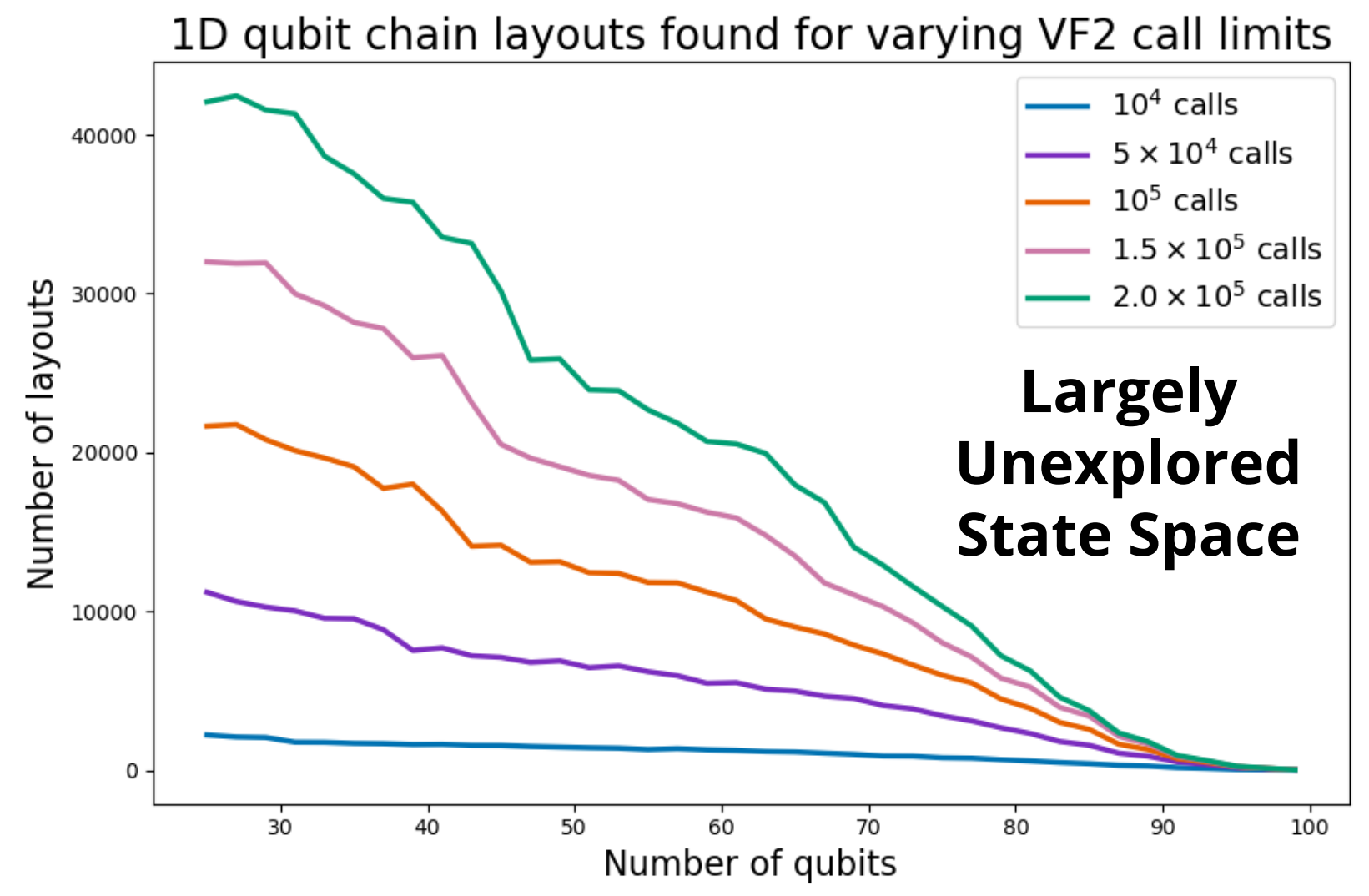
# Active-Noise-Aware Qubit Mapping via Monte Carlo Tree Search and Reinforcement Learning

