

量子技術による無線システムの最適化およびプライバシー保護

代表研究者

李 傲寒

電気通信大学 情報理工学研究科 助教

1 はじめに

量子コンピューティングの高い計算能力により、従来の暗号化アルゴリズムは脅威にさらされるリスクが増加している [1]。量子通信は、測定誤差と擾乱、量子複製不可能定理、不確定性原理などの量子力学の原理に基づいており、盗聴の試みをリアルタイムで検出できるため、物理層における情報暗号化と通信セキュリティを実現する [2]。さらに、量子コンピュータの計算能力は、量子ビットの数に対して指数関数的に増加し、暗号解読、量子シミュレーション、量子最適化などの特定の問題において高速化の利点をもたらす [3]。そのため、量子技術は次世代の情報通信システムにおいて、データ伝送速度と信頼性の向上に重要な役割を果たしている [4]。諸外国は「量子技術」を経済・社会に大きな変革をもたらす革新技術と位置づけ、研究開発投資を大幅に拡充、拠点形成や人材育成等の戦略的取組を展開している [5]。2022 年 4 月には政府の新国家戦略「量子未来社会ビジョン」が策定され、量子技術によるデジタルトランスフォーメーション (DX) の加速はクォンタムトランスフォーメーション (QX) と強調された [6]。量子技術は AI やインターネットなど現代の ICT と組み合わせられることで、私たちの暮らしを豊かにし、大きな変革をもたらすと期待されている。また、量子技術は社会・経済的課題の解決に加え新しい学術分野の誕生にも貢献し、研究開発の新しい潮流となり将来の成長・発展を主導する重要技術でもある。しかしながら、量子情報技術はまだ発展途上にあり、技術的な複雑さや実装コスト、専用ハードウェアの必要性などの課題に直面している [7]。

現在、量子コンピューティングは Noisy Intermediate-Scale Quantum (NISQ) 時代にあり [8]、個々の量子処理ユニット (Quantum Processing Unit: QPU) の規模が限られていることが重要なボトルネックとなっている。一般に、現在の QPU は数百から数千個程度の物理量子ビットしか搭載していない [9], [10]。さらに、量子固有のノイズやハードウェア設計上の制約により、単一のモノリシックチップ上で量子ビット数を急速に増加させることは困難である。これらの制約を克服するため、分散量子計算 (Distributed Quantum Computing: DQC) は、産業界および学術界の双方から重要なアプローチとして認識され、注目を集めている [11]。DQC では、複数の QPU を量子ネットワークを介して相互接続し、単一の統合システムとして動作させることで、量子資源を集約する。これにより、単一の量子プロセッサの処理能力を超える大規模な量子回路の実行が可能となる。DQC のワークフローは、基本的に、量子ビット割り当て、エンタングルメント管理、およびネットワークスケジューリングという、相互に依存する三つの要素から構成される [12]。量子ビット割り当てでは、論理回路上の量子ビットを分散配置された物理量子ビットにマッピングする。この過程では、量子ビット間のゲートを、同一 QPU 内で実行されるローカルゲート (intra-QPU) と、異なる QPU 間で実行されるリモートゲート (inter-QPU) に分類する。リモートゲートは、高コストなエンタングルメント資源を必要とするとともに、大きな通信コストを伴うため、その実行回数を最小化することが主な目的となる。エンタングルメント管理では、リモート操作に必要な量子チャネルを確立する。これは一般に Cat-comm プロトコル [13] を用いて実現される。最後に、ネットワークスケジューリングでは、ゲート間の依存関係を厳密に考慮しながら、各操作の実行順序を調整する。例えば、同一の量子ビットを共有する操作は同時に実行することができず、順次実行する必要がある。

DQC では、論理回路をネットワーク接続された QPUs に効率的にマッピングしつつ、高コストな通信を最小化するという複雑な最適化問題が生じる。既存手法としては、統合型の UNIQ (Unified Nonlinear Integer) フレームワークなどがあるが、これらは静的なヒューリスティックに依存しているため、高次元の組合せ探索空間を効率的に探索することが困難であり、多くの場合において最適性に乏しい実行計画を生成する。そこで、本研究では、DQC コンパイルを自動化する新たな学習ベースのフレームワークとして、GARL (GNN-Attention-RL) を提案する。GARL は、関係グラフ畳み込みネットワーク (Relational Graph Convolutional Networks: R-GCNs) と深層強化学習 (Deep Reinforcement Learning: DRL) を組み合わせることで、DQC における複雑な最適化問題に対処する。ルールベース手法とは異なり、GARL はシステム状態を異種グラフとしてモデル化し、論理ゲート間の先行制約と、厳密に遵守すべき物理的資源制約との複雑な相互依存関係を捉えることができる。具体的には、各 QPU における有限の量子ビット容量と、動的に変化するエンタングルメント生成帯域幅を考慮する。さらに、ロバストな二段階学習パイプラインを導入する。第一段階では、模倣

