Scalable Quantum-Inspired Optimization Through Dynamic Qubit Compression

Co Tran[*1], **Quoc-Bao Tran**[*2], **Hy Truong Son**[3], **Thang N Dinh**[2†]

[1]University of Texas at Austin
[2]Virginia Commonwealth University
[3]University of Alabama at Birmingham
co.quoc.tran.2@gmail.com, tranq3@vcu.edu, thy@uab.edu, tndinh@vcu.edu

Abstract

Hard combinatorial optimization problems, often mapped to Ising models, promise potential solutions with quantum advantage but are constrained by limited qubit counts in near-term devices. We present an innovative quantum-inspired framework that dynamically compresses large Ising models to fit available quantum hardware of different sizes. Thus, we aim to bridge the gap between large-scale optimization and current hardware capabilities. Our method leverages a physics-inspired GNN architecture to capture complex interactions in Ising models and accurately predict alignments among neighboring spins (aka qubits) at ground states. By progressively merging such aligned spins, we can reduce the model size while preserving the underlying optimization structure. It also provides a natural trade-off between the solution quality and size reduction, meeting different hardware constraints of quantum computing devices. Extensive numerical studies on Ising instances of diverse topologies show that our method can reduce instance size at multiple levels with virtually no losses in solution quality on the latest D-wave quantum annealers.

Introduction

Combinatorial optimization problems are ubiquitous in various domains, including portfolio optimization (Mugel et al. 2021; Grozea et al. 2021), car manufacturing scheduling (Yarkoni et al. 2021), and RNA folding (Fox, Branson, and Walker 2021; Fox et al. 2022). These problems often involve finding the optimal solution among a vast number of possibilities, making them computationally challenging. Many of these problems can be mapped to Ising models (Lucas 2014), which encode the optimization objective in terms of interacting spins. However, a significant number of these problems fall into the NP-hard complexity class (Gary and Johnson 1979), meaning they are intractable for classical computers as the problem size grows. This intractability has motivated the exploration of alternative computing paradigms, such as quantum annealing (Kadowaki and Nishimori 1998a; Zhou and Zhang 2022), which leverages quantum mechanics to potentially solve these problems more efficiently.

Recent years have witnessed remarkable advances in quantum computing, bringing us closer to the realm of "quantum supremacy" (Preskill 2018), where quantum processors solve problems intractable for classical computers. Notable milestones include Google's Sycamore and Willow processors demonstrating supremacy in a sampling task (Arute et al. 2019; Acharya et al. 2024). Beyond these proof-of-concept demonstrations, there have been efforts to showcase quantum utility on more practical problems. Quantum annealing, implemented in D-WavDe's systems (Boothby et al. 2020), has demonstrated quantum advantages for certain types of problems (King et al. 2021; Tasseff et al. 2022, 2024; King et al. 2024). Gate-based algorithms such as the Quantum Approximate Optimization Algorithm (QAOA) (Bauza and Lidar 2024) and Variational Quantum Eigensolver (VQE) (Peruzzo et al. 2014) offer alternative routes for tackling optimization challenges. Additionally, quantum-inspired specialized hardware, including optical Ising machines (Honjo et al. 2021), digital annealers, and FPGA-based solvers (Patel et al. 2020), provide complementary approaches to address complex optimization problems.

Despite the rapid progress in quantum computing, qubit count remains a significant limiting factor for solving practical optimization problems. Current state-of-the-art quantum annealers, such as D-Wave's Advantage platform, offer over 5000 qubits (Boothby et al. 2020). However, many real-world applications require even more qubits. For example, performing MIMO channel decoding with a 60Tx60R setup on a 64-QAM configuration would necessitate about 11,000 physical qubits (Tabi et al. 2021), exceeding the capabilities of existing hardware. The challenge is further compounded by limited qubit connectivity, which necessitates complex minor embedding techniques (Choi 2008, 2011), significantly increasing the number of physical qubits required. These hardware constraints substantially limit the size and complexity of problems that can be directly solved on quantum processors with a clear advantage over classical methods.

While waiting for quantum hardware advances, a parallel challenge emerges: efficiently reducing Ising models to fit limited qubit capacities. Current reduction techniques, ranging from classical roof duality (Hammer, Hansen, and Simeone 1984; Boros and Hammer 2002) and extended roof duality (Rother et al. 2007) to recent graph-based approaches (Thai et al. 2022), show promise but face significant limitations. These methods, constrained by the need for an exact reduction, can compress only a fraction of problem in-

[*]These authors contributed equally.
[†]Corresponding author: tndinh@vcu.edu

stances—with the best-performing algorithms reducing less than half of tested cases, achieving an average compression of about 20% among instances (Thai et al. 2022). This variability in effectiveness highlights a notable gap: the absence of a universal, tunable reduction method to compress Ising models to arbitrary sizes. Such flexibility could accommodate diverse quantum hardware constraints and potentially enhance the applicability of quantum annealing across a broader spectrum of real-world optimization problems.

We present **G**raph **N**eural **I**sing **T**ransformer for **E**fficient Quantum Optimization (GRANITE), a novel framework that leverages Graph Neural Networks (GNNs) to dynamically compress large Ising models for quantum annealing. Our approach automates the discovery of combinatorial rules for qubit reduction by training a GNN to predict ground-state qubit alignments and identify optimal contraction candidates. This data-driven method enables progressive spin merging while preserving solution integrity, offering a tunable trade-off between compression ratio and solution quality. Unlike previous compression techniques that relied on manual heuristics (Thai et al. 2022), our GNN-based approach captures complex patterns in both local and global spin interactions, identifying compressible qubit groups that elude detection by conventional methods. In contrast to GNN-based methods to directly solve Ising models and combinatorial optimization problems (Dai et al. 2017; Li, Chen, and Koltun 2018; Gasse et al. 2019; Joshi, Laurent, and Bresson 2019; Schuetz, Brubaker, and Katzgraber 2021; Schuetz et al. 2022), these methods face challenges, including sensitivity to graph structure and connectivity, and poor performance on sparse graphs (Pan et al. 2021). Moreover, by acting as a preprocessing phase for quantum computing, our approaches preserve the potential for a quantum advantage.

Our contributions. We summarize below our contributions

- We present GRANITE, a GNN-based framework that dynamically compresses large Ising models, automating the discovery of qubit reduction rules. This method efficiently predicts ground state alignments and identifies optimal contractions, offering tunable trade-offs between model size and solution quality, and accommodating diverse quantum hardware constraints. The compression preserves the Ising structure and can work with any quantum technology that solves Ising models, including both quantum annealers and gate-based quantum computers via Quantum Approximate Optimization Algorithm (QAOA).
- We demonstrate GRANITE's effectiveness in providing substantial size reductions across various Ising topologies while preserving solution accuracy.
- By significantly reducing qubit requirements, our approach expands the scope of tractable problems for current quantum annealers, potentially accelerating practical quantum advantage in optimization tasks. This work provides a powerful tool for exploring the quantum-classical computational boundary, addressing a critical challenge in near-term quantum computing.