By: Nabil Lathiff, Kameron Shahabi, and Essance Ray

## Motivation & Objective

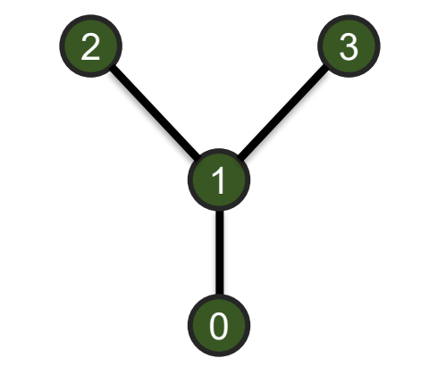
Qubit mapping is a critical step in quantum computing, where logical qubits in a circuit are assigned to physical qubits on hardware through optimization algorithms. Current approaches such as VF2++ rely on graph-matching heuristics that struggle to **scale** and often lack **noise-awareness**.

Our work addresses these limitations by introducing a stochastic, search-based framework that combines Monte Carlo Tree Search (MCTS) coupled with reinforcement learning (RL) to improve circuit performance. Success is then evaluated through circuit fidelity, which shows consistent improvement across benchmark circuits and at larger scales.



## Background- Qubit Mapping via VF2++ & SABRE

Input: Logical Pattern

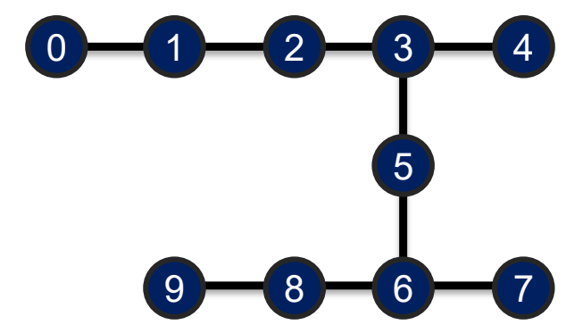


Nodes: Logical qubits
Edges: 2-qubit gates

VF2++/SABRE Process

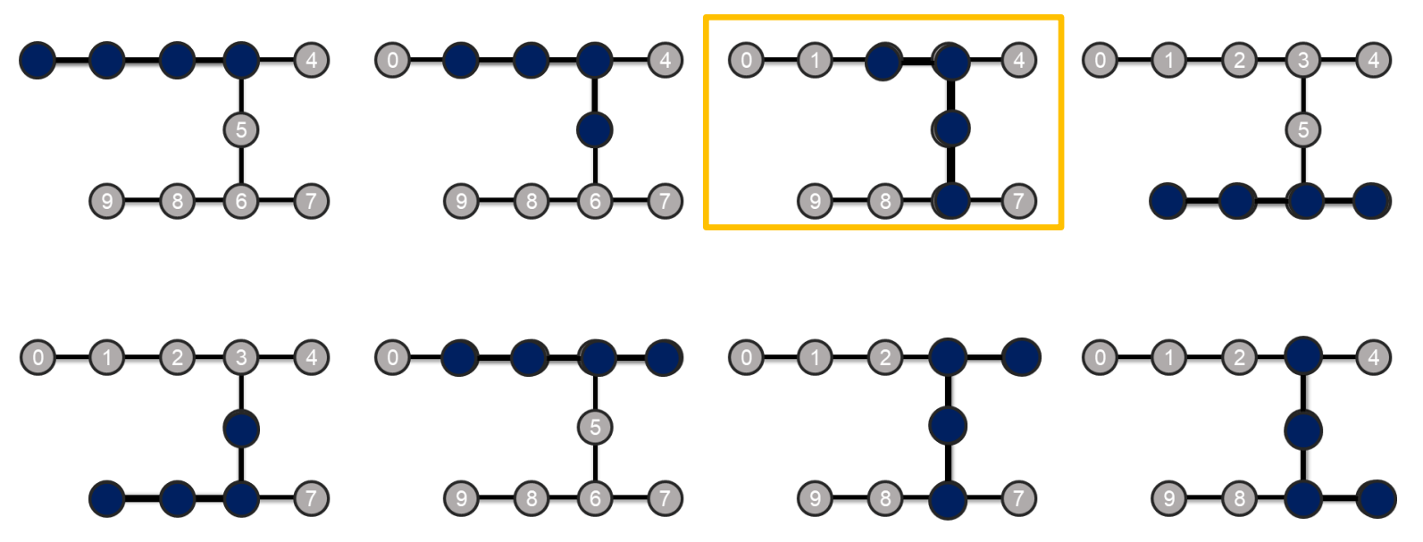
1. Find valid qubit placements using graph matching
2. Prefer placements with minimal swap overhead
3. Score each w/ hardware metrics (gate fidelity, T1, T2, etc.)
4. Select the highest-scoring placement
5. Dynamic swap-based routing after placement