学習 (Imitation Learning) により、エキスパートの行動を模倣することで方策をウォームスタートさせる。続く第二段階では、近接方策最適化 (Proximal Policy Optimization: PPO) を用いた強化学習により探索を行い、より優れた戦略を発見する。実験評価の結果、GARL は最先端の UNIQ ベースラインを大幅に上回り、大規模ベンチマークにおいて通信コストを平均 15%以上削減できることが示された。さらに、学習された方策は高いゼロショット汎化性能を示し、再学習を行うことなく、未知の回路トポロジや異なるハードウェア制約に効率的に適応できることが確認された。

量子プロセッサは、単一チップ上でのスケーリングの限界に近づきつつある。その有力な選択肢の一つとして、量子データセンター (Quantum Data Center: QDC) が考えられる。QDC では、複数の QPU が光インターコネクトによって接続され、モジュール間でエンタングルメント (量子もつれ) が生成される。このようなエンタングルした量子ビット対は、自由に利用できる通信リンクではない。エンタングルメントは、損失のある光伝送路および Bell 状態測定 (Bell-State Measurement: BSM) 検出器を介して、確率的に生成される必要がある。そのため、インターコネクトファブリックのトポロジは、システムの性能とコストに関わる二つの重要な側面に影響を及ぼす。これまでに、光 QDC ファブリックの提案およびプロトタイプ実装が行われており [14]、所与のファブリック上で通信をスケジューリングするコンパイラ手法 [15] や、異なるワークロードの下で固定トポロジを比較するベンチマーク手法 [16] も提案されている。これらの研究はいずれも、ファブリックのトポロジが既に構築・決定された後に、その性能を評価するものである。これとは別に、トポロジそのものを設計する研究も行われている。例えば、回路の通信要求に基づいて QDC のトポロジを生成する手法 [17]、量子ビット割り当てとトポロジを協調的に設計する手法 [18]、さらに、所与の需要行列を最小コストで処理するために、トポロジ、量子メモリ、およびエンタングルメントスワッピングのスケジュールを同時に選択する手法 [19] などがある。これらの研究では、量子メモリやスワッピング機構を含むシステム構成を設計し、対象となる需要を満たすための総コストを最小化することを目指している。これに対して、本研究では、デプロイメントの問題を逆の観点から捉える。すなわち、量子データセンターで実行されるワークロードのクラスが与えられ、さらにスイッチおよび検出器に対する予算があらかじめ定められている場合に、どのようなインターコネクトトポロジを構築すべきかという問題を扱う。

2 分散量子計算のための最適化フレームワーク

2-1 予備知識

図 1 に、3 台の QPU から構成される QDC システムの一例を示す。QDC のワークフローは、量子ビット割り当て、エンタングルメント管理、およびネットワークスケジューリングという、相互に依存する三つのフェーズから構成される複雑なシステム最適化問題である。

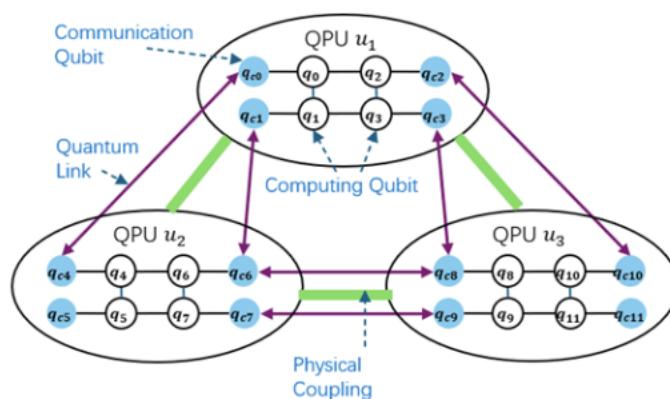


図 1 3 台の QPU から構成される QDC システムの一例

2-2 問題設定

QDC 最適化問題は、本質的に、厳密な物理的資源制約の下で、異種グラフ上における多目的マッピングおよびスケジューリング問題として捉えることができる。グラフ理論を用いてシステムモデルを厳密に定式化し、整数線形計画法により最適化目的を定義する。最終的に、この問題を逐次的なマルコフ意思決定過程と

して再定式化する．この再定式化により，深層強化学習を活用して組合せ解空間を探索することが可能となる．さらに，この解空間が非決定性多項式時間 (Nondeterministic Polynomial time: NP) 困難であることを証明する．全体の最適化問題は、以下に示す二つの目的関数を組み合わせた形で定式化される．

$$\min_{\mathbf{x}, \mathbf{s}} J = \alpha \cdot T_{\max} + \beta \cdot C_{\text{comm}},$$

ここで， α および β は，実行速度とネットワークオーバーヘッドとのバランスを調整するハイパーパラメータである． C_{comm} は以下の式で表される．

$$C_{\text{comm}} = \sum_{g_j \in \mathcal{G}} \sum_{u_k \neq u_l} w_{kl} \cdot x_{a,k} \cdot x_{b,l},$$