This work addresses the qubit limitation challenge, offering a powerful, flexible tool for researchers and practitioners in quantum optimization.

Related Work

Ising Models and Combinatorial Optimization. Ising models, which naturally lend themselves to graph representations, have been a focal point in statistical physics and combinatorial optimization (Carleo et al. 2019; Tanaka, Tomiya, and Hashimoto 2023). Solving Ising models is challenging due to their NP-hardness and has applications across various domains, including computer science and machine learning.

Traditional Approaches to Solving Ising Models. Solving computationally difficult Ising models has traditionally relied on heuristic algorithms and physics-inspired techniques. Simulated Annealing (SA) (Kirkpatrick, Gelatt Jr, and Vecchi 1983) has been a cornerstone approach. More recently, Ising machines based on algorithms such as SimCIM (simulated coherent Ising machine) (Tiunov, Ulanov, and Lvovsky 2019; King et al. 2018) and simulated bifurcation (SB) (Goto et al. 2021; Oshiyama and Ohzeki 2022) have shown impressive results in finding ground states of Ising models.

Machine Learning Approaches for Ising Models. Machine learning techniques have been increasingly applied to solve Ising problems. Variational Autoregressive Networks (VANs) (Wu, Wang, and Zhang 2019) and Variational Classical Annealing (VCA) (Hibat-Allah et al. 2021, 2020) have shown promise. VCA, in particular, outperforms traditional SA but faces scalability issues, being limited to problems with up to 32 spin variables in challenging scenarios like the Wishart Planted Ensemble (WPE) (Hamze et al. 2020). Reinforcement learning approaches (Angelini and Ricci-Tersenghi 2023; Panchenko 2013) offer an alternative by directly optimizing for ground state configurations.

Graph Neural Networks for Ising Models. Graph Neural Networks (GNNs) have emerged as a promising approach for solving Ising models and related combinatorial optimization problems (Dai et al. 2017; Li, Chen, and Koltun 2018; Gasse et al. 2019; Joshi, Laurent, and Bresson 2019; Schuetz, Brubaker, and Katzgraber 2021; Schuetz et al. 2022). Some GNN-based methods have demonstrated the ability to handle large-scale instances with millions of variables (Schuetz, Brubaker, and Katzgraber 2021). However, these approaches face challenges, including sensitivity to graph structure and connectivity, and poor performance on sparse graphs (Pan et al. 2021). Recent works have questioned the effectiveness of GNNs compared to classical heuristic algorithms for certain problems (Boettcher 2022; Angelini and Ricci-Tersenghi 2022). Our proposed approach, unlike existing GNN-based methods that attempt to solve Ising models directly, focuses on using GNNs to compress Ising models. This novel perspective aims to address the scalability issues faced by current methods while maintaining the ability to capture complex interactions in the Ising system.

Background

Ising Models and NP-hard Problems

Ising models, originally developed in statistical physics, have become a powerful framework for representing combinatorial optimization problems (Lucas 2014). An Ising model consists

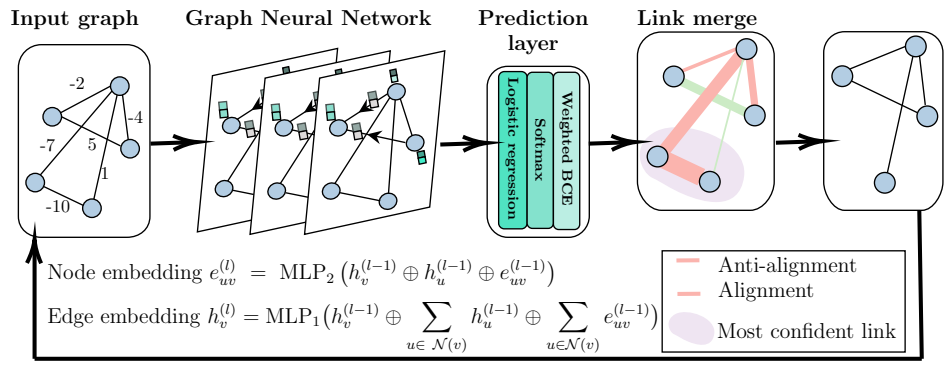

Figure 1: **GRANITE: Graph Neural Ising Transformer for Efficient Quantum Optimization.** The model comprises three key components: **a,** A GNN that learns edge (e_{uv}) and node (h_v) representations, capturing the Ising model's structure and interactions. **b,** A prediction layer using logistic regression with softmax to calculate weighted binary cross-entropy, assigning confidence scores to potential actions. **c,** A link contraction process that executes the highest-confidence merge or flip-merge operation. During inference, the contracted graph is iteratively fed back into GRANITE until the desired reduction ratio is achieved, enabling the transformation of large-scale Ising problems into quantum-compatible formats.

of binary variables (spins) $s_i \in -1, +1$, with energy given by the Hamiltonian:

$$H = -\sum_{i,j} J_{ij} s_i s_j - \sum_i h_i s_i \tag{1}$$

where J_{ij} represents the coupling strength between spins i and j, and h_i is the external field acting on spin i. Many NP-hard problems can be reformulated as Quadratic Unconstrained Binary Optimization (QUBO) problems (Glover, Kochenberger, and Du 2018), which are closely related to Ising models. The QUBO formulation uses binary variables $x_i \in 0, 1$:

$$\min_x \sum_{i,j} Q_{ij} x_i x_j + \sum_i c_i x_i \tag{2}$$

QUBO can be mapped to Ising models through the transformation: $s_i = 2x_i - 1$. NP-hard problems are characterized by exponential time complexity for exact classical algorithms (Gary and Johnson 1979). This intractability for large instances has motivated the exploration of quantum annealing as an alternative computing paradigm (Kadowaki and Nishimori 1998a). By mapping NP-hard problems to Ising models, researchers aim to leverage quantum effects to explore complex solution spaces more efficiently, potentially overcoming limitations of classical computing such as exponential state space growth and local optima traps.

Quantum Annealing

Quantum annealing (QA) is a metaheuristic for solving optimization problems, leveraging quantum mechanical effects such as tunneling and superposition (Kadowaki and Nishimori 1998b; Farhi et al. 2000; Santoro and Tosatti 2006). It is particularly well-suited for solving problems formulated as Ising models or Quadratic Unconstrained Binary Optimization (QUBO), as discussed in the previous section.