Device Fragment

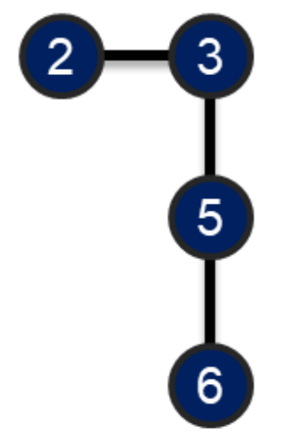


Nodes: Physical qubits
Edges: 2-qubit gates

Sample Random Embeddings

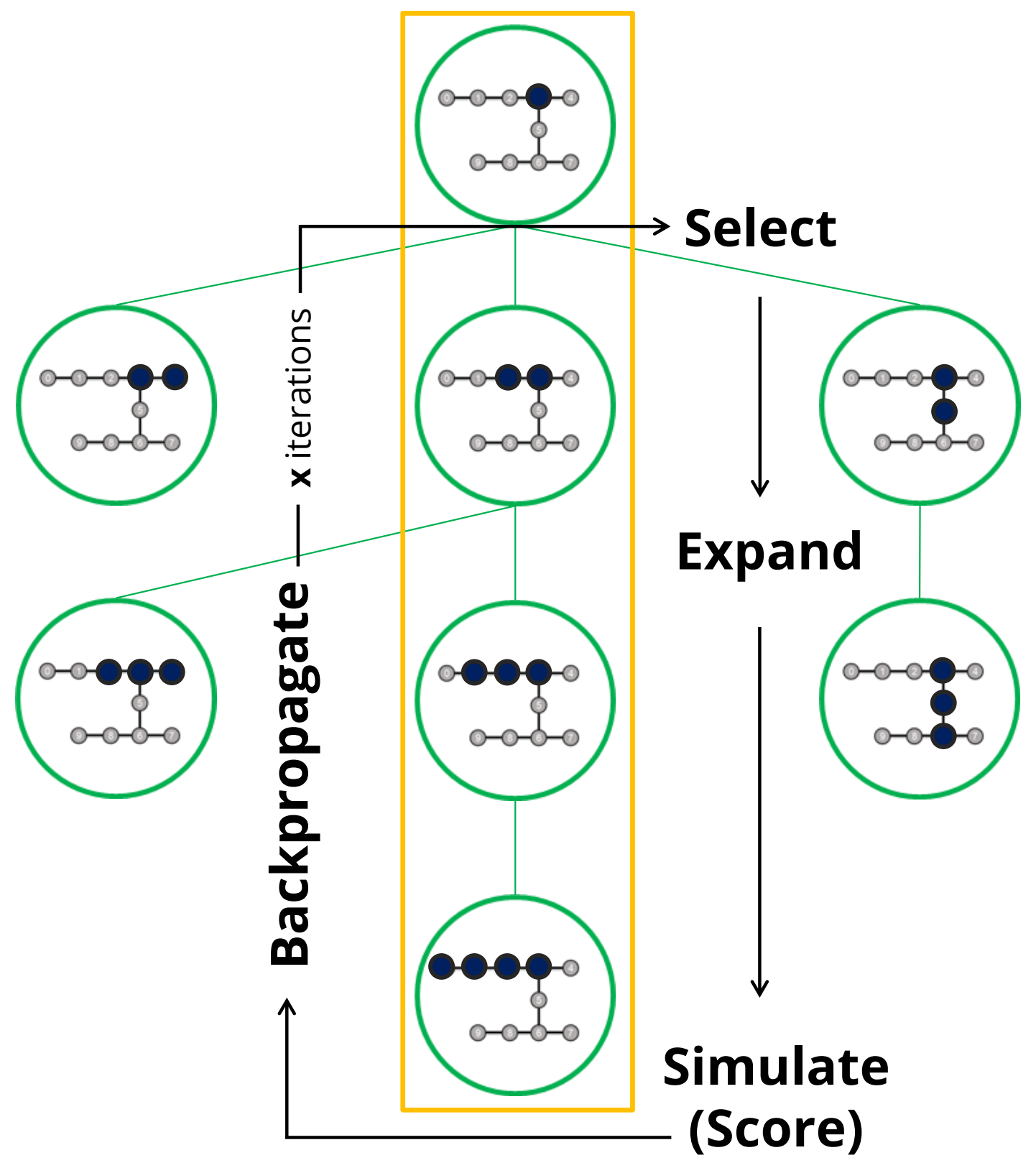


Optimal Layout



**Best Score**

## Background/Methods- Qubit Mapping via MCTS



**Reward