ここで， w_{kl} は QPU k と l の間の通信遅延コスト， $x_{a,k}$ と $x_{b,l}$ は二値マッピング変数です．量子ビット a が QPU k にマッピングされる場合 $x_{a,k}$ は1となり，量子ビット b が QPU l にマッピングされる場合 $x_{b,l}$ は1となる． $T_{\max}$ は以下の式で表される．

$$T_{\max} = \max_{g_j \in \mathcal{G}} (s_j + d(g_j))$$

ここで， s_j は整数スケジューリング変数であり，ゲート j の開始時刻である． $d(g_j)$ はゲート j の実行時間である．

2-3 提案法

定式化された NP 困難な組合せ最適化問題を解決するため，GARL フレームワークを提案する．本フレームワークでは，従来のヒューリスティック手法とは異なり，意思決定方策を深層ニューラルネットワークとしてパラメータ化し，量子回路および分散アーキテクチャが有する非ユークリッド・トポロジーを捉えられるように設計する．さらに，ニューラルネットワークアーキテクチャを数学的に厳密に記述し，関係グラフ畳み込みネットワークを用いたエンコーダと，Attention に基づくデコーダの構成を詳細に示す．加えて，提案アーキテクチャの理論的性質を分析するとともに，包括的な二段階学習パイプラインを通じて，その学習ダイナミクスを定式化する．システムワークフローは，三つの統合されたステージを連続的なサイクルとして構成する．まず，状態構築フェーズでは，論理量子回路の有向非巡回グラフと物理ハードウェアトポロジーからなる入力情報を統合し，統一された異種グラフを構築する．この構造化された表現をニューラルネットワークエージェントに投入し，エンコーダ・デコーダアーキテクチャによって処理する．具体的に，関係グラフ畳み込みネットワークエンコーダによりグラフから高次元のトポロジー特徴を抽出し，これらの特徴を Attention 機構に基づくデコーダによって処理することで，動的な行動空間上の有効な確率分布を算出する．最後に，最適化ループへ移行し，エージェントが行動をサンプリングして DQC 環境と相互作用する．得られたスカラー報酬は，模倣学習と Proximal Policy Optimization (PPO) を組み合わせたハイブリッド学習戦略を通じて，方策パラメータの更新にフィードバックされる．初期フェーズでは，既存の UNIQ ヒューリスティックソルバをエキスパート実演者として利用する．その目的は，有効かつ十分に効率的なスケジューリング戦略に対応するパラメータ空間の領域に，ニューラル方策を初期化することである．まず，エキスパート軌跡のデータセットを生成する．最適化の目的は，学習方策とエキスパート方策との間の Kullback-Leibler ダイバージェンスを最小化することであり，これは負の対数尤度損失の最小化に帰着する．これは，模倣学習において標準的に用いられるアプローチである．この教師あり事前学習により，エージェントは DQC スケジューリングにおける基本的な「文法」，すなわち依存関係や容量制約を遵守する方法を迅速に習得できる．模倣学習によって良好な初期解を得られる一方で，エージェントの性能は，ヒューリスティックなエキスパートが有する非最適性によって制約される．そこで，エキスパートを上回る性能を実現するため，PPO を用いた強化学習へ移行する．ここでは，割引期待累積報酬の最大化を目的とする．PPO では，方策更新のステップサイズを制約する代理目的関数を最適化することで，過度に大きな方策更新を抑制し，単調な性能改善を実現する．

2-4 性能評価

性能評価環境は，DQC に特化した高忠実度の離散イベントシミュレータとして構築する．単純化された抽象モデルとは異なり，本シミュレータでは，定義した相互に結合した物理的制約を厳密に適用する．具体的には，各 QPU が有する有限の量子ビット容量と，エンタングルメント生成リンクにおける動的な帯域幅制約をモデル化する．また，シミュレータは個々のゲート操作単位でシステム状態を追跡し，マルコフ意思決定過程におけるすべての状態遷移が，量子ハードウェア状態の物理的に妥当な変換に対応することを保証する．

シミュレーションプラットフォームおよびニューラルエージェントは、PyTorch 深層学習フレームワークを用いて Python で実装する。強化学習における大規模なロールアウトフェーズでの計算効率を確保するため、学習パイプラインは NVIDIA A100 GPU を搭載した高性能計算ノード上で実行し、テンソル演算およびグラフ畳み込みを高速化する。

