
Neural Approaches to SAT Solving: Design Choices and Interpretability

David Mojžíšek², Jan Hůla^{1,2}, Ziwei Li¹, Ziyu Zhou¹, Mikoláš Janota¹

¹Czech Institute of Informatics, Robotics and Cybernetics, Czech Technical University in Prague, Czechia

²University of Ostrava, Ostrava, Czechia

david.mojzisek@osu.cz

Abstract

In this contribution, we provide a comprehensive evaluation of graph neural networks applied to Boolean satisfiability problems, accompanied by an intuitive explanation of the mechanisms enabling the model to generalize to different instances. We introduce several training improvements, particularly a novel closest assignment supervision method that dynamically adapts to the model’s current state, significantly enhancing performance on problems with larger solution spaces. Our experiments demonstrate the suitability of variable-clause graph representations with recurrent neural network updates, which achieve good accuracy on SAT assignment prediction while reducing computational demands. We extend the base graph neural network into a diffusion model that facilitates incremental sampling and can be effectively combined with classical techniques like unit propagation. Through analysis of embedding space patterns and optimization trajectories, we show how these networks implicitly perform a process very similar to continuous relaxations of MaxSAT, offering an interpretable view of their reasoning process. This understanding guides our design choices and explains the ability of recurrent architectures to scale effectively at inference time beyond their training distribution, which we demonstrate with test-time scaling experiments.

Keywords: Graph Neural Networks, Boolean Satisfiability, Diffusion Models, Test-time scaling, Interpretability

1 Introduction

Reasoning is a cognitive ability which allows humans to solve problems with previously unseen combinations of constraints. For a long time, it has been debated whether artificial neural networks can obtain such generalization skills or whether they can only learn to detect superficial patterns Fodor and Pylyshyn [1988], Marcus [2003, 2018] without being able to generalize to novel combinations of constraints. With the arrival of Large Language Models (LLMs) specially trained for reasoning Guo et al. [2025], Jaech et al. [2024], it became harder and harder to claim that these models can only detect superficial patterns. Nevertheless, the exact mechanism by which they are able to solve tasks that typically require reasoning is largely unknown and the robustness of the solving process is also not understood.

In this contribution, we focus on a restricted class of problems that require reasoning, concretely on solving Boolean formulas in CNF form. This could be viewed as a prototypical task where the goal is to solve problems with novel combinations of constraints, and where detecting superficial patterns seen during training would be insufficient. It has already been demonstrated that Graph Neural Networks (GNNs) can successfully learn to solve such problems and generalize to larger problems Selsam et al. [2018], even though they are still not competitive when compared to state of the art SAT solvers.

Understanding the underlying mechanisms GNNs employ to successfully solve problems, as well as their limitations, would offer significant practical and theoretical value. On the practical side, the trained model can be used as a guessing heuristic inside classical solvers, improving

their performance and on the theoretical side, understanding how a GNN can solve a CNF formula could help us to elucidate the reasoning ability of Transformers Vaswani et al. [2017] because Transformers can be viewed as GNNs in which the graph connectivity is given by the attention map and is learned from data Cai et al. [2023]. Our aim in this contribution is to provide an experimental evaluation of different design choices for GNNs in the context of Boolean satisfiability together with an intuitive explanation of the inner workings of these models. Our main contributions are as follows:

- We provide an experimental comparison of different architectures and training regimes.
- We introduce a novel supervision method based on the closest assignment, resulting in significant improvements.
- We demonstrate that these architectures scale well at test time.
- We extend the graph neural network to a diffusion model and show how it relates to the base model.
- We provide an intuitive explanation for the inner workings of these models.

The rest of the text has the following structure: Section 3 (Relevant Background) provides the necessary context on Boolean satisfiability problems, SAT solving approaches, graph neural networks, theoretical connection to approximation algorithms, and diffusion models. Section 4 (Experimental Setup) describes our methodology, including data representation choices, architecture variants, supervision methods, and benchmark generation. Section 5 (Experimental Results) presents our comprehensive evaluation, comparing different graph representations and training methods (Section 5.2), demonstrating test-time scaling capabilities (Section 5.3), and introducing our diffusion model extension (Section 5.4). Section 6 (Interpreting the Trained Model) offers analysis of the embedding space and explains the networks’ behavior through the lens of approximation algorithms based on continuous relaxation. Section 2 (Related Work) positions our contribution within the broader research landscape, and Section 7 contains a discussion of our findings and directions for future research. We conclude in Section 8. Additional implementation details and mathematical derivations are provided in the Appendix.

2 Related Work

Our research builds directly upon NeuroSAT Selsam et al. [2018], which introduced the first end-to-end neural approach for SAT solving using a recurrent message-passing architecture. While we maintain the core iterative design of NeuroSAT (allowing variable numbers of message-passing iterations through weight sharing), we explore simplified variants using RNNs and LSTMs and incorporate techniques like curriculum learning to improve training efficiency.

Several other works have explored different directions in neural SAT solving. Li et al. [2023] developed G4SATBench to benchmark various GNN architectures (GCN, GGNN, GIN) across different graph representations and supervision objectives. Unlike their broader exploration across architecture types, our work focuses on the recurrent message-passing paradigm from NeuroSAT and investigates how different training objectives and graph representations affect performance within this specific framework. We also mention the work by Warde et al. [2023] who developed a recurrent architecture based on a Restricted Boltzmann Machine.

Hybrid approaches that integrate neural networks with traditional solvers include NeuroCore by Selsam and Bjørner [2019], which uses neural predictions to guide variable branching in CDCL solvers. Similarly, Wang et al. [2021] proposed NeuroComb to enhance CDCL solvers through GNN-based identification of important variables and clauses.

These approaches differ from our end-to-end model but demonstrate alternative applications of neural methods to SAT solving.

The connection between neural networks and continuous relaxations is particularly relevant to our work. Kyrillidis et al. [2020] introduced FourierSAT, which transforms Boolean SAT problems into continuous optimization using the Walsh-Fourier transform. This approach provides a theoretical foundation for understanding how neural networks might implicitly convert discrete search problems into continuous optimization. Similar technique was introduced by Hosny et al. [2024] who develop GPU-accelerated approaches for MaxSAT problems. Hula et al. [2024] and Yau et al. [2024] explore the connection between GNNs and semidefinite programming relaxations, demonstrating empirically and theoretically that message-passing can implement gradient-based optimization of SDP relaxations.

In the broader domain of combinatorial optimization, Sun et al. [2023] used diffusion models based on GNNs to solve problems such as traveling salesman.

3 Relevant Background

3.1 Boolean Satisfiability and Maximum Satisfiability

3.1.1 Boolean Satisfiability as a Constraint Satisfaction Problem

Boolean satisfiability (SAT) is a fundamental problem in computer science that asks whether a given Boolean formula has a satisfying assignment. The formula is built from propositional variables x_1, x_2, \dots that can take values from $\{0, 1\}$, representing false and true respectively, and logical connectives: conjunction (\wedge), disjunction (\vee), and negation (\neg). While other connectives like implication (\rightarrow) and equivalence (\leftrightarrow) exist, they can be expressed using these basic operators.

A literal is either a propositional variable x or its negation $\neg x$. While Boolean formulas can take arbitrary form, the most common representation is the conjunctive normal form (CNF), where a formula is a conjunction of clauses, and each clause is a disjunction of literals. For example, $(x_1 \vee \neg x_2) \wedge (x_2 \vee x_3)$ is a CNF formula with two clauses. We note that any Boolean formula can be transformed into an equisatisfiable CNF formula, albeit potentially requiring additional variables.

An assignment σ maps each propositional variable to either 0 or 1. We say σ satisfies a CNF formula if at least one literal in each clause evaluates to true under σ . For instance, the assignment $\sigma(x_1) = 1, \sigma(x_2) = 0, \sigma(x_3) = 1$ satisfies the formula $(x_1 \vee \neg x_2) \wedge (x_2 \vee x_3)$ as both clauses contain a true literal.

SAT is a special case of the more general Constraint Satisfaction Problem (CSP) framework Brailsford et al. [1999]. A CSP consists of a set of variables, each with a domain of possible values, and a set of constraints that specify allowed combinations of values for groups of variables. While SAT variables are restricted to Boolean values and constraints take the form of clauses, CSPs can accommodate richer variable domains and constraint types.

3.1.2 MaxSAT: The Optimization Variant

The Maximum Satisfiability problem (MaxSAT) is the optimization version of SAT. Given a CNF formula ϕ , the goal is to find an assignment that maximizes the number of satisfied clauses. This formulation is particularly useful when a formula is unsatisfiable, as MaxSAT still yields the best possible solution.

MaxSAT has several variations that differ in their expressiveness and the way they handle the importance of clauses. In unweighted MaxSAT, all clauses have equal importance. Weighted MaxSAT assigns a positive weight to each clause, with the objective being to maximize the sum

of weights of satisfied clauses. In some variants (Partial MAX-SAT Fu and Malik [2006]), clauses are categorized as hard or soft, where hard clauses must be satisfied (often with infinite weight), and soft clauses are those that can be violated but contribute to the objective based on their weight. While weighted variants exist, in this paper we focus exclusively on the unweighted formulation.

Formally, for a CNF formula $\phi = c_1 \wedge c_2 \wedge \dots \wedge c_m$ with m clauses, the unweighted MaxSAT problem seeks an assignment σ^* such that:

$$\sigma^* = \arg \max_{\sigma} \sum_{i=1}^m \mathbf{1}(\sigma \text{ satisfies } c_i) \quad (1)$$

where $\mathbf{1}(\cdot)$ is the indicator function.

3.2 SAT Solving Approaches

SAT solving algorithms are generally categorized into complete and incomplete approaches, each with distinct characteristics and applications.

Complete Solvers Complete solvers theoretically guarantee definitive answers: either finding a satisfying assignment or proving that none exists. The Davis-Putnam-Logemann-Loveland (DPLL) algorithm forms the foundation for most modern complete solvers Biere et al. [2009]. It systematically explores the search space through backtracking while employing unit propagation to deduce logical consequences.

