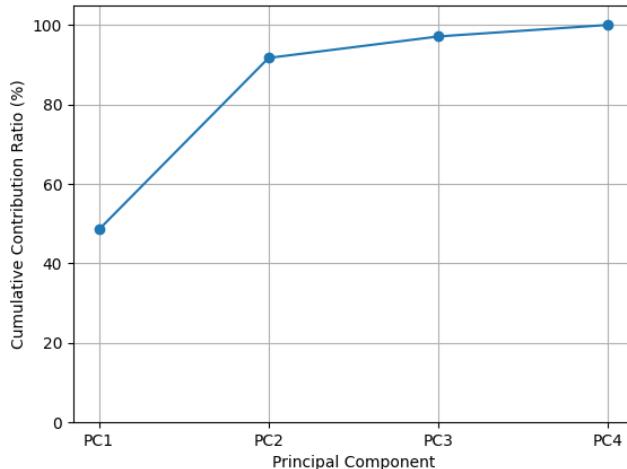


図4 各銘柄のPCA分析における累積寄与率

このように国債の場合は、寄与率が第1主成分だけで86.6%であったが、今回は第一主成分のみだと48.5%まで減少した。しかし、第一主成分、第二主成分の2つで分析することで91.7%の寄与率を達成できるので今回は2量子ビットで分析する。今回の第一主成分、第二主成分は以下の通りとなった。

表 6 主成分分析によって得られた年次 PC スコア

年	PC1 スコア (Z_1)	PC2 スコア (Z_2)
2016	58.86	-1.24
2017	23.61	-20.55
2018	22.08	78.51
2019	36.44	-13.94
2020	90.83	-29.66
2021	-32.41	62.51
2022	-51.97	11.61
2023	-31.87	-12.47
2024	-38.00	-38.96
2025	-77.56	-35.82

表 7 主成分スコアの標準偏差

主成分	PC1 (Z_1)	PC2 (Z_2)
標準偏差 σ	51.30	38.29

PCA によって次元を削減した場合

$$V_{\text{theory}} = w^\top \Sigma w = \sum_{k=1}^4 \lambda_k (w^\top v_k)^2 \quad (8)$$

である。生データから理論上のポートフォリオでは $\lambda_k = 5.0$ として計算している。計算結果は以下のとおりである。

表 8 主成分近似によるリスク評価指標の比較

評価モデル	説明できるリスク	期待値 E	標準偏差 σ	95% VaR (損失)	真値との誤差
真値 (生データ)	100%	3.1540	1.0269	1.4648	-
PC1 近似 (1bit)	64.3%	3.1540	0.6603	2.0678	0.6030
PC1+PC2 近似 (2bit)	94.7%	3.1540	0.9725	1.5543	0.0895

PCA で次元を削減したデータを用いて量子計算を行った。量子回路は以下のとおりである。

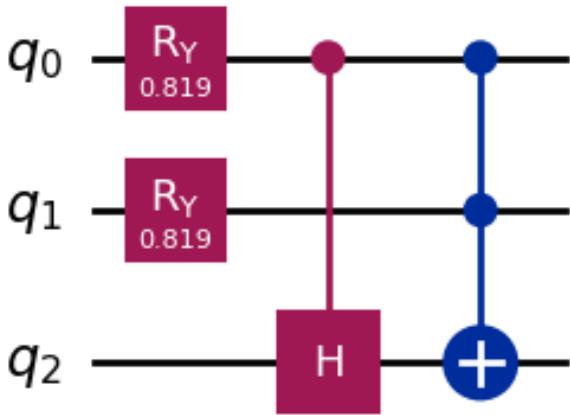


図 5 2 次元データ量子計算分析のための量子回路

この回路を用いて 2 文探索法によって損失閾値を更新しながら VaR,95% になるように測定した。2 量子ビットの量子計算では理論値 $\lambda = 5.0$ を考慮しスケーリングを行った。量子回路シミュレーションによる VaR の結果は以下のとおりである。