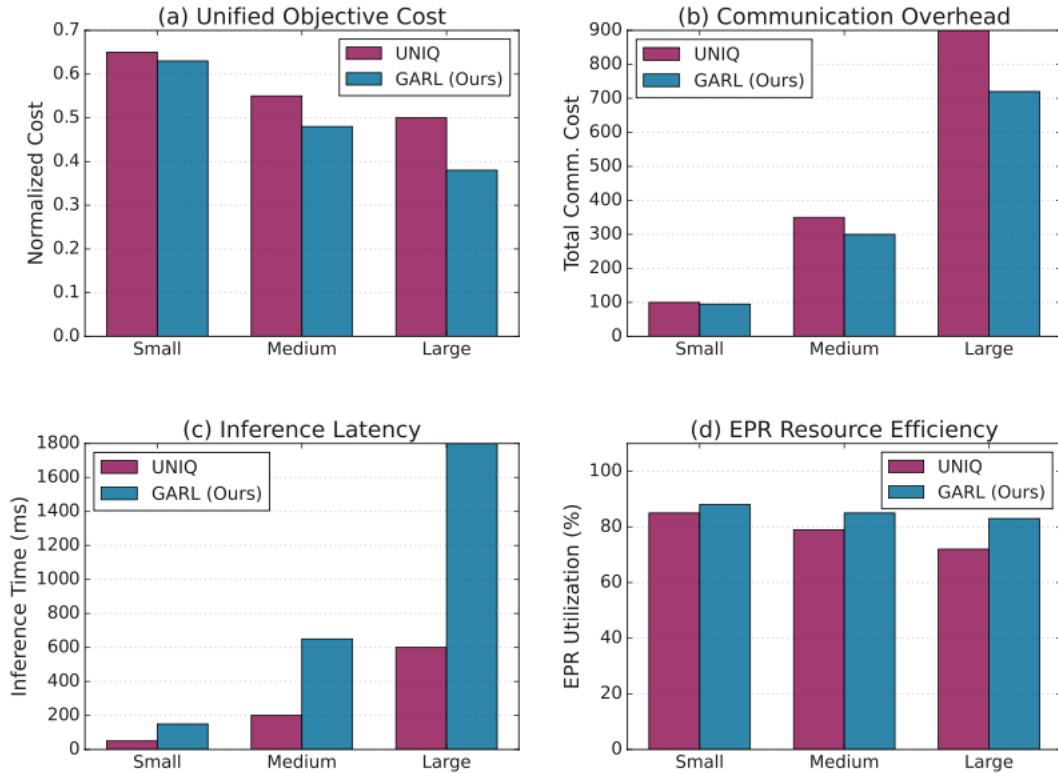


図 2 性能評価

(1) スケール全体にわたる総合的な有効性

GARL のロバスト性を検証するため、合成データセットで定義された全ての問題規模を対象として、その性能を評価する。図 2(a) に、GARL と UNIQ エキスパートを直接比較した正規化統合コストを示す。結果から、GARL は全ての問題規模において統計的に有意な性能向上を達成していることが分かる。ヒューリスティック手法が一般に最適解に近い性能を達成する小規模な回路では、性能差は比較的小さい。一方、問題の複雑性が増加するにつれて、GARL の優位性は大幅に拡大する。大規模な問題インスタンス (80～120 量子ビット) では、GARL は UNIQ と比較して平均 20% 以上のコスト削減を達成する。この性能差の拡大は、提案するグラフベースの方策が、スケーラブルなヒューリスティック手法では局所的な意思決定の性質ゆえに捉えることが困難な、グローバルなトポロジー依存関係を効果的に捉えられることを示している。

(2) 通信オーバーヘッド、推論遅延および EPR 利用率

これらの性能向上の要因を明らかにするため、図 7(b)～(d) に示すように、各問題規模における性能を具体的な評価指標ごとに詳細に分析する。その結果、GARL は UNIQ と比較して一貫して低い通信オーバーヘッドを実現しており、特に大規模な問題領域 (Large) において、その削減効果が顕著であることが分かる。この傾向は、Attention 機構が高いエンタングルメント性を有する量子ビットを効果的にクラスタリングし、高コストな QPU 間転送を最小限に抑えていることを示している。一方、GARL では、ニューラルネットワークのフォワードパスに伴う計算コストにより、軽量のヒューリスティック手法 (約 200 ms) と比較して推論遅延が増加する (Medium 規模で約 650 ms)。しかし、このオーバーヘッドはコンパイル時に一度だけ発生するオフラインコストであるため、コンパイル処理全体に占める影響は限定的である。量子回路は通常、統計的な読み出しのために数千回 (shots) 繰り返し実行されることを考慮すると、この

ミリ秒オーダーのコンパイル時間の増加は、ワークロード全体にわたる通信遅延の大幅な削減と実行忠実度の向上によって十分に償却される。さらに、問題規模が増大した場合でも、GARL のアインシュタイン・ポドルスキー・ローゼン(EPR)利用率は約 85%という高い水準を安定して維持しているのに対し、ヒューリスティック手法ではその効率が低下する。この結果は、GARL がゲート実行とエンタングルメント生成サイクルを効果的に同期させる、より高度な時間的スケジューリング方策を学習していることを示している。