Conflict-Driven Clause Learning (CDCL) extends DPLL by analyzing conflicts to learn new clauses, which helps prune large portions of the search space. When a conflict occurs, the solver identifies the "reasons" for the conflict and adds a new clause that prevents similar conflicts in the future. Modern CDCL solvers incorporate sophisticated heuristics for variable selection, restart strategies, and efficient data structures to improve performance.

Incomplete Solvers Incomplete solvers focus on finding satisfying assignments but cannot prove unsatisfiability. These algorithms are particularly effective for large satisfiable instances where complete methods might be inefficient.

Local search algorithms, such as WalkSAT Selman et al. [1993], start with a random assignment and iteratively modify it to satisfy more clauses. These methods employ heuristics to decide which variables to flip at each step, balancing between greedy choices and random moves to escape local optima. For MaxSAT, local search algorithms often use scoring functions that prioritize flipping variables that maximize the increase in satisfied clauses.

Stochastic algorithms including simulated annealing and genetic algorithms have also been applied to SAT and MaxSAT problems. These approaches can effectively explore search spaces in certain problem classes where deterministic methods struggle.

3.2.1 Continuous Relaxations

A specific type incomplete solvers that have been explored in recent research Kyrillidis et al. [2020], Hosny and Reda [2024] are explored continuous relaxations of MaxSAT, that transform the discrete problem into a continuous optimization task. These methods map Boolean variables to continuous domains, enabling the application of gradient-based optimization techniques. The Fourier-SAT method Kyrillidis et al. [2020], for instance, transforms Boolean formulas into multilinear polynomials through Walsh-Fourier transform and then optimize the continuous variables w.r.t. the resulting polynomial.

Continuous relaxations can also be obtained by making the objective function convex as often done when designing approximation algorithms which provide guarantees for their performance.

The guarantees can be improved by lifting the variables into a high-dimensional vector space and optimizing vectors instead of scalar values. The optimized vectors are finally rounded to discrete values.

Semidefinite Programming (SDP) relaxation, particularly for MAX-2-SAT or MAX-3-SAT, illustrates this approach elegantly. In SDP relaxation, Boolean variables $x_i \in \{0, 1\}$ are transformed into unit vectors \mathbf{y}_i in a high-dimensional space. An additional vector \mathbf{y}_0 is introduced to represent the value "true." The Boolean variable x_i is considered true if \mathbf{y}_i is close to \mathbf{y}_0 (positive inner product) and false if it is far from \mathbf{y}_0 (negative inner product).

The optimization process for these vectors follows a pattern:

1. Initialize random unit vectors for each variable
2. Optimize these vectors to maximize the number of satisfied clauses, expressed as a function of inner products between vectors
3. Round the resulting vectors to discrete assignments (typically based on the sign of inner products with \mathbf{y}_0)

This relaxation enables the application of powerful continuous optimization techniques while providing approximation guarantees. For MAX-2-SAT, this approach yields an approximation ratio of 0.878, meaning the solution will satisfy at least 87.8% of the maximum possible number of clauses.

3.2.2 Learning-Based Approaches

Machine learning (ML) is also being heavily utilized for SAT solving. Many approaches have been developed to guide traditional solvers Selsam and Bjørner [2019], Yolcu and Póczos [2019] or to solve SAT problems directly Selsam et al. [2018], Li and Si [2022]. To guide a solver, a neural network can be used to replace heuristics such as variable selection or restart policies. Importantly, Graph Neural Networks (GNNs) can also be trained to solve SAT problems end-to-end without relying on traditional algorithmic solvers Amizadeh et al. [2018]. These GNN-based approaches can operate directly on the graph representation of Boolean formulas, with variables and clauses forming nodes in a bipartite graph, and learn to predict satisfiability or produce satisfying assignments Li et al. [2023]. In this work, we focus on variants of GNNs that are recurrent and this allows us to scale the computation during inference or adapt the number of iterations for each instance separately.

In Section 6 we will show an evidence that one can view the end-to-end ML approaches as bi-level optimization methods because during inference, the GNN behaves as a continuous solver trying to maximize the number of satisfied clauses. Therefore, during training, the outer loop of the bi-level optimization optimizes the weights of the network which then runs an inner loop that optimizes the values of variables to maximize the number of satisfiable clauses.

3.3 Graph Neural Networks

Graph Neural Networks (GNNs) extend deep learning to graph-structured data, enabling learning on irregular data structures that classical neural architectures cannot directly process. A graph $G = (V, E)$ consists of nodes V and edges E , where each node $v \in V$ may have associated features x_v .

GNNs compute node representations through message passing, where each node iteratively aggregates information from its neighbors and updates its features. Formally, at layer l , a node v updates its representation h_v^l according to:

$$h_v^{l+1} = \text{UPDATE}(h_v^l, \text{AGGREGATE}(\{h_u^l : u \in \mathcal{N}(v)\}))$$

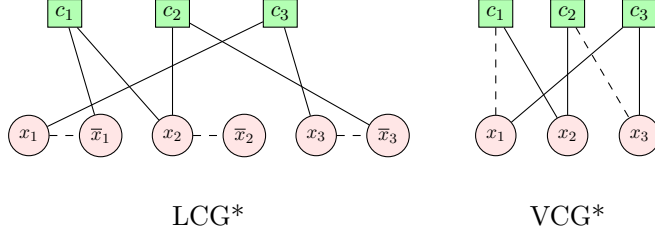


Figure 1: LCG and VCG of the CNF formula $(\bar{x}_1 \vee x_2) \wedge (x_2 \vee \bar{x}_3) \wedge (x_1 \vee x_3)$.

where $\mathcal{N}(v)$ denotes the neighbors of node v . The UPDATE and AGGREGATE functions are typically neural networks, often implementing permutation-invariant operations like sum or max. Through multiple layers of message passing, GNNs can capture both local structure and longer-range dependencies in the graph, making them suitable for processing SAT formulas represented as bipartite graphs.

3.4 Diffusion-based Assignment Generation

In Section 5.4 will show how the GNNs we use can be extended to diffusion models which have in recent years emerged as a powerful approach for generative modeling across domains Ho et al. [2020]. These models learn to transform a random noise distributions (such as multi-variate Gaussian distribution) to complex distributions behind the given domain (i.e., distribution of images of human faces). For practical applications, diffusion models are typically conditioned on an input so that the generated sample has specific characteristics. In our case, we will condition the model by the bipartite graph of the CNF formula.

3.4.1 Categorical Diffusion Process

While continuous diffusion models have gained prominence in image generation and other domains, discrete diffusion processes well-suited for combinatorial optimization problems like MAX-SAT, where the state space is inherently discrete. Our approach presented in Section 5.4 leverages a discrete diffusion process with categorical noise to model the generation of variable assignments. We adapt a concrete form of discrete diffusion first presented by Austin et al. Austin et al. [2021] and later leveraged for combinatorial optimization with GNNs by Sun et al. Sun and Yang [2023].

On a high level, diffusion models are trained to denoise noisy version of the training samples. These noisy versions are obtained by running a forward diffusion process for several steps and the model is then trained to predict the original sample. For a SAT problem with n variables, we represent each variable assignment as a binary value and the vector of these binary values represent the sample. The diffusion process gradually corrupts this sample until it becomes pure noise.

More concretely, the process that progressively adds noise to the initial assignment $\mathbf{x}_0 \in \{0, 1\}^n$ over T timesteps, produces a sequence of increasingly more corrupted assignments $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T$. For categorical diffusion, this corruption process is defined by a Markov chain with the following transition matrices:

$$\mathbf{Q}_t = \begin{pmatrix} 1 - \beta_t & \beta_t \\ \beta_t & 1 - \beta_t \end{pmatrix} \quad (2)$$

where $\beta_t \in (0, 1)$ represents the noise schedule, controlling how quickly the assignments become corrupted. The matrix \mathbf{Q}_t defines the probability of transitioning between states at time t , with the property that as t approaches T , the distribution of \mathbf{x}_t approaches a uniform distribution over all possible assignments.

To simplify inference, the cumulative transition matrices $\overline{\mathbf{Q}}_t = \mathbf{Q}_1 \mathbf{Q}_2 \cdots \mathbf{Q}_t$, which directly gives us $p(\mathbf{x}_t|\mathbf{x}_0)$ are being used. For the Boolean case, this allows us to efficiently sample \mathbf{x}_t given \mathbf{x}_0 using:

$$p(\mathbf{x}_t|\mathbf{x}_0) = \text{Cat}(\mathbf{x}_t; \mathbf{p} = \tilde{\mathbf{x}}_0 \overline{\mathbf{Q}}_t) \quad (3)$$

where $\tilde{\mathbf{x}}_0 \in \{0, 1\}^{n \times 2}$ is the one-hot encoding of \mathbf{x}_0 , with each variable represented by a vector $(1, 0)$ for value 0 or $(0, 1)$ for value 1. The Cat operation refers to the categorical distribution, which samples \mathbf{x}_t based on the probability vector \mathbf{p} .

3.4.2 Learning the Reverse Process

The core idea of diffusion models is to learn the reverse process - how to gradually denoise a corrupted sample to recover the original data distribution. In our case, we train a GNN to progressively recover a satisfiable assignment \mathbf{x}_0 starting from a random initial assignment. The trained model is used to sample from a distribution $p(\mathbf{x}_{t-1}|\mathbf{x}_t)$ which can be used to obtain an assignment \mathbf{x}_0 from random assignment \mathbf{x}_T as explained below. There are multiple ways of training the neural network used in the diffusion model. One can train it to directly model the distribution $p(\mathbf{x}_{t-1}|\mathbf{x}_t)$. In the method introduced by Austin et al. [2021], the network is trained to predict the original uncorrupted input \mathbf{x}_0 which is then used to sample from the posterior $p(\mathbf{x}_{t-1}|\mathbf{x}_t)$ using Bayes' rule. This approach provides stronger learning signals during training, as the target \mathbf{x}_0 remains fixed regardless of a timestep and we use it within this work.