Function:**

Single-qubit fidelity contribution:

$$F_{1q} = \prod (1 - e_{1q}[g(p)])^{n_{1q}[p]}$$

Two-qubit fidelity contribution:

$$F_{2q} = \prod (1 - e_{2q}[g(u), g(v)])^{n_{2q}[u,v]}$$

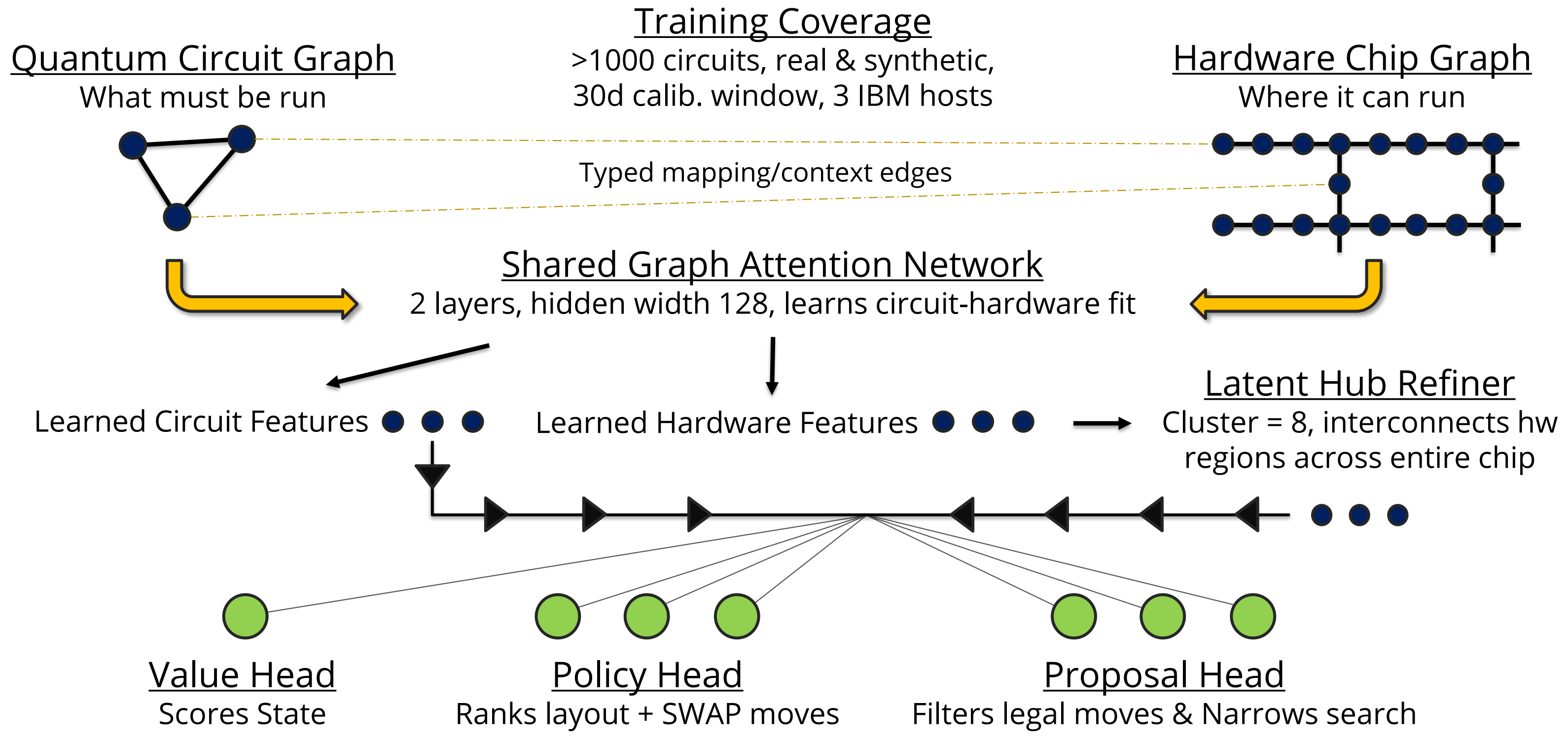
Final fidelity w/out routing:

$$F_{Static} = F_{1q} \cdot F_{2q}$$

Final fidelity w/ routing SWAPS:

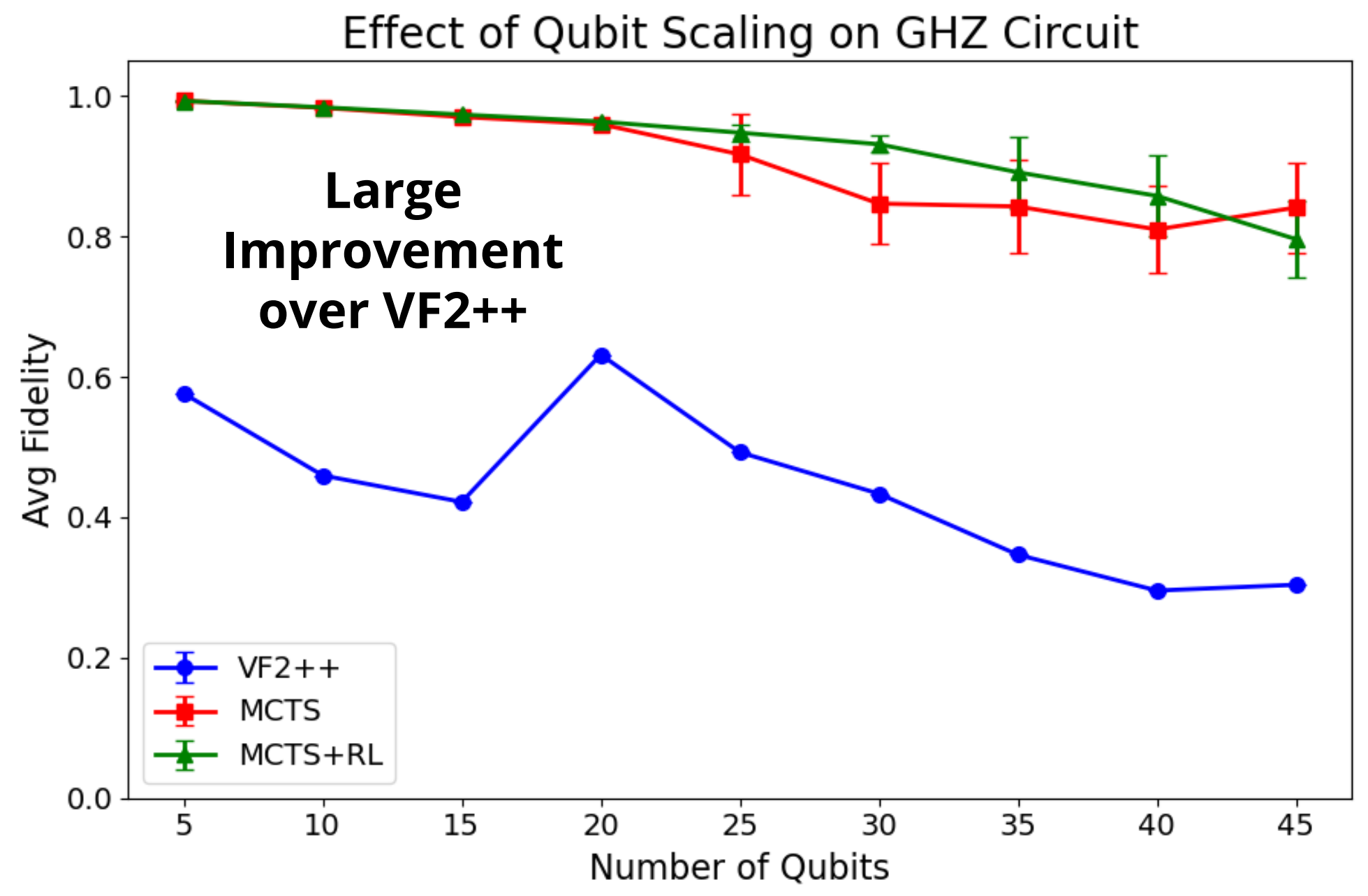
$$F_{SWAP} = \prod (1 - e_{2q}[a,b])^{3n_{swap}[a,b]}$$