(3) 詳細な性能評価

現実的なワークロードに対する提案手法の有効性を厳密に検証するため、QASMBench に含まれる 12 種類の異なる量子アルゴリズムを対象として、大規模な実験を実施する。表 1 に、UNIQ エキスパートと GARL との比較結果を示す。提案モデルは、多様な回路トポロジーにおいて、一貫してエキスパートを上回る性能を達成している。qugan_n71 のように、長距離依存関係が顕著で、かつ深く高いエンタングルメント性を有する回路では、GARL は通信オーバーヘッドを 12.4%大幅に削減している。この結果は、Attention 機構が回路の接続性において中心的な役割を果たす「ハブ」量子ビットを効果的に識別し、その配置を優先できることを示している。同様に、局所的な相互作用が密な qaoa_n14 回路においても、提案モデルは 9.8%の性能向上を達成している。さらに分析した結果、GARL は通信コストの削減を実行時間の犠牲によって達成しているわけではなく、むしろメイクスパン (makespan) も改善する場合が多いことが明らかになった。例えば、supremacy_n120 では、実行時間を 13.8%短縮している。このような同時最適化により、GARL が多目的最適化問題におけるパレートフロンティアを効果的に探索できることが確認される。標準的なコストおよび実行時間の指標に加えて、資源利用の品質を評価する指標として、EPR 効率を導入する。表 1 に示すように、GARL は一貫して高い効率を達成しており、平均効率は UNIQ の 76%に対して 82%に達する。特に、supremacy_n120 のような通信負荷の大きい回路において、その優位性が顕著である。この結果は、強化学習方策が単にコストを最小化するだけでなく、資源競合を能動的に回避することで、より円滑な実行フローを実現していることを示している。

表 1 詳細な性能比較

Circuit	Qubits	Gates	Communication Cost			Runtime (slots)		EPR Efficiency (%)		Objective Improv. (%)
			UNIQ	GARL	Δ	UNIQ	GARL	UNIQ	GARL	
adder_n4	4	23	120	115	-4.2%	45	44	85.0	87.5	+3.8%
qaoa_n14	14	56	450	406	-9.8%	82	78	78.5	83.2	+9.8%
bv_n30	30	29	210	195	-7.1%	35	35	88.0	90.5	+5.5%
qft_n50	50	320	3,200	2,950	-7.8%	410	395	75.4	81.0	+6.2%
cat_n60	60	60	600	580	-3.3%	12	12	92.0	94.5	+3.3%
ghz_n80	80	79	850	790	-7.1%	15	14	81.5	86.0	+7.5%
vqe_n80	80	450	4,100	3,650	-11.0%	520	495	72.0	79.5	+9.2%
qugan_n71	71	620	6,800	5,950	-12.5%	680	645	70.5	78.8	+12.4%
qpe_n90	90	850	9,200	8,100	-12.0%	950	890	69.0	77.5	+10.5%
bv_n100	100	99	1,100	980	-10.9%	45	42	84.5	89.0	+11.2%
supremacy_n120	120	1,200	15,400	13,200	-14.3%	1,400	1,280	65.0	75.2	+13.8%

3 ワークロード駆動型エンタングルメントファブリック設計

本研究では、ワークロード駆動型の QDC エンタングルメントファブリック設計手法である WEFT を提案する。WEFT では、各候補ファブリックを以下の 3 段階で評価する。第 1 に、コンパイル済みの量子回路を、モジュール間のエンタングルメント需要を表す行列に変換する。第 2 に、各候補経路について、ヘラルド型エンタングルメント生成試行の成功確率に基づいて評価する。この成功確率には、光損失および検出器効率が考慮される。第 3 に、線形計画問題を解く。この線形計画問題では、限られたリンク容量および BSM 試行レートを各エンタングルメント需要に割り当てる。これにより、持続可能な需要係数、すなわち、ファブリックの実行可能性を維持したままワークロードの需要を最大で何倍まで拡大できるかを表す係数が得られる。その後、外側の探索において、この評価結果に基づき、所定のハードウェア予算の下でファブリックを選択する。WEFT は、トポロジーファミリの選択、そのパラメータ化、およびワークロード駆動型ショートカットの追加を統合的に行う。本研究の主な貢献は以下のとおりである。1) 定式化: QDC エンタングルメントファブリックの設計を、ワークロード、物理特性、および予算を考慮した最適化問題として定式化する。モジュール間のエンタングルメント需要が与えられた場合、WEFT は損失を考慮した最大同時フローモデルに基づ

いて構造化されたファブリックを評価し、その持続可能な需要係数を算出する。2) 物理特性を考慮した分析：物理特性に基づく重み付きカット指標を導入し、代表的なファブリックファミリーに対する閉形式の容量表現を導出する。特定の iso-QPU, iso-BSM 条件下におけるトラス/ファットツリーの組合せについて、これらの結果から、スイッチ透過率に関する閉形式のクロスオーバー閾値を導出する。3) ロバストな設計と知見：予算制約を満たすトポロジーファミリーと、ワークロード駆動型ショートカットによる拡張を対象として探索を行う。候補の順位が保証される $N=16$ および $N=64$ では、ミニマックス後悔に基づく選択により、QLS ファミリが選択される。QLS は、リーフノード間を循環グラフで接続したファットツリーである。一方、 $N=256$ では、保証された区間が重なっている。