3.4.3 Categorical Posterior Sampling

As mentioned above, our model is trained to predict \mathbf{x}_0 directly and we use this prediction during inference to sample \mathbf{x}_{t-1} given \mathbf{x}_t . This is accomplished through categorical posterior sampling, which uses the distribution $p_\theta(\mathbf{x}_0|\mathbf{x}_t, t)$ to compute the posterior $p(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0)$.

By applying Bayes' rule and the Markov property of the diffusion process, we can derive:

$$p(\mathbf{x}_{t-1}|\mathbf{x}_t) \approx \sum_{\mathbf{x}_0} p(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0) p_\theta(\mathbf{x}_0|\mathbf{x}_t, t) \quad (4)$$

For the categorical case, this is computed using:

$$p(\mathbf{x}_{t-1}|\mathbf{x}_t) \approx \sum_{\mathbf{x}_0} \frac{p(\mathbf{x}_{t-1}|\mathbf{x}_0) p(\mathbf{x}_t|\mathbf{x}_{t-1})}{p(\mathbf{x}_t|\mathbf{x}_0)} p_\theta(\mathbf{x}_0|\mathbf{x}_t, t) \quad (5)$$

The diffusion model replaces the distribution $p_\theta(\mathbf{x}_0|\mathbf{x}_t, t)$ with a function approximator (GNN in our case) $f_\theta(\mathbf{x}_t, t)$. Therefore, we can train the model using a simple procedure (predicting \mathbf{x}_0) and during inference, we can use a sampling process (iteratively sampling \mathbf{x}_{t-1} given \mathbf{x}_t), which tries to recover a uncorrupted input in several steps. A useful feature of diffusion models is that the number of sampling steps during inference can be chosen by the user after the model is already trained.

3.4.4 Inference Schedule

During inference, we can accelerate the generation process by using fewer denoising steps than were used during training or use more denoising steps with the hope to increase the quality of outputs. The tuple of time steps used for inference $(T, T-1, \dots, t_0)$ is called a *schedule*. The function approximator in the diffusion model is normally conditioned by the sample at a given time step and also the time step itself ($f_\theta(\mathbf{x}_t, t)$) but as we show in Section 5.4.1, the time step conditioning is not needed. This means that in our case the schedule is defined only by the number of time steps used.

4 Experimental Setup

4.1 Data Representation and Graph Structure

Boolean formulas in CNF form can be naturally represented as bipartite graphs where clauses and variables (or literals) form two distinct sets of nodes. In this work, we explore two different graph representations:

Literal-Clause Graph (LCG) In the literal-clause graph representation, each literal (both positive and negative polarity of a variable) is represented as a separate node. For a formula with n variables, this results in $2n$ literal nodes. Each literal node is connected to all clause nodes containing that literal. Formally, for a CNF formula ϕ with variables x_1, \dots, x_n and clauses c_1, \dots, c_m , we construct a bipartite graph $G_{LC} = (L \cup C, E)$ where:

- $L = \{l_1, \dots, l_n, \bar{l}_1, \dots, \bar{l}_n\}$ is the set of literal nodes
- $C = \{c_1, \dots, c_m\}$ is the set of clause nodes
- $(l_i, c_j) \in E$ if and only if literal l_i appears in clause c_j

Variable-Clause Graph (VCG) In the variable-clause graph representation, each variable (rather than each literal) is represented as a node. For a formula with n variables, this results in exactly n variable nodes. Each variable node is connected to all clause nodes containing either the positive or negative literal of that variable. To retain information about the polarity of literals, we assign edge features $p_{ij} \in \{-1, 1\}$ to each edge (x_i, c_j) , where $p_{ij} = 1$ if the positive literal x_i appears in clause c_j , and $p_{ij} = -1$ if the negative literal \bar{x}_i appears in clause c_j . Formally, we construct a bipartite graph $G_{VC} = (V \cup C, E, P)$ where:

- $V = \{x_1, \dots, x_n\}$ is the set of variable nodes
- $C = \{c_1, \dots, c_m\}$ is the set of clause nodes
- $(x_i, c_j) \in E$ if and only if variable x_i appears in clause c_j (in either polarity)
- $P : E \rightarrow \{-1, 1\}$ maps each edge to its corresponding polarity

Both graph representations capture the structure of the Boolean formula, but they differ in how they handle variable polarity. The literal-clause graph explicitly represents both polarities as separate nodes, which increases the number of nodes but simplifies the message passing process of the GNN. The variable-clause graph is more compact but requires handling polarity information through edge features. For the GNNs we use, the variable-clause graph representation is more computationally efficient than the literal-clause graph, reducing both memory requirements and processing time. This efficiency comes from having half as many variable nodes (compared to literal nodes) and avoiding an expensive operation during message passing as will be described in Section 4.2.

In our experiments, we compare both representations together with different message passing operations and different training regimes.

4.2 Architecture Variants

Our GNN architecture variants are derived from the NeuroSAT architecture Selsam et al. [2018] which demonstrated the possibility of using GNNs for SAT solving. The main advantage of this architecture is that it is recurrent and therefore the number of message passing iterations is theoretically not limited. This is not the case for the non-recurrent alternatives with fixed number of layers. We will demonstrate the usefulness of this feature in Section (5.3).

Node Embeddings Each node in the bi-partite graph of the formula is associated with a d -dimensional embedding vector ($d = 64$ in most of our experiments as a conclusion from an experiment in A.1.4). We initialize these embeddings randomly from a standard normal distribution. For a formula with n variables and m clauses, we have:

- In the literal-clause graph: $2n$ literal embeddings $\mathbf{l}_i \in \mathbb{R}^d$ and m clause embeddings $\mathbf{c}_j \in \mathbb{R}^d$
- In the variable-clause graph: n variable embeddings $\mathbf{v}_i \in \mathbb{R}^d$ and m clause embeddings $\mathbf{c}_j \in \mathbb{R}^d$

Message Passing Mechanism The core of our architecture is a two-phase message passing procedure that alternates between updating clause representations and unknown node representations (literals or variables, depending on the graph type). This process is repeated for a configurable number of iterations T .

We primarily use an RNN-based update mechanism, where the node embeddings are the hidden states of the RNN that evolve through message passing iterations. For the variable-clause graph, the message passing at iteration t is defined as:

$$\mathbf{h}_c^{(t)} = \text{RNN}_c \left(\sum_{v \in \mathcal{N}(c)} \mathbf{M}_{vc}(\mathbf{h}_v^{(t-1)}, p_{vc}), \mathbf{h}_c^{(t-1)} \right) \quad (6)$$

$$\mathbf{h}_v^{(t)} = \text{RNN}_v \left(\sum_{c \in \mathcal{N}(v)} \mathbf{M}_{cv}(\mathbf{h}_c^{(t)}, p_{vc}), \mathbf{h}_v^{(t-1)} \right) \quad (7)$$

Here, $\mathbf{h}_c^{(t)}$ and $\mathbf{h}_v^{(t)}$ are the hidden states that serve as the actual clause and variable node embeddings for clause nodes and variable nodes respectively. \mathbf{M}_{vc} and \mathbf{M}_{cv} are the message transformation functions that operate on the source node embedding and the edge polarity. For the variable-clause graph, we implement these transformation functions as two MLPs that process positive and negative edges differently:

$$\mathbf{M}_{vc}(\mathbf{h}_v, p) = \begin{cases} \text{MLP}_{\text{pos}}(\mathbf{h}_v) & \text{if } p > 0 \\ \text{MLP}_{\text{neg}}(\mathbf{h}_v) & \text{if } p < 0 \end{cases} \quad (8)$$

For the literal-clause graph, the message passing mechanism also uses operation, called “Flip” bellow, that enforces the logical relationship between complementary literals:

$$\mathbf{h}_c^{(t)} = \text{RNN}_c \left(\sum_{l \in \mathcal{N}(c)} \mathbf{h}_l^{(t-1)}, \mathbf{h}_c^{(t-1)} \right) \quad (9)$$

$$\mathbf{h}_l^{(t)} = \text{RNN}_l \left(\left[\sum_{c \in \mathcal{N}(l)} \mathbf{h}_c^{(t)}, \text{Flip}(\mathbf{h}_l^{(t-1)}) \right], \mathbf{h}_l^{(t-1)} \right) \quad (10)$$

where $[\cdot, \cdot]$ denotes vector concatenation. The $\text{Flip}(\cdot)$ operation exchanges the embeddings of positive literals with their corresponding negative literals and vice versa. The update function for a given literal embedding can therefore take into account the embedding of the complementary literal.

We note, that the $\text{Flip}(\cdot)$ operation incurs a significant computational cost, particularly for large formulas. In contrast, the variable-clause graph representation eliminates this expensive operation by dedicating only one node for each variable and directly encoding its polarity in edge features. This efficiency makes the variable-clause approach particularly well-suited for larger formulas where computational demands become a critical factor.

Apart from the RNN-based update functions, we also experiment with LSTM-based update functions which have been used in the original NeuroSAT architecture Selsam et al. [2018]. The LSTM-based updates follow a similar pattern but maintain an additional cell state alongside the hidden state. In Section 5.2 we show that different update functions are suitable for different settings.

After each update step, we apply L2 normalization to all node embeddings to stabilize training:

$$\mathbf{h}_i^{(t)} = \frac{\mathbf{h}_i^{(t)}}{\|\mathbf{h}_i^{(t)}\|_2} \quad (11)$$

Node classification After T iterations of message passing, we use the final node embeddings to predict variable assignments. For the variable-clause graph, we apply a linear layer to each variable embedding to produce two logits (representing scores for value true and false): $\mathbf{y}_v = \mathbf{W}\mathbf{h}_v^{(T)} + \mathbf{b}$. The assignment is then determined by applying softmax and taking the argmax: $\hat{a}_v = \arg \max_i (\text{softmax}(\mathbf{y}_v)_i)$.

For the literal-clause graph, we focus on the embeddings of positive literals only, as they directly correspond to variables. During training, we use cross-entropy loss between these predicted assignments and the ground truth assignments.

For satisfiability prediction, we can determine whether a formula is satisfiable by checking if the predicted assignment satisfies all clauses. The model is thus trained to find assignments that minimize the number of unsatisfied clauses, effectively solving the MaxSAT problem even when trained only with assignment supervision.

