

물리 지식 기반 그래프 트랜스포머 강화학습을 이용한 강건한 양자 배터리 충전 최적화 연구

박범도¹, 박준성¹, 장현석¹, 정훈², 허태욱², *이상금¹

*국립한밭대학교¹, 한국전자통신연구원²

{pbeomdo, js03093351, seokchu123}@gmail.com, {hjeong, htw398}@etri.re.kr,
sangkeum@hanbat.ac.kr

Robust Charging of Quantum Batteries via Physics-Informed Graph Transformer-Based Reinforcement Learning

Beomdo Park¹, Junseong Park¹, Hyeonseok Jang¹, Hoon Jeong², Taewook Heo²,
*Sangkeum Lee¹

*Hanbat National University¹, Electronics and Telecommunications Research Institute²

요약

스핀 체인 구조의 양자 배터리는 큐비트 수가 증가할수록 에너지를 시스템 전체로 전달하고 상태를 유지하는 제어 난이도가 상승한다. 기존의 완전 연결 기반 그래프 신경망은 큐비트 간의 물리적 위상을 고려하지 않아, 불필요한 장거리 상관관계를 학습함으로써 잡음이 섞인 불안정한 제어 신호를 생성하는 경향이 있다. 이를 개선하기 위해 물리적 거리에 따라 상호작용 가중치를 조절하는 거리 기반 어텐션 가중치를 적용한 Physics-GTN(Graph Transformer Networks)과, 목표 도달 및 유지를 동시에 고려한 조화평균 보상 함수를 제안한다. 8 큐비트 환경 실험 결과, 제안 기법은 기존 Standard GTN 대비 제어 신호의 총 변동량을 약 16 배 감소시켜 신호 왜곡을 억제하고, 최대 충전율 0.93 및 유지 비율 0.99를 달성하여 안정적인 충전 제어를 검증했다.

I. 서 론

양자 배터리는 양자 얹힘 현상을 이용하여 기존 배터리보다 빠르게 에너지를 충전할 수 있는 차세대 에너지 저장 장치이다. 특히 1 차원 스핀 체인 모델은 물리적 구현이 용이하여 주목받고 있으나, 제한된 연결성으로 인해 시스템 전체에 에너지를 고르게 분산시키면서 해당 상태를 안정적으로 유지하는 최적 제어 문제를 해결하는 것이 도전적이다. 최근 기계학습 분야에서는 이러한 고차원 제어 문제를 해결하기 위해 어텐션 메커니즘을 강화학습에 적용하는 접근법이 주목받고 있다[1]. 그러나 표준 트랜스포머의 전역적 어텐션 메커니즘은 모든 노드 간의 관계를 동등하게 계산하므로, 물리적으로 인접하지 않은 큐비트 간의 미세한 잡음까지 과도하게 학습하여 제어 신호의 품질을 저하시키는 원인이 된다[2]. 또한 큐비트 수 N 이 증가함에 따라 연산량이 $O(N^2)$ 으로 급증하여 확장성에 제약이 발생한다. 이에 본 연구에서는 양자 소자의 물리적 구조 정보를 내재화한 Physics-GTN(Graph Transformer Networks)을 제안한다. 제안 모델은 큐비트 간 거리가 멀어질수록 어텐션 영향력을 지수적으로 감소시키는 거리 기반 어텐션 가중치를 도입하고, 연산 복잡도를 선형 수준으로 경량화한다. 아울러 에이전트가 학습된 고에너지 상태를 유지하지 못하고 방전되는 현상을 방지하기 위해 최대 충전량과 최종 충전량의 조화평균을 보상으로 설계하여 학습 안정성을 확보했다.