## Methods- RL w/ AlphaZero Graph Neural Network

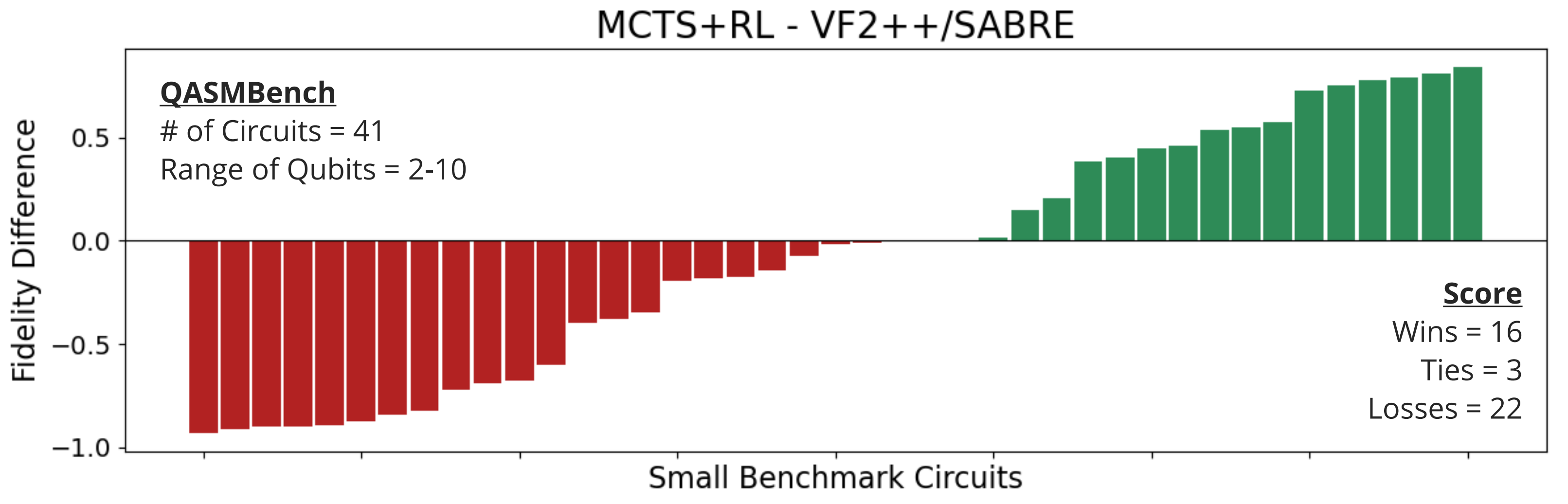


## Results- MCTS + RL

Initial results were obtained by replacing the VF2++ mapping algorithm with the MCTS-based approach while keeping the same SABRE routing procedure. The MCTS was optimized using scalable GHZ circuits that generate maximally entangled states across (N) qubits as well as the QASM benchmarking small circuits. After optimizing the MCTS