表 9 量子計算・生データによる VaR 比較

量子計算	生データ（真値）
1.7673	1.4648

このように、生データと比較して推定値が真値を下回った。これは、量子レジスタは、図 6 のように 2 量子ビット 4 状態による離散化で行ったためである。これによって 5% の確率の間の表現ができなかったからだといえる。したがって、量子ビット数を多くすることにより正確なデータなる。以下に量子ビット数と VaR の推移を示す。

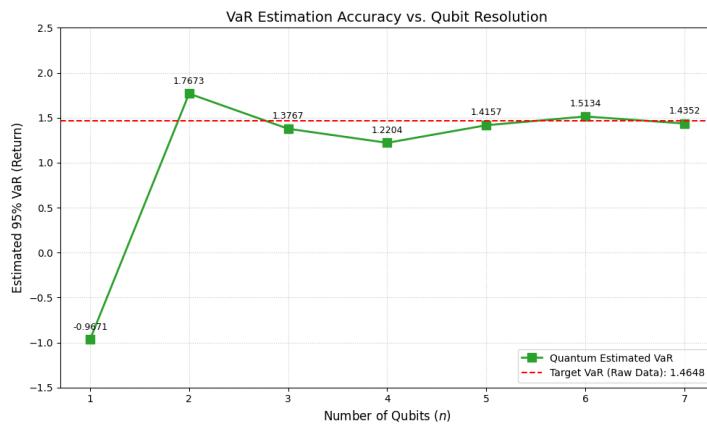


図 6 量子ビット数と VaR の推移

このように、量子ビット数を多くするほど理論値へと収束していることがわかる。またこのグラフから今回の場合 5 量子ビットで VaR の値が安定化していることがわかる。

シミュレーションで得られた知見に基づき、IBM Quantum の実機 (ibm torino) を用いた検証を実施し

た。PCA 済の理論上の期待リターンは 3.154, VaR は 1.4648 なので

$$x = 3.154 - 1.4648 = 1.6892$$

となり、この値は損失側へ x 軸が 1.6892 推移した場合 VaR が求まることを表している。この値を x として ibm torino でこの値を用いて 5 量子ビットで真値と比較し精度を求める。今回の場合シミュレーションで 1.6892 の場合 VaR=95% つまり確率 $\alpha = 0.05$ となるはずである。従って、既知の 1.6892 を入力し逆算して QPU に $\alpha = 0.05$ を出力するようにプログラムを行った。結果は以下のとおりである。

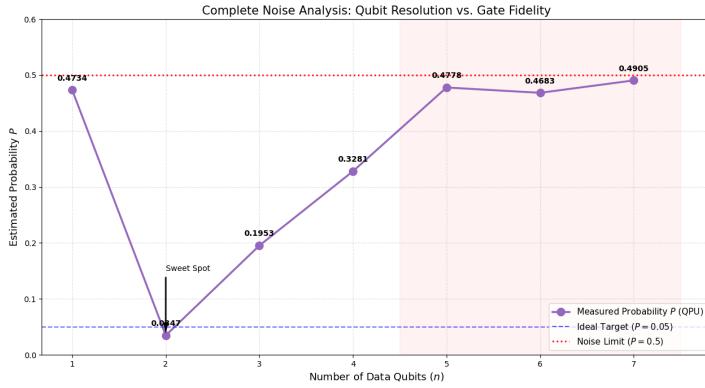


図 7 ibm torino の量子ビット数を上げることで $\alpha = 0.5$ へ収束している様子。量子ビット数の増加に伴い、デコヒーレンスおよびゲート誤差が累積することで、量子状態は最大混合状態に近づく。その結果、測定確率は理論的な値から乖離し、基底状態 $|0\rangle$ および $|1\rangle$ がほぼ等確率で観測される状態、すなわち確率 $1/2$ への収束が生じた。