II. 물리 지식 기반 그래프 트랜스포머 (Physics-GTN)

2-1. 1 차원 스핀 체인 환경

연구 대상은 N 개의 큐비트가 선형으로 연결된 1 차원 하이젠베르크 스핀 체인이다. 시스템 해밀토니안은 다음과 같이 정의된다.

$$H = J \sum_{i=1}^{N-1} Z_i Z_{i+1} + \sum_i h_x^{(i)}(t) X_i + h_y^{(i)}(t) Y_i$$

여기서 $J = 0.5$ 는 인접 큐비트 간 결합 상수이며, $h_x^{(i)}(t)$ 와 $h_y^{(i)}(t)$ 는 에이전트가 제어하는 외부 구동장이다. 이율 $\gamma_1 = 0.002$, 위상 결맞음율 $\gamma_2 = 0.001$ 이 적용된 개방 양자계로 모델링되며, 에이전트는 바닥 상태 $|00 \dots 0\rangle$ 에서 완전 충전 상태 $|11 \dots 1\rangle$ 로 시스템을 전이시켜야 한다.

2-2. 희소 그래프 트랜스포머 구조

1 차원 스핀 체인의 물리적 특성을 신경망 구조에 반영하기 위해 다음 세 가지 기법을 적용했다.

첫째, 거리 기반 어텐션 편향 D_{ij} 를 어텐션 점수에 주입한다.

$$A_{ij} = \text{softmax} \left(\frac{Q_i K_j^T}{\sqrt{d_k}} + D_{ij} \right)$$

D_{ij} 는 $e^{-0.3|i-j|}$ 형태로 초기화되어 인접 큐비트에는 높은 가중치를, 원거리 큐비트에는 낮은 가중치를 부여한다. 이는 하이젠베르크 체인에서 상관관계가 거리에 따라 지수적으로 감소하는 물리적 특성을 모방한 것으로, 학습 과정에서 원거리 큐비트의 잡음 영향을 억제한다.

둘째, k -hop 희소 어텐션을 적용하여 $|i-j| \leq k$ 인 이웃에만 어텐션을 계산한다. 표준 트랜스포머의 $O(N^2)$ 복잡도와 달리 제안

모델은 $O(kN)$ 연결만 처리하여, 8 큐비트 기준 파라미터 수가 55% 감소한다.

셋째, 전역 요약 토큰을 도입하여 양자 얹힘에 의한 비국소적 상관을 포착한다. 이 가상 노드는 모든 큐비트 정보를 집계하여 희소 어텐션으로 인한 장거리 정보 손실을 보완한다.

2-3. 조화평균 기반의 안정적 보상 설계

목표 충전율을 도달 후 제어 실패로 인한 상태 붕괴를 방지하기 위해, 현재 충전율 F_{end} 와 에피소드 내 최대 충전율 F_{max} 의 조화평균을 최종 보상에 반영한다.

$$R_{final} = \frac{2 \cdot F_{max} \cdot F_{end}}{F_{max} + F_{end}}$$

조화평균은 두 값 중 하나라도 낮으면 전체 보상이 급격히 감소하므로, 높은 충전율에 도달하더라도 이를 유지하지 못하면 낮은 보상을 받게 된다. 이 설계는 고에너지 상태의 안정적 유지를 유도한다[4].

III. 실험 및 결과

3-1. 실험 설계

제안 기법 검증을 위해 8 큐비트 환경에서 실험을 진행한다. 1 에피소드는 최대 120 스텝이며, 시간 간격 (dt)은 0.02이다. 에이전트가 충전율 0.98에 도달하면 초기 종료되도록 설계하여, 효율적인 제어 경로 학습을 유도한다. 비교 모델로는 표준 그래프 트랜스포머(Standard GTN)를 사용하며, 동일한 SAC(Soft Actor-Critic) 알고리즘 하에서 220 에피소드 동안 학습했다[3].