The quantum annealing process is based on adiabatic quantum computation. The system is initialized in the ground state of an easily prepared Hamiltonian, typically the transverse field Hamiltonian, and then evolves according to:

$$H_{\text{system}}(s) = -\frac{A(s)}{2}\left(\sum_i^n \sigma_i^x\right) + \frac{B(s)}{2}\left(H_{\text{problem}}\right), \tag{3}$$

where $s \in [0, 1]$ is the anneal fraction, σ_i^x is the Pauli x-matrix for the i-th qubit, and H_{problem} is the problem Hamiltonian, equivalent to the Ising model Hamiltonian in Eq. (1). $A(s)$ and $B(s)$ define the anneal schedule (Hauke et al. 2020).

Quantum annealing harnesses quantum effects like superposition to tackle optimization problems more efficiently than classical methods. By exploring multiple states simultaneously and "tunneling" through energy barriers, it can potentially find better solutions that classical algorithms might miss. Recent studies have demonstrated a quantum advantage for certain problem classes (Harris et al. 2018; King et al. 2022, 2024), highlighting the potential of QA as a powerful tool for solving Ising model-based optimization problems.

Despite the promise of quantum annealing, current systems face several challenges including the key challenge of limited qubit counts. The latest quantum annealer from D-Wave features fewer than 6000 flux qubits arranged, in a Pegasus topology (Boothby et al. 2020). As quantum annealing technology continues to advance, it promises to tackle increasingly complex optimization problems that are intractable for classical computers.

Method

Qubit Alignment at Ground States

Ising models, fundamental in statistical physics and optimization, can be elegantly represented as graphs. This representation not only captures the model's structure but also enables powerful reduction techniques.

Graph representation. Consider an Ising Hamiltonian (h, J) over spins s_1 to s_n. We construct an undirected

weighted graph $G_H = (V, E)$ where $V = \{0, 1, \ldots, n\}$ and $E = \{(i, j) \mid J_{ij} \neq 0\} \cup \{(0, i) \mid h_i \neq 0\}$ (Fig. 1a). The auxiliary vertex 0 represents linear biases, unifying the treatment of linear and quadratic interactions. Edge weights encapsulate both interaction types:

$$w(i,j) = \begin{cases} J_{ij} & \text{if } i \neq j \\ h_j & \text{if } i = 0 \end{cases} \tag{4}$$

Our choice of representation simplifies the analysis by treating all interactions uniformly as edge weights.

Qubits alignment. Ground states, configurations with minimum energy, reveal crucial structural information. We classify each edge (i, j) based on the behavior of connected spins across all ground states:

- **Alignment**: s_i and s_j always have the same value.
- **Anti-alignment**: s_i and s_j always have different values.
- **Neutral**: s_i and s_j alignment varies among ground states.

This classification enables targeted graph reductions that preserve ground-state properties. For an aligned edge (i, j), we have $s_i = s_j$ in all ground-states, thus, we can replace $s_i = s_j$ and remove s_i from the Ising. Equivalently, we perform a merge operation on the graph. This operation combines two nodes that always have the same value in ground states, effectively reducing the number of variables in our system. The merge operation removes one node and redirects its connections to the remaining node. Formally, we define the merge operation $M(i, j)$ as:

$$M(i,j): \begin{cases} V' = V \setminus \{j\}, \\ E' = E \setminus \{\, (j,k) \mid (j,k) \in E \,\}, \\ w'(i,k) = w(i,k) + w(j,k),\ \forall\, (i,k) \in E', \\ w'(u,v) = w(u,v) \text{ otherwise.} \end{cases}$$

An anti-aligned edge (i, j) means $s_i = -s_j$ in all ground-states, thus, we can replace s_i with $-s_j$ and remove s_i, completely. In the graph, the two nodes i and j undergo a two-step flip-merge operation. This operation first flips the sign of all interactions involving one node (to account for the constant difference in spin values), then merges the nodes. The flip-merge operation $FM(i, j)$ consists of:

1. *Flip*: Negate the weights of all edges incident to j, i.e., $w'(j,k) = -w(j,k), \quad \forall\, (k,j) \in E$.
2. *Merge*: Apply $M(i, j)$ as defined above.

Finally, neutral edges allow either merge or flip-merge, offering flexibility in reduction strategy.

These operations can significantly simplify Ising Hamiltonians while maintaining their essential properties. Iterative application potentially reduces problem complexity, guiding the development of efficient solution methods or revealing underlying system structure.

Hardness of Predicting Qubit Alignment. Predicting the alignment of qubits is, however, intractable.

Theorem 1. *The problem of classifying a single edge in an Ising model as alignment or non-alignment is Co-NP-hard. Consequently, there is no polynomial-time algorithm for this problem unless P = NP.*

The proof is done by a polynomial-time reduction from a Co-NP-hard problem of determining whether all truth assignments of a 3-SAT formula satisfy $x_i = x_j$ (or $x_i \neq x_j$) for some pair of variables x_i and x_j. A complete proof with detailed construction of the Hamiltonian and analysis is provided in our extended version of the paper.

The co-NP-completeness of edge classification in Ising models underscores its computational intractability, with exact solutions requiring exhaustive examination of all spin configurations to identify the complete set of ground states. While this approach remains viable for small instances, it becomes infeasible as system size grows, limiting its applicability in practical scenarios. This computational barrier motivates the exploration of alternative strategies, particularly in the realm of machine learning.

Graph Neural Ising Transformer for Efficient Quantum Optimization (GRANITE)

Graph Neural Networks (GNNs) emerge as a promising candidate, offering a unique ability to capture the intricate spatial relationships and interactions inherent in Ising models. By leveraging the graph structure of the Ising Hamiltonian, GNNs can potentially learn to approximate edge classifications without explicit enumeration of ground states, opening avenues for scalable analysis of larger systems. This approach not only promises computational efficiency but also the potential to uncover hidden patterns and heuristics in edge behavior across diverse Ising instances, potentially leading to new insights into the structure of complex spin systems.

The **G**raph **N**eural **I**sing **T**ransformer for **E**fficient Quantum Optimization (GRANITE) leverages Graph Neural Networks (GNNs) to navigate the complex landscape of Ising model reduction (Fig. 1). GRANITE iteratively predicts optimal graph contraction operations—merge or flip-merge—for each edge in the Ising model's graph representation. In each iteration, the GNN processes the current graph structure, learning edge and node representations that capture local and global spin interactions. These representations feed into a prediction layer, which assigns confidence scores to potential merge and flip-merge operations for each edge. The edge with the highest confidence score is selected, and its associated operation is performed, reducing the graph by one node. This process repeats until the Ising model is sufficiently small to be handled by quantum hardware, effectively bridging the gap between large-scale classical problems and limited-size quantum processors. The GRANITE workflow can be summarized in Algorithm 1.