4.3 Supervision Tasks and Objectives

There are several obvious supervision objectives and prediction tasks which can be used to train the model. The original NeuroSAT model was trained to predict the satisfiability status of a given formula using binary cross-entropy. Later, several authors tried different training tasks and objectives which have been summarized in a review paper by Li et al. [2023]. We reimplement these objective and task for our setup and also introduce a novel training objective which in certain settings results in significant improvements of the model performance. These objective are briefly described below.

Satisfiability Classification This is the task which was used by Selsam et al. [2018] for training the original NeuroSAT architecture. The model is trained to predict whether the formula is satisfiable or not through graph-level embedding aggregation using global mean pooling. The loss is computed by binary cross-entropy between the prediction \hat{y} and ground truth $y \in \{0, 1\}$: $\mathcal{L}_{\text{sat}} = -(y \log \hat{y} + (1 - y) \log(1 - \hat{y}))$.

Unsupervised Training For unsupervised training, we define the loss using clause validity Ozolins et al. [2022], where \hat{x}_i represents the model’s predicted continuous probability of a variable being true:

$$V_c(\hat{x}) = 1 - \prod_{i \in c^+} (1 - \hat{x}_i) \prod_{i \in c^-} \hat{x}_i, \quad \mathcal{L}_\phi(\hat{x}) = - \sum_{c \in \phi} \log(V_c(\hat{x})), \quad (12)$$

where c^+ and c^- are the sets of variables that occur in clause c in positive and negative form respectively. This loss reaches its minimum only when the prediction \hat{x} is a satisfying assignment. We note that alternative unsupervised formulations exist Amizadeh et al. [2018], and comprehensive evaluations reported by Li et al. [2023] suggest that these two different

approaches perform similarly in practice. Another training option would be to directly optimize a convex loss function derived from SDP relaxation, but this approach is limited because SDP formulations work well for MAX-2-SAT and can be extended to MAX-3-SAT, but become increasingly difficult to formulate for general MaxSAT problems with larger clauses.

Assignment Prediction For satisfiable formulas, we can train the model to predict the satisfiable variable assignments directly. We tried to use either mean squared error or cross-entropy loss between the predicted assignments and the ground truth assignments: $\mathcal{L}_{\text{assign}}^{MSE} = \|\hat{a} - x\|_2^2$ and $\mathcal{L}_{\text{assign}}^{CE} = -\sum_i x_i \log \hat{x}_i + (1 - x_i) \log(1 - \hat{x}_i)$ where x is the ground truth assignment and \hat{a}, \hat{x} are the predicted assignments which differ by application of softmax (i.e. \hat{a} are just logits without a softmax applied).

Closest Assignment Training One problem with assignment prediction is that satisfiable formulas can have a lot of solution and the network is penalized even if it predicts satisfiable solution which differs from the one which is used as a ground truth. We therefore introduce a novel supervision method which uses a MaxSAT solver to always compute the solution which is closest to the solution predicted by the model. We then update then model with respect to this solution. In Section 5.2, we show that this method works particularly well when the solution space is large.

For each formula in a batch, a valid assignments that minimize the Hamming distance to the model’s current predictions is found by the RC2 MaxSAT solver. For satisfiable formulas it finds an assignment that satisfies all clauses while being closest to current prediction. For unsatisfiable formulas, it finds an assignment that maximizes the number of satisfied clauses while minimizing distance to prediction.

This approach allows the model to explore different regions of the solution space while maintaining valid solutions for SAT instances or optimal partial solutions for UNSAT instances. The supervision signal adapts to the model’s current state rather than forcing it toward a single pre-determined assignment. The disadvantage of this method is that the computation of the loss is slower then with the precomputed solution. This could be solved by pre-computing solutions or by using an approximate MaxSAT solver.

SAT-Only Instance Filtering After initially training with both satisfiable and unsatisfiable instances, we experimented with formula-type specialization by restricting training to only satisfiable instances. In Table 3, we show that this filtering can lead to higher accuracy of the trained model.

4.4 Benchmarks and Data Generation

We utilize two complementary benchmark generators for evaluating the tested variants: the SR generator and a 3-SAT generator with the ratio between variable and clauses set close to the phase transition point.

SR Generator The SR generator by Selsam et al. [Selsam et al., 2018] produces pairs of satisfiable and unsatisfiable formulas that differ by negating only a single literal. This design specifically prevents models from exploiting superficial features for classification. Intuitively, it works by iteratively sampling random clauses and adding them to a formula. After each addition, a SAT solver checks if the formula remains satisfiable. When adding a clause that finally makes the formula unsatisfiable, the generator saves this instance and creates its satisfiable counterpart by flipping a single literal in the last clause. To create each clause, it samples a small integer k based on a mix of Bernoulli and geometric distributions, then randomly selects k variables without replacement, negating each with 0.5 probability. This solver-driven approach ensures

that satisfiability classification requires understanding the logical structure rather than statistical properties. As reported in the review by Li et al. [2023], the models trained on problems from this generator transfer the best to other problem distributions.

3-SAT Generator We also employ a 3-SAT generator configured at the critical clause-to-variable ratio of 4.26, known as the phase transition point where SAT problems are empirically the most challenging to solve [Crawford and Auton, 1996]. At this ratio, approximately half of the generated instances are satisfiable. Each clause contains exactly 3 literals selected uniformly from the available variables, with each literal negated with 0.5 probability. Unlike the SR generator, 3-SAT focuses on generating naturally difficult problems rather than explicitly preventing superficial feature learning.

5 Experimental Results

5.1 Training and Evaluation Methodology

For training, we generate 50,000 instances: 25,000 pairs for SR and 50,000 instances for 3-SAT. We annotate each dataset by the maximum number of variables appearing in the training formulas. For SR, we test two variations, SR40 for which the training examples are sampled with 3-40 variables and SR100 for which the training examples contain 10-100 variables. For 3-SAT, the training samples contain 10-100 variables (3SAT100). The SR dataset is well suited for training SAT/UNSAT prediction models due to its design that prevents learning from superficial features, making it harder for models to exploit statistical shortcuts rather than learning true logical reasoning. We also create versions of training data which contain only satisfiable instances (denoted SAT only). The size of these datasets is half of the original datasets (i.e. 25000 examples). To evaluate generalization, we validate exclusively on problems with exactly the maximum number of variables in each category, therefore SR40 for evaluation means that the problems have always exactly 40 variables (not a range of 3-40), SR100 test contains only problems with exactly 100 variables, and so on.¹

Table 1 summarizes the key statistics of our evaluation datasets.

Table 1: Statistics of benchmark test sets. SAT% indicates the percentage of satisfiable instances in each dataset. Avg. Gap represents the average number of unsatisfied clauses when using random variable assignments. SAT Gap and UNSAT Gap show this metric separated by instance satisfiability. SR datasets are generated using the SR generator with the indicated number of variables (e.g., SR40 contains instances with 40 variables), while 3SAT datasets contain instances near the phase transition point with the specified number of variables. All datasets maintain a balanced distribution of satisfiable and unsatisfiable instances.

Dataset	SAT%	Avg. Gap	SAT Gap	UNSAT Gap	Avg. Clauses
SR40	50.0%	21.29	21.59	20.99	228.40
SR100	50.0%	51.31	50.64	51.98	547.49
SR200	50.0%	100.31	101.03	99.59	1083.81
SR400	50.0%	198.74	198.53	198.95	2152.32
3SAT100	53.5%	52.78	53.00	52.54	426.00
3SAT200	55.5%	107.65	107.45	107.90	852.00

The Gap metric represents the average number of unsatisfied clauses when using random

¹This is consistent with the original experiments by Selsam et al. but different from the experiments reported by Li et al. [2023] where SR40 in evaluation contains problems from the same distribution as the training set and therefore they report a better performance because smaller problems are easier to solve.

variable assignments. This metric has the same definition for both SAT and UNSAT instances; it simply counts how many clauses remain unsatisfied with random assignments on average. Larger gaps indicate more challenging problems where random guessing performs poorly.

5.2 Quantitative Evaluation

We conducted a comprehensive evaluation that compares different architectural choices and supervision methods. Our evaluation focuses on five key performance metrics:

- **Average Gap:** The average number of unsatisfied clauses across all test instances. Lower values indicate better performance, with 0 representing perfect satisfaction (i.e., no unsatisfied clauses) on satisfiable instances. For unsatisfiable instances, this metric reflects how close the model gets to minimizing unsatisfied clauses.
- **Gap on SAT:** The average number of unsatisfied clauses computed only over satisfiable instances.
- **Gap on UNSAT:** The average number of unsatisfied clauses computed only over unsatisfiable instances.
- **SAT Accuracy:** The percentage of satisfiable instances for which the model correctly finds a satisfying assignment, computed only over satisfiable instances.
- **Decision Accuracy:** The percentage of instances for which the model correctly predicts whether the formula is satisfiable. Since our approach does not formally refute unsatisfiable instances, we classify an instance as unsatisfiable when the model fails to find a satisfying assignment. This means unsatisfiable instances are always classified correctly under this assumption. This applies specifically in the case of assignment-based evaluation.

5.2.1 Comparison of Graph Representations, Update Functions and Training Methods

Table 2 presents a comprehensive comparison of different architectural configurations trained exclusively on the SR40 dataset. This comparison includes different graph representations (Literal-Clause Graph vs. Variable-Clause Graph), update functions (RNN vs. LSTM), and supervision approaches (SAT/UNSAT classification, assignment supervision, and unsupervised objective training), all evaluated on instances with 40 variables. All models were evaluated using Exponential Moving Average (EMA) of parameters during validation only, as detailed in A.1.2, which helps reduce fluctuations in validation metrics and provide more reliable model selection. Importantly, curriculum learning (A.1.1) was employed only for training models with SAT/UNSAT classification objectives, as it proved unnecessary for models trained with assignment prediction or unsupervised learning approaches.

Graph Representation Impact: Our results demonstrate that Literal-Clause Graph (LCG) and Variable-Clause Graph (VCG) representations exhibit different strengths. VCG shows better performance for assignment-based training with RNN updates, achieving a SAT accuracy of 68.8% compared to 48.6% for LCG. Additionally, VCG’s more compact representation (using one node per variable rather than two for positive and negative literals) provides computational advantages for larger formulas, making it our preferred choice for scaling to more complex problems.