2-1 システムモデル

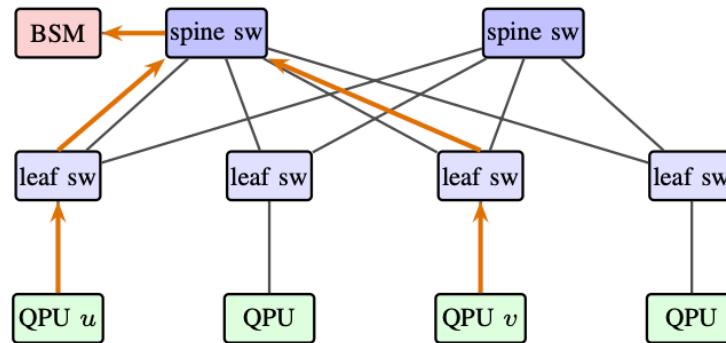


図3 システムモデル

QDC を全光型の回線交換ファブリックによって接続された N 個の計算モジュールから構成されるシステムとしてモデル化する。各計算モジュールは、 C_{mod} 個の量子ビットと光インタフェースを備えた QPU である。ファブリックは、光スイッチから構成されるグラフとして表現される。直接接続型では、各 QPU に小ポート数の光スイッチが併設される。トラスおよびハイパーキューブは直接接続型に分類される。一方、間接接続型では、QPU をリーフスイッチに接続する。ファットツリーおよびドラゴンフライは間接接続型に分類される。BSM ステーションはスイッチに接続される。2つのモジュール間でエンタングルメントを生成する際には、両方の QPU が光子を放出する。光子は光ファイバを伝搬し、透過型スイッチを通過して、2つのアームを経由して共通の BSM ステーションに到達する。BSM ステーションでは、2回の光子検出によってエンタングルメント対の生成がヘラルドされる。本ファブリックには量子メモリが存在せず、エンタングルメントスワッピングも行わない。この構成は、スイッチ型 QDC を対象としたコンパイラレベルの研究で仮定されているメモリレススイッチの前提と整合する。図3に、小規模なファットツリーを例として、このファブリックの構成を示す。本研究では、ファブリック設計によって、スイッチグラフ、QPU の接続構成、BSM ステーションの配置、および各デバイスの容量が定まるものとする。一方、トポロジーはグラフファミリーのみを指すものとする。本モデルは、ワークロード層、物理層、および運用層の3つの層から構成される。

2-2 提案法

WEFT は、予算制約を満たす構造化されたトポロジーファミリーを探索し、そのパラメータとワークロード駆動型の拡張を統合的に最適化する。そして、単一のワークロードに特化して調整された設計ではなく、複数のワークロードにわたる最悪後悔値が最小となる設計を選択する。外側探索についてですが、設計問題は、二つのレベルから構成される。内側の線形計画問題では、各シナリオにおける個々のファブリックを評価し、

外側の探索では、予算制約の下で設計空間 Θ に含まれる候補から最適な設計を選択する。シナリオ $s=(\bar{d}, \beta, \psi)$

は、ワークロードの需要記述子と物理層におけるエンタングルメント生成特性 ψ を組み合わせたものである。選択の目的指標 α_c は、需要記述子に依存する。一方、バースト性 β は、ブロッキング解析に用いるため、シナリオの構成要素として保持される。ここで、 ψ は $\eta_{sw} \in \{0.93, 0.79, 0.63\}$ の値を取り、それぞれ-0.3dB, -1 dB, および -2 dB のスイッチ損失に対応する。探索計算量を削減するため、三つの枝刈り規則を適用する。まず、コスト制約を満たさない候補は、評価前に除外する。次に、物理特性重み付き二分レートで与えられる閉形式の上界をカット需要で除算することで α_c の上界を求め、この上界によって劣化する線形計画問題の求解を省略する。さらに、対称なパラメータ化を持つ候補については、重複するものを除外する。また、サン

プリングされたカット比 ρ を用いて実行不可能性を判定する。 $\rho < 1$ であれば $\alpha_c < 1$ が成立するため、当該候補が実行不可能であることを保証できる。したがって、 α_c に加えて、この ρ も併せて報告する。横方向リンク予算制約付きショートカット拡張についてですが、横方向のショートカットを1本追加することで、 $H = 4$ のスパイン経路を $H = 3$ の経路に短縮でき、各ペアについて η_{sw}^{-1} の利得が得られるとともに、スパイン側の容量予算に対する負荷を軽減できる。ファットツリーと s_l 本の横方向リンクに対する予算が与えられた場合、この拡張問題では、リンク部分集合を選択することによって α_c を最大化する。最小比率に基づく目的関数は劣モジュラではないため、この性質を考慮する必要がある。実際、以下に示すあるワークロードでは、単一のリンクを追加しても α_c は改善されない一方、2本のリンクを組み合わせて α_c が7.4%向上する。これは、1本目のリンクによってボトルネックが移動し、その結果、2本目のリンクによって新たなボトルネックを緩和できるためである。そこで、探索には貪欲法による逐次追加にペア単位の先読みを組み合わせる。すべてのリンク部分集合は、単一の固定された経路集合上で評価する。具体的には、すべての候補リンクを含む状態で経路集合を列挙することにより、各部分集合に対応する実行可能集合が包含関係を満たすようにする。列挙可能なすべてのインスタンスにおいて、先読み付き貪欲法は全探索によって得られる最適解を達成する。より大規模な問題では、分数緩和によって保証された上界を導出する。各候補リンク e に対して、 $y_e \in [0,1]$ を満たす容量変数 y_e を導入し、その試行予算をスケールリングする。また、 $\sum_e y_e \leq s_l$ とする。整数値の s_l 本のリンク部分集合はすべてこの緩和問題の実行可能解であるため、緩和問題の最適値は整数問題の最適値に対する上界となる。実験で測定された整数性ギャップは最大7.4%である。この拡張方式に基づき、 Θ には2つのパラメータ化されたハイブリッドトポロジーファミリを導入する。QLSファミリは、ファットツリーのリーフノード間に循環グラフを追加し、そのリング次数をパラメータとして選択する。DLSファミリは、リングによって拡張されたドラゴンフライであり、通常の全対全グローバルリンクに加えて、リング上で隣接するグループ間にリンクを追加する。両ファミリでは、追加されるポートおよび光ファイバのコストをコストモデルに基づいて計上し、 Θ に含まれる通常の候補として評価する。