The key advantage of leveraging GNNs in this process is their ability to learn complex graph structures and capture both local and global information, enabling accurate identification of non-separable qubit groups. By exploiting the representational power of GNNs, our approach aims to improve upon the greedy merging strategies employed in previous methods to lead to more effective Qubits reduction and facilitate the solution of larger optimization problems on quantum annealing hardware.

GNN Model. We start with the representation of Ising model with Hamiltonian graph $G_H = (V, E)$ where V is

Algorithm 1: GRANITE - Graph Neural Ising Transformer

Require:
Ising model graph $G = (V, E)$
Desired reduction ratio α
Number of GNN layers L
Ensure: Reduced Ising model graph G'
Initialize the GNN model with specified hyperparameters
while size$(G) > \alpha \times$ (initial size of G) **do**
  for $\ell = 1$ to L **do**
    Compute node representation $h_v^{(\ell)}$, $\forall v \in V$.
    Compute edge representation $e_{uv}^{(\ell)}$, $\forall (u, v) \in E$.
  end for
  for each edge $(u, v) \in E$ **do**
    $z_{uv} = h_u^{(L)} \oplus h_v^{(L)} \oplus e_{uv}^{(L)}$
    $\hat{y}_{uv} = \sigma(\langle w, z_{uv} \rangle)$
    $C_{uv} = -(\hat{y}_{uv} \log(\hat{y}_{uv}) + (1 - \hat{y}_{uv}) \log(1 - \hat{y}_{uv}))$
  end for
  $(\hat{u}, \hat{v}) = \arg\max_{(u,v) \in E} C_{uv}$
  if $\hat{y}_{\hat{u}\hat{v}} < 0.5$ **then** ▷ Regular merge
    $G = \text{Merge}(G, \hat{u}, \hat{v})$
  else ▷ Flip then merge
    $G = \text{Flip}(G, \hat{u})$
    $G = \text{Merge}(G, \hat{u}, \hat{v})$
  end if
end while
return G

the set of nodes and $E \subseteq V \times V$ is the set of edges. We construct the initial node and edge features are defined as:

$$\mathbf{H}^{(0)} = \{h_v^{(0)} \in \mathbb{R}^{d_h} \mid v \in V\},$$

$$\mathbf{E}^{(0)} = \{e_{uv}^{(0)} \in \mathbb{R}^{d_e} \mid (u, v) \in E\},$$

respectively, in which d_h and d_e are the corresponding numbers of input node and edge features. For each node v, $h_v^{(0)}$ is initialized with the degree, the weighted degree, and the absolute weighted degree. The edge features $e_{uv}^{(0)}$ for edge (u, v) contain the edge weights and the absolute edge weights.

The GNN model is designed to learn the representations of both nodes and edges simultaneously across multiple layers via the message-passing scheme. This model takes into account not only the features of nodes and edges but also their interactions, ensuring that both nodes and edges evolve over time as the network processes information. Let $\mathbf{H}^{(\ell)}$ denote the set of node representations $h_v^{(\ell)}$ and $\mathbf{E}^{(\ell)}$ denote the set of edge representations $e_{uv}^{(\ell)}$ at layer ℓ. At layer ℓ, the message passing scheme updates each node representation based on the neighboring nodes' representations at the previous layer $\ell - 1$; meanwhile, each edge representation is updated based on the edge's two corresponding nodes. Formally, we have:

$$h_v^{(\ell)} = \text{MLP}_1 \left(h_v^{(\ell-1)} \oplus \sum_{u \in \mathcal{N}(v)} h_u^{(\ell-1)} \oplus \sum_{u \in \mathcal{N}(v)} e_{v,u}^{(\ell-1)} \right),$$

$$e_{uv}^{(\ell)} = \text{MLP}_2 \left(h_u^{(\ell-1)} \oplus h_v^{(\ell-1)} \oplus e_{uv}^{(\ell-1)} \right),$$

where $\text{MLP}(\cdot)$ denotes a Multilayer Perceptron[1], $\mathcal{N}(v)$ denotes the set of neighboring nodes of v, and $\oplus$ denotes the operation of vector concatenation. For effective computation of combinatorial properties in the underlying graph, our MLP uses simple ReLU activation functions after each hidden linear transformation layer and a simple linear (identity) activation for the output layer.

An important step in our reduction framework is to predict for if each edge (u, v) is aligned or anti-alignment, i.e., whether it should be merged or flip-merged, respectively. We leverage a logistic regression model for this task. Let L denote the number of layers of message passing, we concatenate the node and edge representations of the last layer into a feature vector z_{uv}

$$z_{uv} = h_u^{(L)} \oplus h_v^{(L)} \oplus e_{uv}^{(L)}.$$

The logistic regression model predicts the probability on each edge (i.e. edge confidence) as:

$$\hat{y}_{uv} = \sigma(\langle w, z_{uv} \rangle), \tag{5}$$

where $\sigma(\cdot)$ denotes the sigmoid function, w is a learnable vector with equal length as the concatenated edge vector z_{uv}, and $\langle \cdot, \cdot \rangle$ denotes the inner product.

Hybrid Loss Function for Iterative Contraction

To simplify the Ising Hamiltonian, our method iteratively contracts one edge per iteration. The process involves combining a binary cross-entropy (BCE) loss for edge classification (alignment or anti-alignment) with confidence-based weighting to prioritize the most certain predictions.

Consider a graph with N edges, represented by the set $\mathbf{L} = \{l_1, l_2, \ldots, l_N\}$, where each edge l_i has a logit $\hat{y}_i$ corresponding to the ith edge (u, v) and a ground-truth label $y_i \in \{0, 1\}$ indicating alignment or anti-alignment, respectively. Neutral links are excluded from the loss computation since their processing (merge or flip-merge) preserves Ising optimality.

The confidence-based weight for each edge l_i is computed directly from $\hat{y}_i$ as $c_i = \frac{\exp(|\hat{y}_i - 0.5|/T)}{\sum_{j=1}^{N} \exp(|\hat{y}_j - 0.5|/T)}$, where T is a temperature hyper-parameter. The absolute difference $|\hat{y}_i - 0.5|$ measures the confidence of the prediction, with larger values indicating greater certainty. The softmax normalization ensures a smooth, differentiable prioritization of edges for contraction.