Message Passing Mechanism: While LSTM-based message passing shows advantages in some configurations, particularly for unsupervised training, we found that RNN-based approaches offer a better balance of performance and interpretability for assignment-based training. RNN updates with VCG representation achieved higher results for finding satisfying assignments, with 68.8% SAT accuracy and 84.4% decision accuracy. The simpler RNN structure also facilitates better analysis of the model’s internal reasoning process. However, we found training RNN-based models for SAT/UNSAT classification particularly challenging, with LSTM being more stable for this specific task.

Supervision Approach: Our experiments reveal distinct advantages for different supervision approaches:

1. **Assignment-based supervision** shows better performance for finding satisfying assignments, especially with VCG+RNN configuration (68.8% SAT accuracy, 84.4% decision accuracy).
2. **Unsupervised learning** achieves the lowest average gaps across configurations (as low as 0.91 for VCG+RNN and 0.84 for VCG+LSTM). This makes unsupervised training useful for applications where minimizing unsatisfied clauses is the priority.
3. **SAT/UNSAT classification** training, while challenging with RNN, enables an interesting property: models trained only for classification develop an implicit ability to separate embeddings for positive and negative literals. This separation allows for retrieving satisfying assignments through clustering techniques, despite the model not being explicitly trained for assignment prediction.

Based on the results reported in Table 2, we identify the VCG+RNN+Assignment configuration as our most effective approach, offering a good balance between assignment accuracy and computational efficiency. This configuration forms the foundation for our further experiments and analysis in subsequent sections.

Assignment Training Refinements: Table 3 highlights the impact of a novel training method we introduce, here called “closest assignment”, with the VCG+RNN configuration across multiple datasets. This method computes assignments that minimize Hamming distance to the model’s current predictions, showing improvements over training with precalculated assignments, especially for formulas with more variables. For SR100, using the closest assignment approach reduces the average gap from 3.81 to 1.43 for SAT+UNSAT training and improves SAT accuracy from 44.8% to 53.2%.

This improvement correlates with the number of possible solutions in the benchmarks (SR100 has a median of 16 solutions per formula compared to SR3-40’s median of 7), supporting our hypothesis that for formulas with larger solution spaces, guiding the model with dynamically selected assignments that align with its current predictions yields better generalization than using fixed predetermined assignments.

The computational challenges of calculating closest assignments during training are noteworthy, particularly for larger benchmarks like 3SAT+UNSAT, where this approach became impractical and we therefore omit this experiment and leave the last row of Table 3 empty. It also highlights an opportunity for future work on more efficient approximation methods for finding near-optimal assignments.

Training Data Composition: Our results also indicate that training exclusively on SAT instances (SAT only) improves performance for finding satisfying assignments. For SR40, this approach with closest assignment training achieves our highest SAT accuracy of 76% and decision accuracy of 88%. However, models trained on both SAT and UNSAT instances (SAT+UNSAT)

with closest assignment supervision demonstrate better gap minimization, achieving an average gap of 0.98 versus 2.68 for SAT-only training on SR40.

Table 2: Performance comparison of GNN architectures for SAT solving on the SR40 dataset. The table compares Literal-Clause Graph (LCG) and Variable-Clause Graph (VCG) representations, RNN and LSTM update mechanisms, and different training objectives. Metrics include average gap (number of unsatisfied clauses) across all instances and separated by satisfiability status (lower is better), SAT accuracy (percentage of satisfiable instances solved by finding assignment), and decision accuracy (percentage of correct satisfiability predictions). Notable findings include: unsupervised training consistently achieves lowest gaps; VCG+RNN with assignment prediction shows highest SAT accuracy (68.8%); and RNN-based models with SAT/UNSAT classification proved challenging to train effectively (indicated by dashes). Asterisks (*) indicate results obtained through clustering of node embeddings rather than direct prediction. This model combination was particularly hard to train in our setup. We found that both for VCG and LCG RNN is very sensitive to hyper-parameter selection. As the model failed to get generalized in our final unified experimental setup we do not include this result (close to random performance) now.

Graph	Update	Loss Function	Avg. Gap ↓	Gap on SAT ↓	Gap on UNSAT ↓	SAT Acc. ↑	Dec. Acc. ↑
LCG	RNN	SAT/UNSAT	—	—	—	—	—
		Assignment	1.83	1.25	2.41	48.6 %	72.8 %
		Unsup	0.93	0.59	1.26	51.4 %	75.7 %
	LSTM	SAT/UNSAT	1.96*	1.27*	2.62*	59.2* %	83.9 / 79.6* %
		Assignment	1.82	1.06	2.58	56.8 %	78.4 %
		Unsup	0.81	0.45	1.16	62 %	81 %
VCG	RNN	SAT/UNSAT	3.62*	1.9*	5.34*	56.6* %	80 / 78.3* %
		Assignment	1.95	0.8	3.05	68.8 %	84.4 %
		Unsup	0.91	0.58	1.23	51.6 %	75.8 %
	LSTM	SAT/UNSAT	2.33*	1.57	3.08	52.2* %	81.9 / 76.1* %
		Assignment	2.05	0.96	3.14	66.4 %	83.2 %
		Unsup	0.84	0.51	1.17	56.4 %	78.2 %

5.3 Test-time Scaling

A key property of our recurrent GNN architecture for SAT solving is the ability to adjust computational effort at inference time. Unlike standard GNNs with fixed number of layers, the weight-shared recurrent design enables flexible scaling through additional iterations and resampling.

5.3.1 Iteration and Resampling Effects

Figure 2 demonstrates how increasing message-passing iterations improves the percentage of solved SAT instances. Similarly, Figure 3 shows how the average gap decreases across iterations for various benchmarks. The heat maps in Figure 4 provide a comprehensive view of how performance metrics improve with both increased iterations and resampling attempts.

For the model trained on SR40, several observations are notable:

- **Iteration benefits:** Increasing iterations from 25 to 125 consistently improves all metrics across benchmarks.
- **Resampling effects:** Multiple inference attempts with different random initializations of node feature vectors further enhance performance. For SR40, decision accuracy improves from 84% with one sample to 93% with five samples at 125 iterations.

Table 3: Performance analysis of VCG+RNN with assignment prediction across different datasets and training methodologies. Our novel "Closest" supervision method (which dynamically selects assignments closest to current model predictions) consistently outperforms training with precalculated assignments. For SR40, SAT-only training with closest assignment supervision achieves the highest SAT accuracy (76%), while SAT+UNSAT training with closest assignment supervision yields the lowest average gap (0.98). The missing data for 3SAT100 with SAT+UNSAT closest supervision is due to prohibitive computational costs. Bold values indicate best results per dataset.

Dataset	Training Mode	Assignment Type	Avg. Gap ↓	Gap on SAT ↓	Gap on UNSAT ↓	SAT Acc. ↑	Dec. Acc. ↑
SR40	SAT only	Precalculated	2.93	1.11	4.75	68.2 %	84.1 %
	SAT only	Closest	2.68	0.88	4.48	76 %	88 %
	SAT+UNSAT	Precalculated	1.95	0.8	3.05	68.8 %	84.4 %
	SAT+UNSAT	Closest	0.98	0.48	1.49	71.2 %	85.6 %
SR100	SAT only	Precalculated	4.42	2.36	6.48	47.4 %	73.7 %
	SAT only	Closest	3.57	1.67	5.48	59.6 %	79.8 %
	SAT+UNSAT	Precalculated	3.81	2.34	5.28	44.8 %	72.4 %
	SAT+UNSAT	Closest	1.43	0.92	1.94	53.2 %	76.6 %
3SAT100	SAT only	Precalculated	5.93	3.40	9.27	25.7 %	57.2 %
	SAT only	Closest	5.23	2.33	9.11	48.4 %	70 %
	SAT+UNSAT	Precalculated	4.22	2.84	6.00	23.9 %	55.8 %
	SAT+UNSAT	Closest	—	—	—	—	—

- **Cross-distribution applicability:** The model trained on SR40 maintains reasonable effectiveness on SR100 and 3SAT100, though with expected performance decrease. This aligns with findings from Li et al. [2023], who demonstrated that models trained on SR distributions generally transfer well to other SAT problem structures.

5.3.2 Train-time vs Test-time Scaling

Tables 4 and 5 present the performance of models trained on SR40 and SR100 distributions when evaluated across benchmarks of varying sizes. The SR40-trained model achieves reasonable generalization to larger instances, though with decreasing effectiveness as problem size increases. For SR100, the model achieves 74.2% decision accuracy despite being trained on smaller instances, showing good generalization capabilities.

The SR100-trained model demonstrates better performance on larger instances compared to the SR40-trained model, as expected. On SR200, it achieves 83.0% decision accuracy compared to 58.5% for the SR40 model. This suggests that while test-time scaling can improve performance on larger problems, there are limits to this approach, and training models on larger instances might be necessary for optimal performance on very large problems.

These results highlight that recurrent GNN architectures allow for a flexible computation-performance tradeoff that can be adjusted at inference time based on available computational resources and desired solution quality.

5.4 Diffusion Model Extension

As we mentioned in Section 3.4.4, one can use the GNN as a function approximator $f_\theta(\cdot)$ inside a diffusion model. This enables another way of scaling the test-time compute. We adapt the diffusion model used by Sun et al. [2023] where the function approximator is trained to predict the ground truth solution $\mathbf{x}_0 = f_\theta(\mathbf{x}_t, t)$ conditioned on a sample \mathbf{x}_t at time t . The predicted assignment is then used to obtain a sample at time $t - 1$ and this process is repeated again $\mathbf{x}_0 = f_\theta(\mathbf{x}_{t-1}, t - 1)$ until we reach $t = 0$. One application of the function approximator together with the sampling is called a diffusion step. The number of diffusion steps T used for inference is a parameter which can be adapted after the model was already

Table 4: Performance of a model trained on SR40 (VCG+RNN with closest assignment supervision) when tested across various benchmarks with a maximum of 100 message-passing iterations and early stopping. The model maintains reasonable performance on SR100 (74.2% decision accuracy) but degrades on larger instances. "UNSAT Instances (gap == 1)" shows the percentage of UNSAT instances where the model achieved a gap of 1, which is always optimal for SR datasets but not always achievable for 3SAT instances. "Steps" columns indicate average/median iterations required to reach solutions, demonstrating the model's efficiency.