また、量子ビットの行き先の詳細は以下のとおりである。

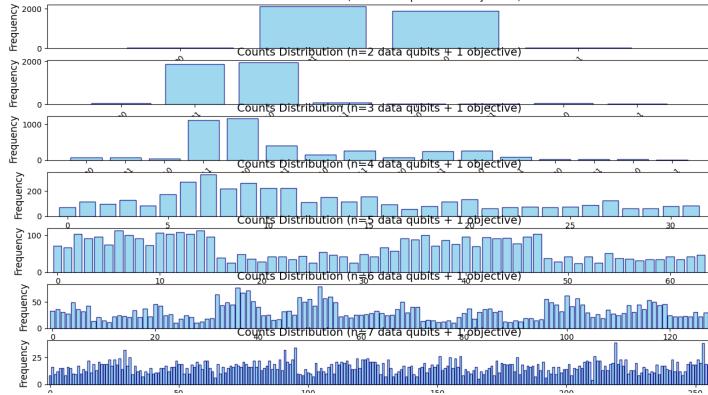


図 8 量子ビット別のヒストグラムによる量子情報の詳細

ショット数と量子ビット数の関係をいかに示す。

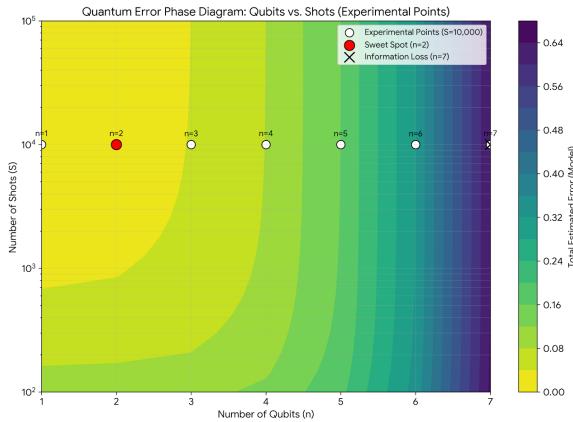


図9 ショット数と量子ビット数の関係。ショット数を上げることは観測回数を上げることなので誤差は小さくなる。また、量子ビット数を上げるとデコヒーレンス化によって量子コンピュータの性質上 2^n の勢いでノイズが増加する。

$n=1$ の場合、用いる量子ビット数が少なく情報量が不足するため、測定結果が過度に平均化され、目的とする有意水準から乖離した値が得られた。一方で、 $n=5$ 以上では量子ビット数の増加に伴い情報量は増加するものの、重ね合わせ状態に付随するノイズが急激に増大し、測定結果がノイズに支配されることで、 $\alpha = 0.5$ 付近に収束してしまった。

これに対し、 $n=2$ の場合は、情報量が過不足なく確保されており、ノイズの影響も比較的抑制されるため、 $\alpha = 0.05$ 付近の値を達成できたと考えられる。この結果は、量子ビット数が少なすぎても多すぎても目的の値を正確に推定できないことを示しており、適切な量子ビット数の選択が重要であることを示唆している。また、具体的なノイズの影響は、図9からわかる通り、指數関数的にノイズが大きくなってしまうので量子ビットを増やすことは慎重にならないといけないことがわかる。

この現象は、光学系において倍率を過度に上げた際に収差や回折限界によって像がぼやける現象と類似している。すなわち、量子計算においても、ノイズがアルゴリズムの解像度を上回ると、本来観測すべき結果が失われてしまうと解釈できる。

また、シミュレーション環境では理想的な条件下で計算が行われるため理論値に収束するが、実機ではノイズ管理が不可欠であることが確認された。量子ビット数を増やすことで理論上の解像度は向上する一方、重ね合わせ状態を維持するためのノイズ耐性が限界を超えると、情報の消失が生じ、結果として推定精度が低下することが明らかとなった。

文献

1. Stefan Woerne Quantum risk analysis,
,
2. Nikitas Stamatopoulos, et al., "Option Pricing using Quantum Computers"
,