2-3 性能評価

評価結果から、設計選択がどの要因に敏感であり、どの要因に対して頑健であるかが明らかになる。設計選択はハードウェア予算に敏感であり、予算の変化によって最適なトポロジーファミリそのものが変化する。また、 $N = 16$ ではワークロードにも敏感であり、ハブクラスでは他のワークロードとは異なるリングパラメータが選択される。一方、モジュールのラベリング、検出器の配置、パーティションのシード、およびコストパラメータの変化に対しては、設計選択は影響を受けない。さらに、クロスオーバー則を検証し、ミニマックス後悔に基づく選択が、ショートカットによる拡張を行わないすべてのベースラインを上回ることを示す。また、自由形式のシンセサイザとの比較を行い、試行レートに基づく指標がスロットレベルの挙動を適切に反映することを確認する。

4 まとめ

本研究では、新たな学習ベースの最適化フレームワークであるGARLを提案し、分散量子計算に内在するスケラビリティ上の重要な課題に取り組んだ。従来の静的なルールベースのヒューリスティック手法とは異なり、GARLでは、量子ビット割り当て、エンタングルメント管理、およびネットワークスケジューリングを含む統合コンパイル問題を、深層ニューラル方策によって制御される逐次的な意思決定過程として再定式化する。この設計により、システムは局所的なヒューリスティック手法の限界を克服し、グローバルなトポロジー情報と長期的な計画能力を活用することで、通信オーバーヘッドと実行効率との間に存在する複雑なトレードオフを効果的に探索できる。これにより、固定的なルールベース手法では捉えることが困難な解を発見することが可能となる。本研究の理論的および実証的な貢献は、以下の3点にまとめられる。第一に、異種DQCシステムの状態を関係グラフ構造へ写像する厳密な数学モデルを導入した。関係グラフ畳み込みネットワークを活用することで、論理ゲート間の実行順序と物理的な資源制約との間に存在する複雑な非ユークリッド的依存関係を効果的に捉え、ノードの順列に対して不変な高忠実度の状態表現を実現する。第二に、Attention機構に基づくデコーディング機構を設計し、組合せ的な行動空間を動的に探索する。ことで、エージェントがクリティカルパス上の処理に重点的に着目できるようにした。第三に、UNIQエキスパートによる模倣学習を通じて初期方策を獲得し、その後PPOによって方策を改善する二段階の学習戦略が、この高次元の最適化問題において安定した収束を実現するために不可欠であることを示した。さらに、本研究では、実行されるワークロードに基づいて、量子データセンタの光インターコネクトをどのように設計すべきかを

検討した。WEFT は、量子回路をモジュール間のエンタングルメント需要に変換し、物理特性を考慮したヘラルド型エンタングルメント生成モデルを用いて候補ファブリックを評価するとともに、スイッチおよび検出器に対する所定の予算制約の下で設計を選択する。問題規模が大きくなるにつれて、ワークロードに応じた設計選択の自由度は小さくなる。16 モジュールでは、2 種類のリングパラメータ設定により、シナリオごとに選択される設計が分かれる。一方、64 モジュールでは、単一の設計がすべてのシナリオにおいて選択される。256 モジュールでは、保証された区間が重なっているため、候補間の優劣を一意に決定することはできない。候補間の順位関係が保証される 2 つの問題規模では、リングによって拡張されたハイブリッドトポロジーが、対応する予算制約の下で選択される。一方、より厳しい予算制約の下では、ランダム正則グラフが選択される。これらの結果は、量子メモリを持たず、エンタングルメントスワッピングを行わない単一 BSM ファブリックを前提とした場合に成立する。本研究は、量子計算を活用した次世代無線システムにおける大規模最適化問題およびネットワークセキュリティ問題の解決に向けて、理論的基盤と技術的基盤を提供するものである。今後の研究では、実際の無線システムを対象とした実験を通じて、提案手法の有効性を検証する予定である。