Dataset	Decision Accuracy	SAT Instances Solved	UNSAT Instances (gap == 1)	SAT Steps (Avg/Med)	UNSAT Steps (Avg/Med)
SR40	89.5%	79.0%	95.6%	16.17/13.0	13.33/10.0
SR100	74.2%	48.3%	91.3%	24.93/21.0	23.47/19.0
SR200	58.5%	17.0%	64.5%	32.44/30.0	33.98/28.0
SR400	51.5%	3.0%	16.0%	41.33/34.0	52.06/44.5
3SAT100	74.8%	52.8%	64.0%	25.63/22.0	29.66/24.0
3SAT200	54.0%	17.1%	22.5%	33.21/25.0	35.10/31.5

Table 5: Performance of a model trained on SR100 (VCG+RNN with closest assignment supervision) when tested on SR benchmarks with a maximum of 100 message-passing iterations and early stopping. Given that the SR40-trained model achieved only 3% SAT accuracy on SR400 (see Table 4), we focused on evaluating how training on larger instances improves scaling. The results show dramatic improvement on larger benchmarks (36.5% vs 3% on SR400), demonstrating that training on larger problems significantly enhances generalization capacity. The "Steps" metrics confirm the SR100-trained model requires fewer iterations on larger problems (e.g., 30.68 vs 41.33 average iterations for SAT instances on SR400).

Dataset	Decision Accuracy	SAT Instances Solved	UNSAT Instances (gap == 1)	SAT Steps (Avg/Med)	UNSAT Steps (Avg/Med)
SR40	90.6%	81.2%	90.0%	14.57/12.0	16.21/12.0
SR100	79.3%	58.7%	85.7%	22.16/18.0	22.64/19.0
SR200	83.0%	68.2%	60.2%	22.29/20.0	27.12/22.0
SR400	68.2%	36.5%	78.0%	30.68/26.0	33.13/28.0

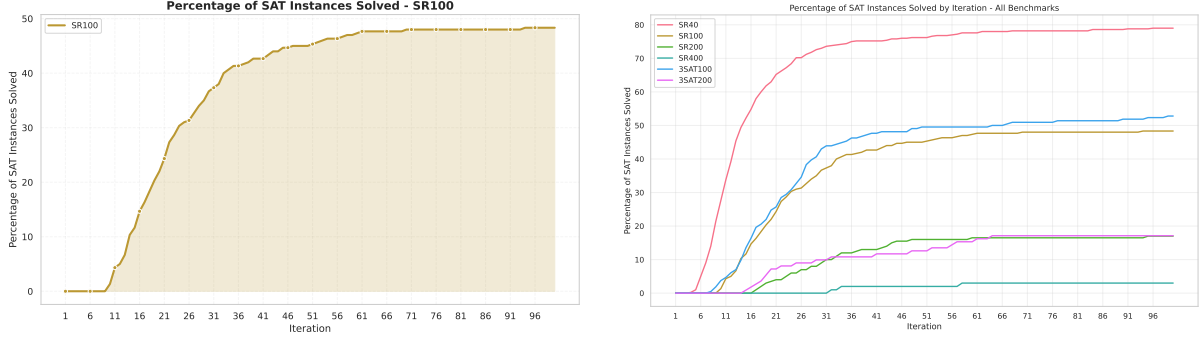


Figure 2: Percentage of SAT instances solved as message passing iterations increase for a model trained on SR40 with SAT+UNSAT closest assignment supervision. Left: Performance on SR100, showing rapid initial improvement. Right: Comparison across benchmarks, demonstrating effectiveness decreases with problem size but benefits from additional iterations, highlighting the recurrent architecture’s inference-time scaling capability.

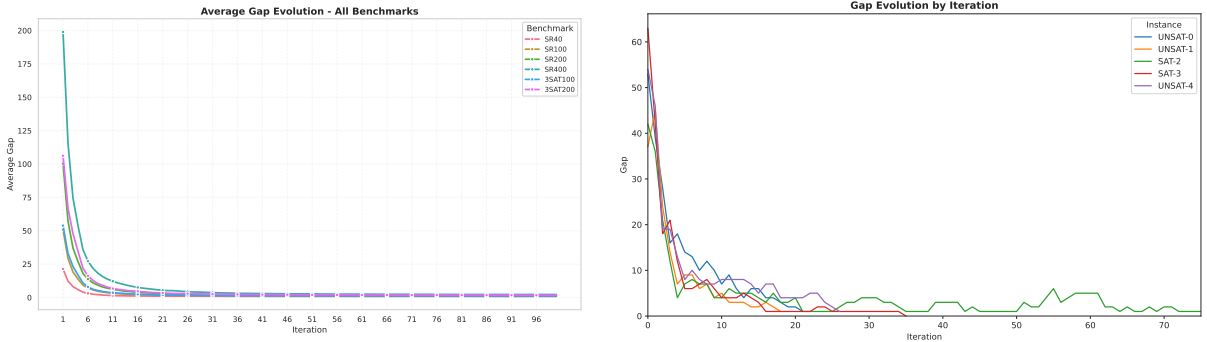


Figure 3: Average gap (unsatisfied clauses) reduction with increasing message passing iterations for a model trained on SR40 with SAT+UNSAT closest assignment supervision. Left: Comparison across benchmarks showing extremely rapid gap reduction in early iterations for all problem sizes, with all benchmarks achieving remarkably low average gaps despite varying SAT-solving performance. Right: Individual instance trajectories revealing different convergence patterns between SAT and UNSAT instances, with occasional fluctuations suggesting potential benefit from monitoring solution quality during inference.

trained and therefore, in this setting we have two types of iterations. One is the number of message-passing iterations and the second is the number of diffusion steps. In Table 6 we report the tradeoff between the number of message-passing steps (referred to as **GNN_Steps**) and the number of diffusion steps. The reported numbers correspond to the dataset **SR100** with 100 variables in each problem. The model was trained on the **SR40** distribution and the tested combinations use around 300 iterations in total distributed between the two types of steps.

The experiments revealed a consistent trend: increasing the number of message-passing steps is generally more important for improving metrics such as Accuracy and Avg. Gap.

5.4.1 Connection to Assignment Prediction Training

We also report an interesting finding which allows to simplify the function approximator used in the diffusion model. Notice that in the expression $\mathbf{x}_0 = f_\theta(\mathbf{x}_t, t)$ it is also conditioned on the timestep t . This conditioning is dictated by the theory of diffusion models Nakkiran et al. [2024] and most of the models, including the one by Sun et al. Sun and Yang [2023] blindly follow this design choice. In our experiments, we found out that this conditioning is not needed and that the model sometimes works even better without it. Therefore, in all reported results, the

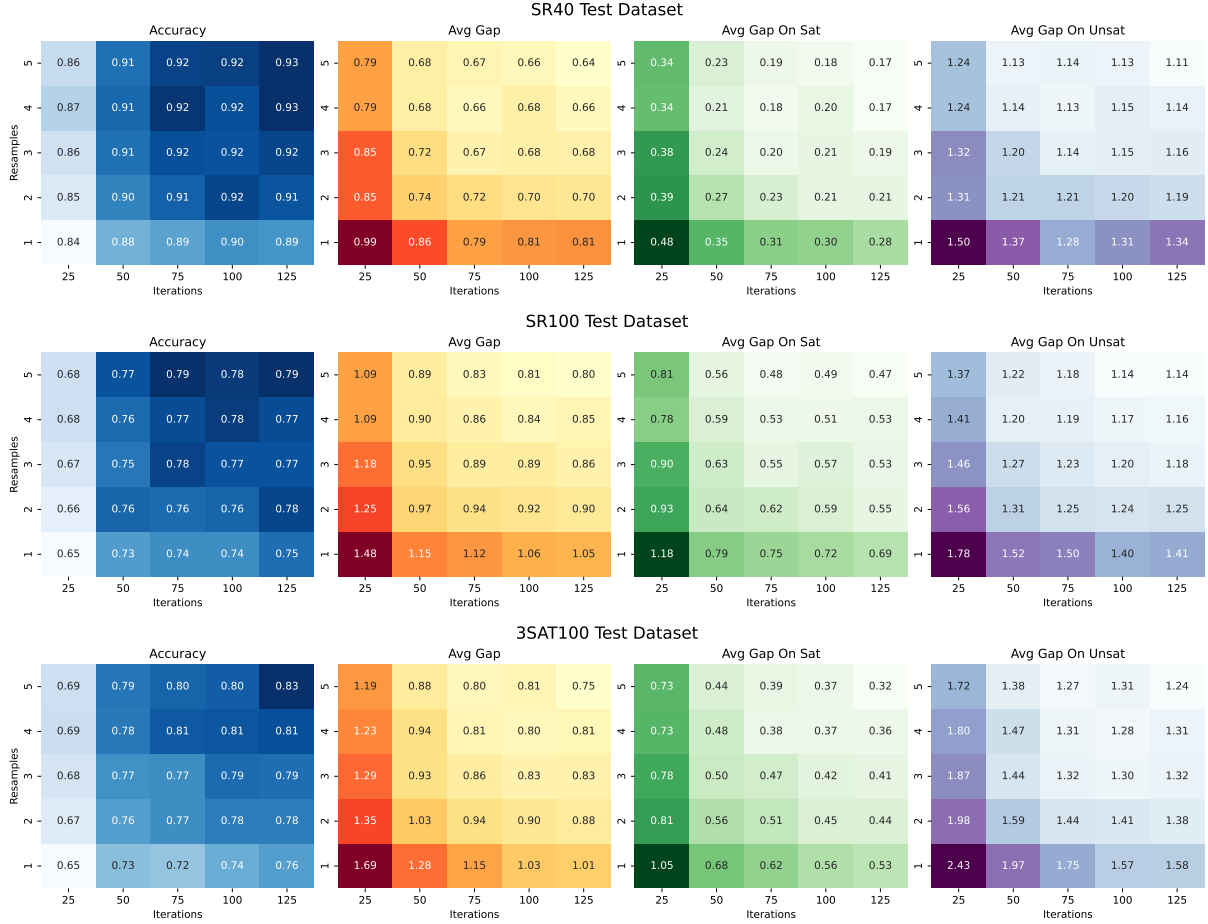


Figure 4: Performance heatmaps for a model trained on SR40 with SAT+UNSAT closest assignment supervision, showing how metrics improve with both increased iterations (columns) and resampling attempts (rows). Testing on SR40 (top), SR100 (middle), and 3SAT100 (bottom) demonstrates significant gains from both scaling dimensions—e.g., SR40 decision accuracy improves from 84% (1 sample, 25 iterations) to 93% (5 samples, 125 iterations). This two-dimensional inference-time scaling capability is consistent across benchmarks but with decreasing returns on larger problems.

model is trained to predict the solution \mathbf{x}_0 only from the sample at timestep t ($\mathbf{x}_0 = f_\theta(\mathbf{x}_t)$). Concurrently to us, this fact was also discovered by Sun et al. Sun et al. [2025] (the same surname is a coincidence) and it’s possible that many of the reported experiments which blindly use this conditioning would result in better values without it.