The weighting mechanism introduces a trade-off parameter, $\lambda \in [0, 1]$, to combine the confidence-based weights c_i with uniform weights [2]. The final weight for each edge is: $w_i = \lambda \cdot c_i + (1 - \lambda)$. The hybrid loss function is defined as:

$$\mathcal{L} = \sum_{i=1}^{N} w_i \cdot \left(-y_i \log(\hat{y}_i) - (1 - y_i) \log(1 - \hat{y}_i) \right).$$

This interpolation enables different loss formulations:

[1]The two MLPs encoding for nodes and edges do not share parameters and are denoted as MLP_1 and MLP_2, respectively.

[2]An alternative formulation for c_i can be defined as $c_i = \frac{(\hat{y}_i - 0.5)^p}{\sum_{j=1}^{N} ((\hat{y}_j - 0.5)^p}$, where even integer $p \geq 2$ is a parameter controlling the sensitivity of the confidence scores.

- $\lambda = 0$: The loss reduces to standard binary cross-entropy (BCE), treating all edges equally.
- $\lambda = 1$: The loss relies entirely on the confidence-based softmax weights, emphasizing high-confidence edges.

Experiments

We conducted experiments to evaluate GRANITE's effectiveness in compressing Ising models while maintaining solution quality. Our evaluation focused on the trade-off between compression levels and solution accuracy across various graph topologies and sizes. The experiments were performed on D-Wave's Advantage Quantum Processing Unit (QPU), featuring the advanced Pegasus topology (P_{16}) with 5,640 qubits.

Experimental Setup

Dataset. We generate random Ising Hamiltonians represented as graphs with the spins as nodes and edges following three graph topologies Erdős-Rényi (ER), Barabási-Albert (BA), and Watts-Strogatz (WS) models. Nodes have zero linear biases, i.e., $h_i = 0$ for all i and the edge weight J_{ij} sampled uniformly randomly in the range $(-5, 5)$.

The dataset includes 97,500 graphs (325 distinct configurations × 100 instances × 3 topologies), split into 80% training and 20% validation sets.For each type of topology, the number of nodes ranges from 2 to 26 (25 distinct sizes), while the average node degree varies from 1 to $n - 1$ (sum up to 325 instances). For each combination of node count and average degree, we generate 100 unique graphs, resulting in a total dataset of 97,500 graphs (325 $\times$ 100 instances $\times$ 3 types of topologies). We split training at a ratio of 80%/20% for training/validation respectively. The evaluation set contains 20,000 graphs to assess the effectiveness of unseen graphs. The ground state labels (alignment, anti-alignment, or neutral) for each edge were computed using an exhaustive search.

Hyperparameters. We employ the Adam optimizer with a learning rate of 0.001. The maximum number of iterations is set to 300, with the best model saved based on the lowest loss performance on the validation set. The size of the MLP layers is relatively small, depending on the size of $\mathbf{H}^{(0)}$ and $\mathbf{E}^{(0)}$, which in this case yields 2 and 2, respectively. By default, the GRANITE consists of three GNN layers and uses a hybrid loss function with $\lambda = 0.5$.

Environment. We conducted experiments on the D-Wave Advantage 4.1 Quantum Processing Unit (QPU), utilizing the advanced Pegasus topology with 5,640 qubits and an annealing time of 40 ns. The minor-embeddings of the Ising Hamiltonians are found with minorminer (Cai, Macready, and Roy 2014). Optimal solutions for each Ising instance is found using the Gurobi Optimizer (Gurobi Optimization, LLC 2024).

Evaluation metrics We report two metrics that measure the solution quality and the qubit reduction levels. The solution quality is measured using optimality, the ratio between the best energy found using D-wave quantum annealer, denoted by E_{best}, and the minimum energy found using Gurobi, denoted by E_{min}. Formally, the optimality is computed as

$$\text{Optimality}(\%) = 1 - \frac{|E_{\text{best}} - E_{\text{min}}|}{|E_{\text{min}}|}, \tag{6}$$

to accommodate for the case when $E_{\text{best}} > 0$ (and $E_{\text{min}} < 0$). When optimal solutions are found, the optimality will be one.

The qubit reduction is measured as

$$\text{reduction} = 1 - \frac{q_{\text{compressed}}}{q_{\text{original}}},$$

where $q_{\text{compressed}}$ represents the number of physical qubits after compression and q_{original} represents the original number of physical qubits before compression.

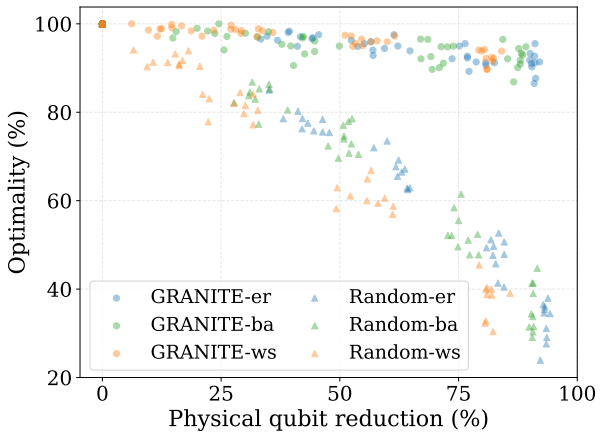

Figure 2: GRANITE vs. random, the random merge and flip-merge of edge for n = 200 across three different topologies.

Experiment Results

Solution quality. Our experimental results demonstrate GRANITE's effectiveness in compressing Ising models while maintaining high solution quality across different graph types and sizes. Table 1 presents optimality for different compression ratios (12.5%, 25%, 50%, 75%) across graph of sizes $n = \{25, 50, 100, 200, 400\}$. The solution quality remains high across all graph types, even at aggressive compression levels. For instance, with 75% edge reduction on graphs with $n = 200$, we maintain optimality above 91% for all three topologies. We observe a slight degradation in solution quality as graph size increases.

Qubit Reduction. As shown in Figure 2, GRANITE demonstrates significant qubit reduction at a slight cost of solution quality. For the largest Ising instances ($n = 400$), the Ising Hamiltonian exceeds the hardware constraint and cannot be solved on the D-Wave quantum annealer for Erdős-Rényi model while taking 3621 and 1545 qubits on average for Barabási-Albert, and Watts-Strogatz instances. At 75% edge reduction, the remaining physical qubit ratios reach as low as 5.3% (Erdős-Rényi), 7.4% (Barabási-Albert), and 16.5% (Watts-Strogatz) of the original number of qubits.