framework, reinforcement learning (RL) was incorporated to further improve performance beyond randomized search strategies.

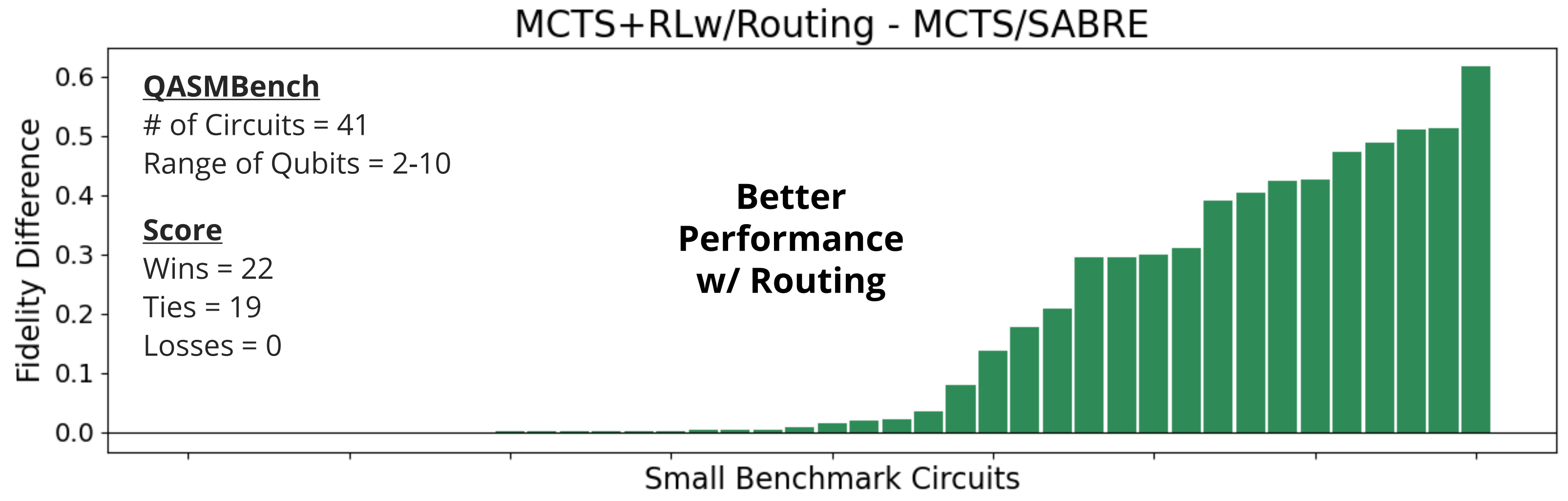


The MCTS-based approaches achieved substantially higher fidelities and scalability than VF2++, although VF2++ did perform better on more smaller benchmark circuits. Some runs from both VF2++ and MCTS produced zero-fidelity mappings, thus further debugging is underway to identify the underlying cause.