3-2. 제어 신호 품질 및 스펙트럼 분석

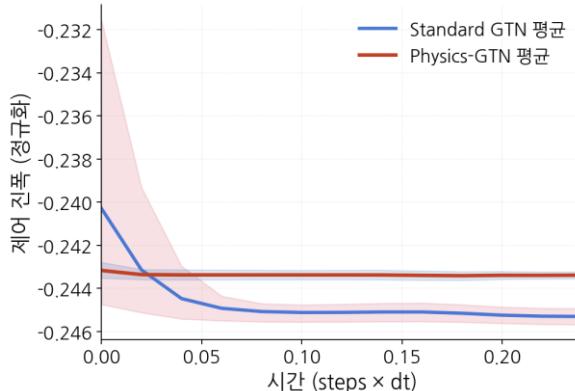


그림 1. 학습된 에이전트가 생성한 제어 진폭

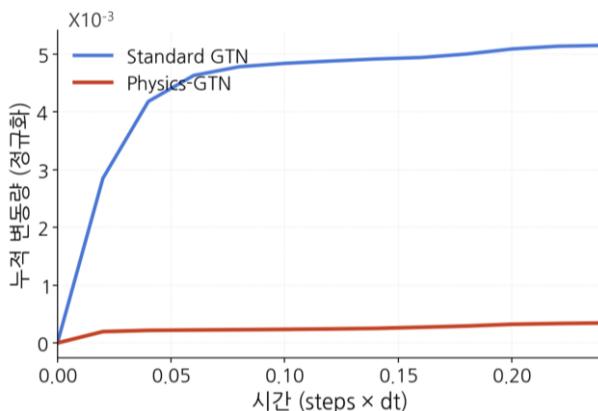


그림 2. 시간 흐름에 따른 제어 신호 누적 변동량

그림 1에서 Standard GTN은 전역 어텐션으로 인해 초반부터 넓은 진폭 분산을 보인다. 반면 Physics-GTN은 거리 기반 가중치가 원거리 영향을 감쇠시켜 안정적인 파형을 형성한다.

그림 2는 제어 신호의 누적 변동량을 비교한 결과이다. Standard GTN은 에피소드 전반에 걸쳐 변동량이 지속적으로 증가하는 반면, Physics-GTN은 초기 충전 구간 이후 변동량 증가가 거의 없어 목표 상태 도달 후에도 안정적으로 제어한다. 정량적으로 Physics-GTN의 평균 변동량은 0.0001로, Standard GTN 0.0017 대비 약 16 배 낮은 수치를 기록했다.

| 지표 | Standard GTN | Physics-GTN |
|-----------|-----------------|-----------------|
| F_{max} | 0.89 ± 0.03 | 0.93 ± 0.02 |
| F_{end} | 0.82 ± 0.05 | 0.92 ± 0.02 |
| 평균 변동량 | 0.0017 | 0.0001 |

표 1. 모델별 성능 비교

표 1은 8 큐비트 환경에서 3 회 시드 실험의 평균 성능을 정리한 결과이다. Physics-GTN은 최대 충전율 0.93, 최종 충전율 0.92를 달성하여 Standard GTN 대비 충전 및 유지 성능 모두 개선되었다. 특히 평균 변동량이 0.0001로 Standard GTN(0.0017) 대비 약 16 배 낮아, 거리 기반 어텐션 가중치가 불필요한 제어 신호 생성을 효과적으로 억제한다.

IV. 결론

양자 배터리 제어의 불안정성과 확장성 문제를 개선하기 위해 물리적 위상 정보를 반영한 Physics-GTN을 제안한다. 거리 기반 어텐션과 조화평균 보상 설계를 통해 안정적인 충전 제어를 달성했다. 8 큐비트 환경 실험에서 제안 모델은 Standard GTN 대비 제어 신호 변동량을 약 16 배 감소시켰으며, 최대 충전율 0.93과 유지 비율 0.99를 달성했다. 향후 대규모 환경과 실제 하드웨어 접속 조건에서 검증할 계획이다.

ACKNOWLEDGMENT

This work was partly supported by Korea Evaluation Institute of Industrial Technology(KEIT) grant funded by the Korea government(MOTIE) (No.RS-2025-04752989, Quantum battery core technology for ultra-fast charging 100x faster than traditional lithium-ion batteries)

참고 문헌

- [1] S. Yun et al. (2019). "Graph Transformer Networks," Advances in Neural Information Processing Systems (NeurIPS), 2019.
- [2] D. García-Martín et al., "Effects of noise on the overparametrization of quantum neural networks," Physical Review Research, 2024.
- [3] A. Katharopoulos et al., "Transformers are RNNs: Fast Autoregressive Transformers with Linear Attention," ICML, 2020.
- [4] M. Seo et al., "Rewards prediction-based credit assignment for reinforcement learning with sparse binary rewards," IEEE Access, vol. 7, pp. 118776-118791, 2019.