Topology	Reduction (%)	n				
		25	50	100	200	400
Erdős-Rényi	0.0% Original Ising	100.00 ± 0.00	100.00 ± 0.00	99.68 ± 0.19	96.77 ± 0.45	NaN
	12.5%	98.21 ± 1.09	98.67 ± 0.44	96.70 ± 0.55	94.87 ± 0.37	89.36 ± 0.37
	25.0%	96.03 ± 1.44	97.12 ± 0.76	94.26 ± 0.56	93.38 ± 0.54	88.58 ± 0.80
	50.0%	94.58 ± 1.64	94.19 ± 1.42	91.23 ± 1.00	91.56 ± 0.63	88.02 ± 0.63
	75.0%	93.20 ± 1.55	92.26 ± 1.40	89.60 ± 0.97	91.07 ± 0.82	88.17 ± 0.79
Barabási-Albert	0.0% Original Ising	100.00 ± 0.00	100.00 ± 0.00	99.93 ± 0.07	97.25 ± 0.25	89.11 ± 0.48
	12.5%	97.69 ± 1.22	97.54 ± 0.91	97.79 ± 0.65	94.71 ± 0.80	88.53 ± 1.11
	25.0%	97.12 ± 1.38	95.63 ± 1.24	97.09 ± 0.75	92.69 ± 1.08	87.84 ± 1.11
	50.0%	94.88 ± 1.59	93.00 ± 1.45	94.87 ± 1.14	92.66 ± 0.92	88.76 ± 0.60
	75.0%	92.43 ± 1.88	91.89 ± 1.21	93.37 ± 1.28	92.30 ± 0.87	88.72 ± 0.96
Watts-Strogatz	0.0% Original Ising	100.00 ± 0.00	100.00 ± 0.00	100.00 ± 0.00	99.14 ± 0.14	96.57 ± 0.26
	12.5%	98.06 ± 1.13	98.85 ± 0.52	99.08 ± 0.40	96.73 ± 0.59	91.96 ± 0.72
	25.0%	95.95 ± 1.30	96.86 ± 0.90	97.40 ± 0.52	96.28 ± 0.56	90.73 ± 1.28
	50.0%	93.39 ± 1.08	96.26 ± 1.07	95.13 ± 0.83	94.69 ± 0.41	90.63 ± 0.94
	75.0%	91.69 ± 1.31	94.69 ± 1.22	92.74 ± 1.26	92.23 ± 0.53	91.03 ± 0.88

Table 1: Solution optimality on D-Wave quantum annealers after and before compressing Ising models with GRANITE.

	BCE	MSE	Hybrid	Softmax
ER	86.90 ± 1.08	75.68 ± 3.85	$\mathbf{91.07 \pm 0.82}$	89.68 ± 1.07
BA	89.05 ± 1.35	92.07 ± 0.69	$\mathbf{92.30 \pm 0.87}$	78.89 ± 2.44
WS	92.58 ± 0.30	$\mathbf{94.02 \pm 0.60}$	92.23 ± 0.53	92.60 ± 0.51

Table 2: Optimality for different loss functions for n=200, edge reduction 75%.

# Layers	ER	BA	WS
1	76.64 ± 0.86	82.95 ± 0.72	85.73 ± 1.05
2	88.09 ± 0.45	86.47 ± 1.07	91.56 ± 1.03
3	$\mathbf{91.07 \pm 0.82}$	$\mathbf{92.30 \pm 0.87}$	$\mathbf{92.23 \pm 0.53}$
4	81.07 ± 1.27	86.47 ± 0.91	85.79 ± 1.39
5	84.48 ± 1.49	90.72 ± 0.85	90.94 ± 0.63

Table 3: Optimality with different number of GNN layers at n = 200 and reduction rate = 75%.

Comparison with Baselines. There is a clear performance gap between GRANITE and random reduction strategies as shown in Figure 2. While GRANITE maintains optimality above 91% across all topologies, random baseline approaches achieve significantly lower performance, under 40% for all network topologies. This stark contrast demonstrates that GRANITE's learned compression strategies significantly outperform random reduction approaches. Moreover, as the first method to offer variable compression ratios, GRANITE provides a unique advantage over existing approaches, which typically achieve less than 20% reduction on average.

Ablation Studies. We conducted comprehensive ablation studies to evaluate the impact of two key architectural choices: loss function and the number of GNN layers. For loss functions, we compared Binary Cross-Entropy (BCE), Mean Squared Error (MSE), our proposed hybrid loss, and softmax-based weighting. The results in Table 2 demonstrate that our hybrid loss function achieves better performance on both Erdős-Rényi (91.07%) and Barabási-Albert (92.30%) graphs while performing competitively on Watts-Strogatz topology (92.23%). GNN depth analysis, shown in Table 3, reveals that a three-layer architecture consistently achieves optimal performance across all graph types, with peak optimality of 91.07% (ER), 92.30% (BA), and 92.23% (WS). Deeper GNN architecture with 4 and 5 layers shows performance degradation, suggesting that three layers provide sufficient representational capacity while avoiding overfitting. These findings guided our choice of hybrid loss and three-layer architecture for the final GRANITE model.

Conclusion and Future Work

While extensive efforts have been made to tackle combinatorial optimization problems from various angles, particularly in solver development, the approach of dynamic qubit compression has remained largely unexplored. This paper introduces GRANITE, an automated, fast system that iteratively compresses large Ising models while maintaining solution accuracy. GRANITE's has demonstrated effective compression of large-scale graphs across different random graph models, resulting in significant physical qubit reduction on D-Wave quantum annealing machines. The potential of GRANITE opens up new avenues for solving large-scale optimization problems on quantum hardware with limited qubit counts. Future work will focus on investigating the effectiveness of our qubit compression model on real-world large graphs and exploring potential applications in various domains. Additionally, adapting GRANITE to other quantum computing technologies while addressing hardware-specific noise challenges presents an important direction for extending its impact across different quantum computing paradigms.

Acknowledgements

This work was partly supported by NSF AMPS-2229075, VCU Quest Award, and Commonwealth Cyber Initiative.

References

Acharya, R.; Aghababaie-Beni, L.; Aleiner, I.; Andersen, T. I.; Ansmann, M.; Arute, F.; Arya, K.; Asfaw, A.; Astrakhantsev, N.; Atalaya, J.; et al. 2024. Quantum error correction below the surface code threshold. *arXiv preprint arXiv:2408.13687*.

Angelini, M. C.; and Ricci-Tersenghi, F. 2022. Modern graph neural networks do worse than classical greedy algorithms in solving combinatorial optimization problems like maximum independent set. *Nature Machine Intelligence*, 29–31.

Angelini, M. C.; and Ricci-Tersenghi, F. 2023. Modern graph neural networks do worse than classical greedy algorithms in solving combinatorial optimization problems like maximum independent set. *Nature Machine Intelligence*, 5(1): 29–31.