In this simplified setup, the training examples $(\mathbf{x}_0, \mathbf{x}_t)$ are sampled by taking a solution of a formula (\mathbf{x}_0), sampling a random t from the diffusion schedule and obtaining a corrupted version of the solution at time t (\mathbf{x}_t). The model is trained to predict \mathbf{x}_0 from \mathbf{x}_t . The GNN is the same as in the case of *assignment prediction* except that it also contains a learnable embedding layer which embeds the Boolean values in the assignment \mathbf{x}_t into a vector space to obtain the initial embeddings of variables (or literals) for the first pass of message-passing.

The only difference from the model trained for *assignment prediction* is therefore that the initial embeddings are not sampled randomly but obtained by embedding the perturbed assignment \mathbf{x}_t . This also means that during test time, these two approaches differ only by rounding, i.e. running the model trained for *assignment prediction* for 100 steps and after every 20 steps rounding the variable embeddings to vectors representing *True* and *False* is same as running the diffusion model for 5 diffusion steps where each step has 20 message-passing iterations.

Table 6: Performance Metrics for Different GNN and Diffusion Step Configurations. Variable names shown in parentheses in the original data source are omitted here for brevity.

GNN Steps	Diffusion Steps	Avg. Gap	Dec. Acc. (%)	Single-Step Avg. Gap	Single-Step Dec. Acc (%)
20	15	1.09	68.7	3.26	60.2
23	13	1.05	70.4	2.80	63.1
27	11	0.96	73.6	2.44	64.7
31	9	0.98	73.8	2.11	68.2
35	8	0.93	75.4	1.96	69.5
38	7	0.92	76.3	1.83	70.2
42	7	0.90	76.7	1.70	71.6
46	6	0.95	76.8	1.64	72.1
50	6	0.94	76.2	1.52	73.0

5.4.2 Interleaving Diffusion Steps with Unit Propagation

The fact that for each diffusion step, the model outputs probabilities for two possible values, allows us to obtain a partial solution and then run a unit propagation to deduce assignment to other variables. The partial assignment can be obtained by fixing a threshold and then assigning only variables for which one of the values has a predicted probability higher than this threshold. The lower the threshold, the more variables will be fixed and the higher the probability that it will not be possible to complete the partial assignment to a satisfiable assignment.

We therefore design a tree-search-like algorithm which first tries a low threshold in each diffusion step and if it does not find a satisfiable assignment it backtracks and increases the threshold to obtain a new partial assignment. The details of this algorithm are described in A.3 and the experimental results are reported in Table 7. As can be seen, interleaving the diffusion steps with unit propagation results in additional improvements over the base diffusion model (approximately 10%). We explicitly mention that this experiment is provided only to show a possible avenue for further improvements and the algorithm in its current form is not optimized for speed.

Table 7: Performance with Unit Propagation. Here we compare the performance with (U.P. Acc.) and without (Dec. Acc.) Unit Propagation, and report the computational cost of Unit Propagation, listing the average number of total recursive function calls, the average number of recursive calls in solved problems, and the average number of recursive calls in unsolved problems.

Problems	Dec. Acc. (%)	U.P. Acc. (%)	Total Rec. Calls	Solved Rec. Calls	Unsolved Rec. Calls
<i>SR40</i>	88.4	94.2	32.864	6.701	53.546
<i>SR50</i>	86.6	93.7	29.539	6.995	47.038
<i>SR60</i>	83.3	92.2	26.414	7.526	40.204
<i>SR70</i>	79.5	89.5	24.162	6.752	35.505
<i>SR80</i>	77.6	88.0	22.604	6.917	32.219
<i>SR90</i>	74.0	85.1	22.140	7.274	30.151
<i>SR100</i>	73.4	83.7	20.363	7.129	27.074
<i>SR150</i>	63.2	75.1	17.388	7.828	20.592
<i>SR200</i>	58.0	67.5	16.270	8.710	17.868

6 Interpreting the Trained Model

6.1 Embedding Space Analysis

Our analysis of variable embeddings reveals patterns that explain how GNNs learn to solve SAT problems. When visualizing these embeddings using dimensionality reduction McInnes et al. [2018], we observe that they form distinct clusters corresponding to optimal variable assignments.

As shown in Figure 5, variable embeddings start randomly distributed but gradually organize into two clusters through message passing iterations. By applying k-means clustering ($k = 2$) to these embeddings, we can recover variable assignments that approximate optimal solutions, even from networks trained only to predict satisfiability status.

6.2 Iterative Optimization Behavior

By tracking clause satisfaction across iterations, we observe that GNNs solve SAT problems through progressive local refinement. The gap (number of unsatisfied clauses) decreases following a trajectory typical of iterative optimization methods: rapid initial improvement followed by gradual refinement.

This behavior supports the interpretation that GNNs implicitly learn to perform continuous optimization in a high-dimensional space similar to SDP relaxations for SAT. The effectiveness of additional message passing iterations during inference further strengthens this connection. A difference from the SDP relaxation is that the objective function which the GNN implicitly optimizes is non-convex because we observed that it can get stuck in local optima or converge to different solutions when initialized multiple times by different random embeddings.

Figure 3 illustrates how the average gap decreases with increasing iterations. The trajectory suggests a rapid improvement phase followed by more gradual refinement. Individual instance trajectories reveal that while most instances show steady improvement toward optimal solutions, some exhibit fluctuations, particularly unsatisfiable instances. This observation supports the potential value of early stopping techniques, as in rare cases, the gap at later iterations might be higher than a previously achieved minimum gap.

The bi-level optimization perspective—where message passing performs an inner optimization loop (finding variable assignments) guided by network parameters optimized at the outer level (during training)—helps explain the network’s ability to generalize to novel problem instances and larger problems than those seen during training. In Section 7, we discuss more details about a possibility of manual derivation of the GNN equations from an explicit objective function.

7 Discussion

In this section, we discuss the limitations of our work along with an outlook for future research. The primary limitation of the methods presented here is that they are not competitive with state-of-the-art SAT solvers on benchmarks derived from real-world problems. Current SAT solvers can handle formulas with millions of variables, which is not feasible for the GNN in its current form. However, as mentioned in the introduction, our motivation for studying these models is to better understand the reasoning capabilities of neural networks in a simplified context.

The test-time scaling experiments clearly demonstrate that the GNNs can successfully generalize beyond their training distribution and do not merely learn superficial statistical patterns. The qualitative results presented in Section 6 further suggest that it is possible to fully understand the mechanisms by which the GNN solves a given formula. Figure 3 illustrates that the trained GNN functions as an implicit MaxSAT solver, incrementally maximizing the number

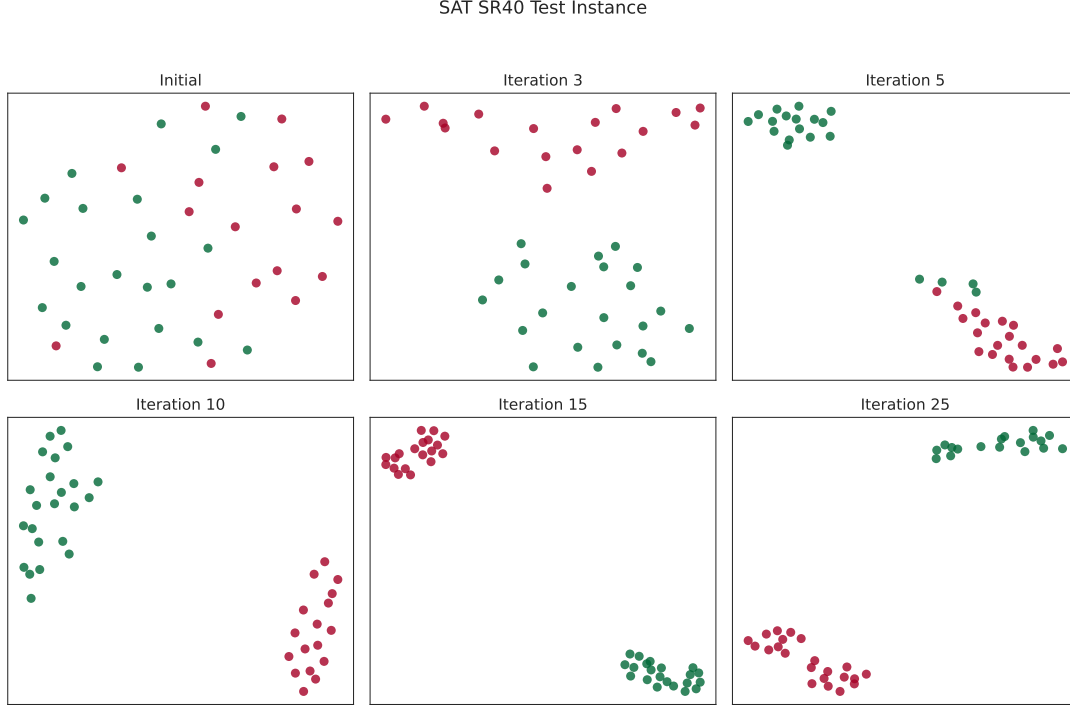


Figure 5: Evolution of variable embeddings during message passing iterations for a satisfiable SR40 instance. The visualization shows 2D projections at different stages (Initial through Iteration 25), colored k-means algorithm in each iteration (green/red). Initially random, embeddings gradually organize into two distinct clusters often corresponding to optimal variable assignments. This clustering behavior was observed across different model architectures and training objectives—notably, even models trained solely for SAT/UNSAT classification (without explicit assignment supervision) develop this embedding separation. This phenomenon supports our interpretation that GNNs implicitly perform continuous optimization similar to SDP relaxation for SAT problems.

of satisfied clauses at each step. These local updates occur in continuous space and can therefore be viewed as gradient updates with respect to an implicit objective function measuring clause satisfaction. Variables are also represented in a high-dimensional vector space, similar to semi-definite programming as explained in B.