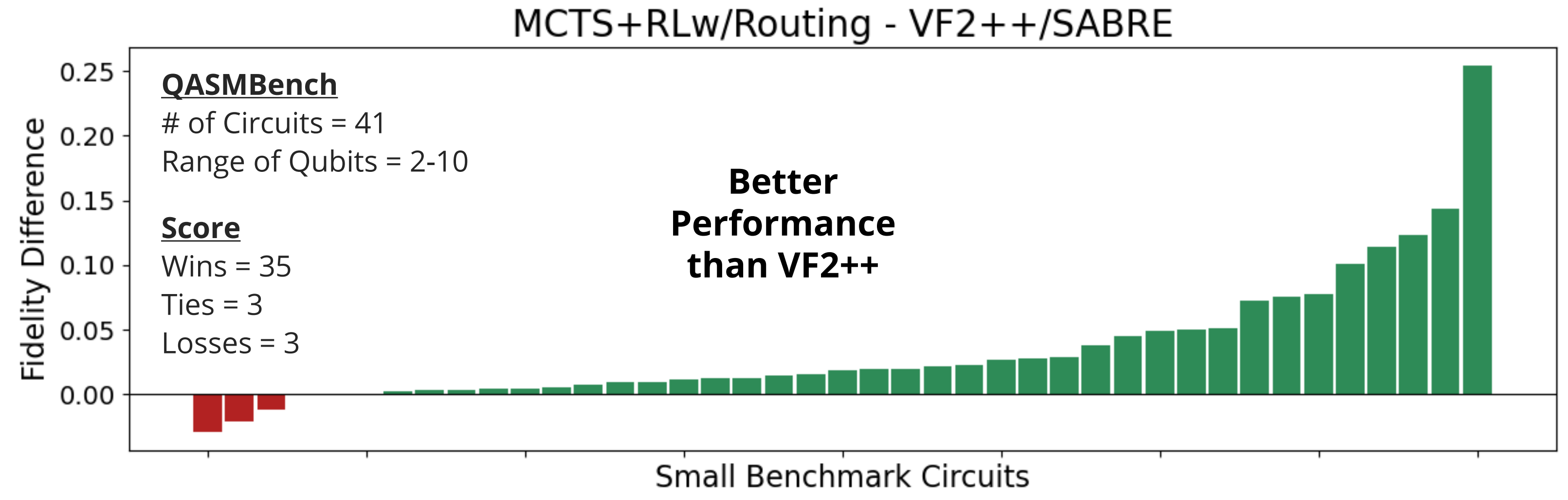


## Results- MCTS + RL w/ Routing

Another framework was also explored in which MCTS replaced both the VF2++ placement algorithm and the SABRE router. In this unified layout-and-routing approach, fidelity is evaluated after SWAP insertion using routed fidelity rather than a static mapping score. This allows the search process to optimize placement and routing simultaneously, aligning the objective more closely with the true hardware fidelity. The bar graph below illustrates the fidelity difference between the two approaches where positive (green) bars denote circuits where the routing-aware framework achieved higher fidelity.



Additionally, the unified approach shows superior performance across many of the small benchmarking circuits with just 3 losses and 3 ties compared to the standard VF2++/Sabre approach. Thus, the unified MCTS approach has proven to be our most effective solution.



## Current/Future Work

Our MCTS+RL routing framework shows promising improvements in circuit performance and scalability compared to the standard VF2++ and SABRE pipeline. Benchmarking has primarily focused on small circuits (≈2–10 qubits) as an initial optimization stage, and we are now extending evaluation to medium (≈10–30 qubits) and large (>30 qubits) circuits.

Training the neural network on larger circuits is computationally expensive, but the current results already demonstrate measurable improvements over the standard approach. Future work includes validating the pipeline on real quantum hardware using mirror circuits to obtain experimental fidelity measurements and further confirm performance gains. If hardware execution or large-scale training cannot be completed within the project timeline, simulation-based results will still provide strong evidence of the framework's effectiveness.

**We would like to thank all our faculty and industry mentors!**

ELECTRICAL & COMPUTER ENGINEERING
UNIVERSITY of WASHINGTON

**Faculty Supervisor: Sara Mouradian**
**Industry Advisers: Henry Zou & Nate Earnest-Noble**
**Sponsor: IBM - International Business Machines Corporation**