Arute, F.; Arya, K.; Babbush, R.; Bacon, D.; Bardin, J. C.; Barends, R.; Biswas, R.; Boixo, S.; Brandao, F. G.; Buell, D. A.; et al. 2019. Quantum supremacy using a programmable superconducting processor. *Nature*, 574(7779): 505–510.

Bauza, H. M.; and Lidar, D. A. 2024. Scaling Advantage in Approximate Optimization with Quantum Annealing. arXiv:2401.07184.

Boettcher, S. 2022. Inability of a graph neural network heuristic to outperform greedy algorithms in solving combinatorial optimization problems. *Nature Machine Intelligence*, 2522–5839.

Boothby, K.; Bunyk, P.; Raymond, J.; and Roy, A. 2020. Next-Generation Topology of D-Wave Quantum Processors. ArXiv:2003.00133.

Boros, E.; and Hammer, P. L. 2002. Pseudo-boolean optimization. *Discrete applied mathematics*, 123(1-3): 155–225.

Cai, J.; Macready, W. G.; and Roy, A. 2014. A practical heuristic for finding graph minors. *arXiv preprint arXiv:1406.2741*.

Carleo, G.; Cirac, I.; Cranmer, K.; Daudet, L.; Schuld, M.; Tishby, N.; Vogt-Maranto, L.; and Zdeborová, L. 2019. Machine learning and the physical sciences. *Rev. Mod. Phys.*, 91: 045002.

Choi, V. 2008. Minor-embedding in adiabatic quantum computation: I. The parameter setting problem. *Quantum Information Processing*, 7(5): 193–209.

Choi, V. 2011. Minor-embedding in adiabatic quantum computation: II. Minor-universal graph design. *Quantum Information Processing*, 10(3): 343–353.

Dai, H.; Khalil, E. B.; Zhang, Y.; Dilkina, B.; and Song, L. 2017. Learning Combinatorial Optimization Algorithms over Graphs. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, NIPS'17, 6351–6361. ISBN 9781510860964.

Farhi, E.; Goldstone, J.; Gutmann, S.; and Sipser, M. 2000. Quantum Computation by Adiabatic Evolution. ArXiv:quant-ph/0001106.

Fox, D. M.; Branson, K. M.; and Walker, R. C. 2021. mRNA codon optimization with quantum computers. *PloS one*, 16(10): e0259101.

Fox, D. M.; MacDermaid, C. M.; Schreij, A. M.; Zwierzyna, M.; and Walker, R. C. 2022. RNA folding using quantum computers. *PLOS Computational Biology*, 18(4): e1010032.

Gary, M. R.; and Johnson, D. S. 1979. Computers and Intractability: A Guide to the Theory of NP-completeness.

Gasse, M.; Chetelat, D.; Ferroni, N.; Charlin, L.; and Lodi, A. 2019. Exact Combinatorial Optimization with Graph Convolutional Neural Networks. In *Neural Information Processing Systems*.

Glover, F.; Kochenberger, G.; and Du, Y. 2018. A tutorial on formulating and using QUBO models. *arXiv preprint arXiv:1811.11538*.

Goto, H.; Endo, K.; Suzuki, M.; Sakai, Y.; Kanao, T.; Hamakawa, Y.; Hidaka, R.; Yamasaki, M.; and Tatsumura, K. 2021. High-performance combinatorial optimization based on classical mechanics. *Science Advances*, 7(6): eabe7953.

Grozea, C.; Hans, R.; Koch, M.; Riehn, C.; and Wolf, A. 2021. Optimising Rolling Stock Planning including Maintenance with Constraint Programming and Quantum Annealing. *arXiv preprint arXiv:2109.07212*.

Gurobi Optimization, LLC. 2024. Gurobi Optimizer Reference Manual.

Hammer, P. L.; Hansen, P.; and Simeone, B. 1984. Roof duality, complementation and persistency in quadratic 0–1 optimization. *Mathematical programming*, 28(2): 121–155.

Hamze, F.; Raymond, J.; Pattison, C. A.; Biswas, K.; and Katzgraber, H. G. 2020. Wishart planted ensemble: A tunably rugged pairwise Ising model with a first-order phase transition. *Physical Review E*, 101(5): 052102.

Harris, R.; Sato, Y.; Berkley, A.; Reis, M.; Altomare, F.; Amin, M.; Boothby, K.; Bunyk, P.; Deng, C.; Enderud, C.; et al. 2018. Phase transitions in a programmable quantum spin glass simulator. *Science*, 361(6398): 162–165.

Hauke, P.; Katzgraber, H. G.; Lechner, W.; Nishimori, H.; and Oliver, W. D. 2020. Perspectives of quantum annealing: methods and implementations. *Reports on Progress in Physics*, 83(5): 054401.

Hibat-Allah, M.; Ganahl, M.; Hayward, L. E.; Melko, R. G.; and Carrasquilla, J. 2020. Recurrent neural network wave functions. *Physical Review Research*, 2(2): 023358.

Hibat-Allah, M.; Inack, E. M.; Wiersema, R.; Melko, R. G.; and Carrasquilla, J. 2021. Variational neural annealing. *Nature Machine Intelligence*, 3(11): 952–961.

Honjo, T.; Sonobe, T.; Inaba, K.; Inagaki, T.; Ikuta, T.; Yamada, Y.; Kazama, T.; Enbutsu, K.; Umeki, T.; Kasahara, R.; et al. 2021. 100,000-spin coherent Ising machine. *Science advances*, 7(40): eabh0952.

Joshi, C. K.; Laurent, T.; and Bresson, X. 2019. An Efficient Graph Convolutional Network Technique for the Travelling Salesman Problem. *ArXiv*, abs/1906.01227.

Kadowaki, T.; and Nishimori, H. 1998a. Quantum annealing in the transverse Ising model. *Physical Review E*, 58(5): 5355.

Kadowaki, T.; and Nishimori, H. 1998b. Quantum annealing in the transverse Ising model. *Physical Review E*, 58(5): 5355–5363.

King, A. D.; Bernoudy, W.; King, J.; Berkley, A. J.; and Lanting, T. 2018. Emulating the coherent Ising machine with a mean-field algorithm. *arXiv preprint arXiv:1806.08422*.