【参考文献】

- [1] L. Hanzo, et. al., “Quantum Information Processing, Sensing, and Communications: Their Myths, Realities, and Futures,” *Proceedings of the IEEE*, DOI: 10.1109/JPROC.2024.3510394, Jan. 2025.
- [2] W. Jiang, Y. Zhang, H. Han, and J. Mu, “Quantum Communication in Self-Organizing Satellite Networks: Challenges and Opportunities,” *IEEE Communications Standards Magazine*, DOI: 10.1109/MCOMSTD.2025.3578320, June 2025.
- [3] M. O. Butt, N. Waheed, T. Q. Duong, and W. Ejaz, “Quantum-Inspired Resource Optimization for 6G Networks: A Survey,” *IEEE Communications Surveys & Tutorials*, DOI: 10.1109/COMST.2024.3519865, Dec. 2024.
- [4] X. Zhou, A. Shen, S. Hu, W. Ni, X. Wang, and E. Hossain, “Towards Quantum-Native Communications Systems: State-of-the-Art, Trends, and Challenges,” *arXiv:2311.05239*, May 2025.
- [5] 佐々木, “量子技術分野の動向と今後の課題,” 情報通信審議会・技術戦略委員会, 総務省, 2022.
- [6] 研究開発の俯瞰報告書 システム・情報科学技術分野, JST 研究開発戦略センター, 2024.
- [7] D. Pan, G. Long, L. Yin, Y. Sheng, D. Ruan, S. Ng, J. Lu, and L. Hanzo, “The Evolution of Quantum Secure Direct Communication: On the Road to the Qinternet,” *IEEE Communications Surveys & Tutorials*, vol. 26, no. 3, 2024.
- [8] K. Ju, H. Zhong, X. Zhang, X. Qin, and M. Pan, “Quantum computing catalyzes future quantum networks: Efficiency and inherent privacy,” *IEEE Network*, vol. 39, no. 1, pp. 124–131, May 2024.
- [9] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. Brandao, D. A. Buell et al., “Quantum supremacy using a programmable superconducting processor,” *Nature*, vol. 574, no. 7779, pp. 505–510, October 2019.
- [10] Y. Li, Y. Zhao, X. Zhang, H. Zhong, M. Pan, and C. Zhang, “Differential privacy preserving quantum computing via projection operator measurements,” in *International Conference on Quantum Communications, Networking, and Computing (QCNC)*, Kanazawa, Japan, July 2024.
- [11] M. Caleffi, A. S. Cacciapuoti, and G. Bianchi, “Quantum internet: From communication to distributed computing!” in *ACM international conference on nanoscale computing and communication (NanoCom’18)*, New York, NY, September 5-7 2018.
- [12] D. Cuomo, M. Caleffi, and A. S. Cacciapuoti, “Towards a distributed quantum computing ecosystem,” *IET Quantum Communication*, vol. 1, no. 1, pp. 3–8, July 2020.
- [13] Y. Mao, Y. Liu, and Y. Yang, “Qubit allocation for distributed quantum computing,” in *IEEE Conference on Computer Communications (INFOCOM’23)*, New York, NY, May 17-20, 2023, pp. 1–10.
- [14] H. Shapourian, E. Kaur, T. Sewell, J. Zhao, M. Kilzer, R. Kompella, and R. Nejabati, “Quantum data center infrastructures: A scalable architectural design perspective,” 2025.

- [15] H. Zhang, Y. Xu, H. Hu, K. Yin, H. Shapourian, J. Zhao, R. R. Kompella, R. Nejabati, and Y. Ding, “SwitchQNet: Optimizing distributed quantum computing for quantum data centers with switch networks,” in Proceedings of the 52nd Annual International Symposium on Computer Architecture, 2025, pp. 1449–1463.
- [16] S. Pouryousef, E. Kaur, H. Shapourian, D. Towsley, R. Kompella, and R. Nejabati, “Benchmarking quantum data center architectures: A performance and scalability perspective,” 2026.
- [17] Y. Mao, Y. Liu, X. Shang, and Y. Yang, “Network topology design for distributed quantum computing,” in 2024 IEEE 44th International Conference on Distributed Computing Systems (ICDCS), 2024, pp. 1213–1223.
- [18] J. Liu, L. Fan, Y. Guo, Z. Han, and Y. Wang, “Co-design of network topology and qubit allocation for distributed quantum computing,” in 2025 International Conference on Quantum Communications, Networking, and Computing (QCNC), 2025, pp. 355–362.
- [19] Y. Gan, R. Zhou, S. Ling, Y. Liu, and C. Qian, “On designing a quantum network topology: Model, methodology, and algorithm,” in 2025 IEEE International Conference on Quantum Computing and Engineering (QCE), 2025, pp. 1207–1218.

〈発 表 資 料〉

題 名	掲載誌・学会名等	発表年月
GARL: A Unified GNN-Attention and Reinforcement Learning Optimization Framework for Distributed Quantum Computing	IEEE Journal on Selected Areas in Communications	2026 年 1 月(査読中)
WEFT: Workload-Driven Design of Entanglement Fabrics for Quantum Data Centers	IEEE INFOCOM 2027	2026 年 6 月(査読中)