King, A. D.; Nocera, A.; Rams, M. M.; Dziarmaga, J.; Wiersema, R.; Bernoudy, W.; Raymond, J.; Kaushal, N.; Heinsdorf, N.; Harris, R.; Boothby, K.; Altomare, F.; Berkley, A. J.; Boschnak, M.; Chern, K.; Christiani, H.; Cibere, S.; Connor, J.; Dehn, M. H.; Deshpande, R.; Ejtemaee, S.; Farré, P.; Hamer, K.; Hoskinson, E.; Huang, S.; Johnson, M. W.; Kortas, S.; Ladizinsky, E.; Lai, T.; Lanting, T.; Li, R.; MacDonald, A. J. R.; Marsden, G.; McGeoch, C. C.; Molavi, R.; Neufeld, R.; Norouzpour, M.; Oh, T.; Pasvolsky, J.; Poitras, P.; Poulin-Lamarre, G.; Prescott, T.; Reis, M.; Rich, C.; Samani, M.; Sheldan, B.; Smirnov, A.; Sterpka, E.; Clavera, B. T.; Tsai, N.; Volkmann, M.; Whiticar, A.; Whittaker, J. D.; Wilkinson, W.; Yao, J.; Yi, T. J.; Sandvik, A. W.; Alvarez, G.; Melko, R. G.; Carrasquilla, J.; Franz, M.; and Amin, M. H. 2024. Computational supremacy in quantum simulation. arXiv:2403.00910.

King, A. D.; Raymond, J.; Lanting, T.; Isakov, S. V.; Mohseni, M.; Poulin-Lamarre, G.; Ejtemaee, S.; Bernoudy, W.; Ozfidan, I.; Smirnov, A. Y.; et al. 2021. Scaling advantage over path-integral Monte Carlo in quantum simulation of geometrically frustrated magnets. *Nature com.*, 12(1): 1113.

King, A. D.; Suzuki, S.; Raymond, J.; Zucca, A.; Lanting, T.; Altomare, F.; Berkley, A. J.; Ejtemaee, S.; Hoskinson, E.; Huang, S.; Ladizinsky, E.; MacDonald, A. J. R.; Marsden, G.; Oh, T.; Poulin-Lamarre, G.; Reis, M.; Rich, C.; Sato, Y.; Whittaker, J. D.; Yao, J.; Harris, R.; Lidar, D. A.; Nishimori, H.; and Amin, M. H. 2022. Coherent quantum annealing in a programmable 2,000 qubit Ising chain. *Nature Physics*, 18(11): 1324–1328.

Kirkpatrick, S.; Gelatt Jr, C. D.; and Vecchi, M. P. 1983. Optimization by simulated annealing. *science*, 220(4598): 671–680.

Li, Z.; Chen, Q.; and Koltun, V. 2018. Combinatorial Optimization with Graph Convolutional Networks and Guided Tree Search. In *Neural Information Processing Systems*.

Lucas, A. 2014. Ising formulations of many NP problems. *Frontiers in physics*, 5.

Mugel, S.; Abad, M.; Bermejo, M.; Sánchez, J.; Lizaso, E.; and Orús, R. 2021. Hybrid quantum investment optimization with minimal holding period. *Scientific Reports*, 11(1): 1–6.

Oshiyama, H.; and Ohzeki, M. 2022. Benchmark of quantum-inspired heuristic solvers for quadratic unconstrained binary optimization. *Scientific reports*, 12(1): 2146.

Pan, F.; Zhou, P.; Zhou, H.-J.; and Zhang, P. 2021. Solving statistical mechanics on sparse graphs with feedback-set variational autoregressive networks. *Phys. Rev. E*, 103: 012103.

Panchenko, D. 2013. *The sherrington-kirkpatrick model*. Springer Science & Business Media.

Patel, S.; Chen, L.; Canoza, P.; and Salahuddin, S. 2020. Ising model optimization problems on a FPGA accelerated restricted Boltzmann machine. *arXiv preprint arXiv:2008.04436*.

Peruzzo, A.; McClean, J.; Shadbolt, P.; Yung, M.-H.; Zhou, X.-Q.; Love, P. J.; Aspuru-Guzik, A.; and O'brien, J. L. 2014. A variational eigenvalue solver on a photonic quantum processor. *Nature communications*, 5(1): 4213.

Preskill, J. 2018. Quantum Computing in the NISQ era and beyond. *Quantum*, 2: 79.

Rother, C.; Kolmogorov, V.; Lempitsky, V.; and Szummer, M. 2007. Optimizing binary MRFs via extended roof duality. In *2007 IEEE conference on computer vision and pattern recognition*, 1–8. IEEE.

Santoro, G. E.; and Tosatti, E. 2006. Optimization using quantum mechanics: quantum annealing through adiabatic evolution. *Journal of Physics A: Mathematical and General*, 39(36): R393.

Schuetz, M. J. A.; Brubaker, J. K.; and Katzgraber, H. G. 2021. Combinatorial Optimization with Physics-Inspired Graph Neural Networks. *Nature Machine Intelligence*, 4: 367–377.

Schuetz, M. J. A.; Brubaker, J. K.; Zhu, Z.; and Katzgraber, H. G. 2022. Graph coloring with physics-inspired graph neural networks. *Phys. Rev. Res.*, 4: 043131.

Tabi, Z. I.; Marosits, Á.; Kallus, Z.; Vaderna, P.; Gódor, I.; and Zimborás, Z. 2021. Evaluation of Quantum Annealer Performance via the Massive MIMO Problem. *IEEE Access*, 9: 131658–131671.

Tanaka, A.; Tomiya, A.; and Hashimoto, K. 2023. *Deep Learning and Physics*. Springer Singapore.

Tasseff, B.; Albash, T.; Morrell, Z.; Vuffray, M.; Lokhov, A. Y.; Misra, S.; and Coffrin, C. 2022. On the Emerging Potential of Quantum Annealing Hardware for Combinatorial Optimization. ArXiv:2210.04291.

Tasseff, B.; Albash, T.; Morrell, Z.; Vuffray, M.; Lokhov, A. Y.; Misra, S.; and Coffrin, C. 2024. On the emerging potential of quantum annealing hardware for combinatorial optimization. *Journal of Heuristics*, 1–34.

Thai, P.; Thai, M. T.; Vu, T.; and Dinh, T. N. 2022. FastHare: Fast Hamiltonian Reduction for Large-scale Quantum Annealing. arXiv:2205.05004.

Tiunov, E. S.; Ulanov, A. E.; and Lvovsky, A. 2019. Annealing by simulating the coherent Ising machine. *Optics express*, 27(7): 10288–10295.

Wu, D.; Wang, L.; and Zhang, P. 2019. Solving statistical mechanics using variational autoregressive networks. *Physical review letters*, 122(8): 080602.

Yarkoni, S.; Alekseyenko, A.; Streif, M.; Von Dollen, D.; Neukart, F.; and Bäck, T. 2021. Multi-car paint shop optimization with quantum annealing. In *2021 IEEE International Conference on Quantum Computing and Engineering (QCE)*, 35–41. IEEE.

Zhou, Y.; and Zhang, P. 2022. Noise-Resilient Quantum Machine Learning for Stability Assessment of Power Systems. *IEEE Transactions on Power Systems*.