From this perspective, Equations 6, 7, and 11 can be interpreted as a gradient descent algorithm searching for an optimal assignment over a high-dimensional unit sphere (due to unit normalization), while the final classification layer corresponds to a rounding step to Boolean values. In future work, we aim to manually derive these equations from a trained GNN using a primal-dual approach, interleaving gradient updates of primal and dual variables associated with constraints. We believe that by utilizing suitable proximal operators and an appropriate metric in the relaxed solution space, the GNN can be effectively interpreted as a primal-dual algorithm optimizing a continuous relaxation of the MaxSAT objective in a high-dimensional space. This points out to another major advantage of using the RNN update function because its simple form is suitable for such derivation.

Deriving equations for such algorithms applicable to arbitrary combinatorial optimization problems would be highly beneficial in practice, allowing these equations to be parameterized by learnable matrices and fine-tuned for specific problem distributions. Such data-driven solvers would be analogous to physics-informed neural networks Cai et al. [2021], where substantial domain knowledge is embedded within the model, followed by fine-tuning to approximate the dynamics of a particular physical system. This approach results in fast numerical solvers tailored to specific domains. We believe that the development of data-driven numerical solvers represents an exciting future direction for combinatorial optimization research. To make these numerical solvers practical, it will still be necessary to integrate them into more complex systems, where they would function as guessing or bounding heuristic.

Another limitation of our work is that the model was tested exclusively on random problems. This decision is justified by the findings of Li et al. [2023], who demonstrated that models trained on random problem instances exhibit superior generalization to other distributions. Since Li et al. already provided experimental results demonstrating the transferability of models across different problem distributions, we chose not to repeat those experiments here.

8 Conclusion

This work provides a comprehensive analysis of graph neural networks for Boolean satisfiability problems. Our evaluation identified key design choices that enhance performance: variable-clause graph representation with RNN updates offers an effective balance of accuracy and efficiency, while our novel closest assignment supervision method significantly improves performance on problems with large solution spaces. The recurrent architecture enables flexible scaling during inference through additional message-passing iterations and resampling. Our diffusion model extension demonstrates another approach to inference-time adaptation, with further improvements possible by integrating classical techniques like unit propagation.

Our analysis of embedding space patterns and optimization trajectories supports the interpretation that these models implicitly implement continuous relaxation algorithms for MaxSAT, explaining their ability to generalize to novel problem instances. This connection provides a theoretical framework for understanding neural reasoning capabilities in structured domains, with implications for the design of hybrid solving approaches.

References

Saeed Amizadeh, Sergiy Matushevych, and Markus Weimer. Learning to solve circuit-sat: An unsupervised differentiable approach. In *International conference on learning representations*, 2018.

- Jacob Austin, Daniel D Johnson, Jonathan Ho, Daniel Tarlow, and Rianne Van Den Berg. Structured denoising diffusion models in discrete state-spaces. *Advances in neural information processing systems*, 34:17981–17993, 2021.
- Armin Biere, Marijn Heule, and Hans van Maaren. *Handbook of satisfiability*, volume 185. IOS press, 2009.
- Sally C Brailsford, Chris N Potts, and Barbara M Smith. Constraint satisfaction problems: Algorithms and applications. *European journal of operational research*, 119(3):557–581, 1999.
- Chen Cai, Truong Son Hy, Rose Yu, and Yusu Wang. On the connection between mpnn and graph transformer. In *International conference on machine learning*, pages 3408–3430. PMLR, 2023.
- Shengze Cai, Zhiping Mao, Zhicheng Wang, Minglang Yin, and George Em Karniadakis. Physics-informed neural networks (pinns) for fluid mechanics: A review. *Acta Mechanica Sinica*, 37(12):1727–1738, 2021.
- James M Crawford and Larry D Auton. Experimental results on the crossover point in random 3-sat. *Artificial intelligence*, 81(1-2):31–57, 1996.
- Jerry A Fodor and Zenon W Pylyshyn. Connectionism and cognitive architecture: A critical analysis. *Cognition*, 28(1-2):3–71, 1988.
- Zhaohui Fu and Sharad Malik. On solving the partial max-sat problem. In *International Conference on Theory and Applications of Satisfiability Testing*, pages 252–265. Springer, 2006.
- Bernd Gärtner and Jiri Matousek. *Approximation algorithms and semidefinite programming*. Springer Science & Business Media, 2012.
- Michel X Goemans and David P Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM (JACM)*, 42(6):1115–1145, 1995.
- Daya Guo, Dejian Yang, Haowei Zhang, Junxiao Song, Ruoyu Zhang, Runxin Xu, Qihao Zhu, Shirong Ma, Peiyi Wang, Xiao Bi, et al. Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning. *arXiv preprint arXiv:2501.12948*, 2025.
- Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33:6840–6851, 2020.
- Abdelrahman Hosny and Sherief Reda. torchmsat: A gpu-accelerated approximation to the maximum satisfiability problem. *arXiv preprint arXiv:2402.03640*, 2024.
- Jan Hůla, David Mojžíšek, and Mikoláš Janota. Understanding gnns for boolean satisfiability through approximation algorithms. In *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management*, pages 953–961, 2024.
- Aaron Jaech, Adam Kalai, Adam Lerer, Adam Richardson, Ahmed El-Kishky, Aiden Low, Alec Helyar, Aleksander Madry, Alex Beutel, Alex Carney, et al. Openai o1 system card. *arXiv preprint arXiv:2412.16720*, 2024.
- Anastasios Kyrillidis, Anshumali Shrivastava, Moshe Vardi, and Zhiwei Zhang. Fouriersat: A fourier expansion-based algebraic framework for solving hybrid boolean constraints. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 1552–1560, 2020.

- Zhaoyu Li and Xujie Si. Nsnet: A general neural probabilistic framework for satisfiability problems. *Advances in Neural Information Processing Systems*, 35:25573–25585, 2022.
- Zhaoyu Li, Jinpei Guo, and Xujie Si. G4satbench: Benchmarking and advancing sat solving with graph neural networks. *arXiv preprint arXiv:2309.16941*, 2023.
- Gary Marcus. Deep learning: A critical appraisal. *arXiv preprint arXiv:1801.00631*, 2018.
- Gary F Marcus. *The algebraic mind: Integrating connectionism and cognitive science*. MIT press, 2003.
- Leland McInnes, John Healy, and James Melville. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*, 2018.
- Preetum Nakkiran, Arwen Bradley, Hattie Zhou, and Madhu Advani. Step-by-step diffusion: An elementary tutorial. *arXiv preprint arXiv:2406.08929*, 2024.
- Emils Ozolins, Karlis Freivalds, Andis Draguns, Eliza Gaile, Ronalds Zakovskis, and Sergejs Kozlovics. Goal-aware neural sat solver. In *2022 International joint conference on neural networks (IJCNN)*, pages 1–8. IEEE, 2022.
- Motakuri Ramana and Alan J Goldman. Some geometric results in semidefinite programming. *Journal of Global Optimization*, 7(1):33–50, 1995.
- Bart Selman, Henry A Kautz, Bram Cohen, et al. Local search strategies for satisfiability testing. *Cliques, coloring, and satisfiability*, 26:521–532, 1993.
- Daniel Selsam and Nikolaj Bjørner. Guiding high-performance sat solvers with unsat-core predictions. In *Theory and Applications of Satisfiability Testing–SAT 2019: 22nd International Conference, SAT 2019, Lisbon, Portugal, July 9–12, 2019, Proceedings 22*, pages 336–353. Springer, 2019.
- Daniel Selsam, Matthew Lamm, Benedikt Bünz, Percy Liang, Leonardo de Moura, and David L Dill. Learning a sat solver from single-bit supervision. *arXiv preprint arXiv:1802.03685*, 2018.
- Qiao Sun, Zhicheng Jiang, Hanhong Zhao, and Kaiming He. Is noise conditioning necessary for denoising generative models? *arXiv preprint arXiv:2502.13129*, 2025.
- Zhiqing Sun and Yiming Yang. Difusco: Graph-based diffusion solvers for combinatorial optimization. *Advances in neural information processing systems*, 36:3706–3731, 2023.
- Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.
- W Wang, Y Hu, M Tiwari, S Khurshid, KL McMillan, and R Miikkulainen. Neurocomb: improving sat solving with graph neural networks, 2021.
- David Warde-Farley, Vinod Nair, Yujia Li, Ivan Lobov, Felix Gimeno, and Simon Osindero. Solving maxsat with matrix multiplication. *arXiv preprint arXiv:2311.02101*, 2023.
- Morris Yau, Nikolaos Karalias, Eric Lu, Jessica Xu, and Stefanie Jegelka. Are graph neural networks optimal approximation algorithms? *Advances in Neural Information Processing Systems*, 37:73124–73181, 2024.
- Emre Yolcu and Barnabás Póczos. Learning local search heuristics for boolean satisfiability. *Advances in Neural Information Processing Systems*, 32, 2019.

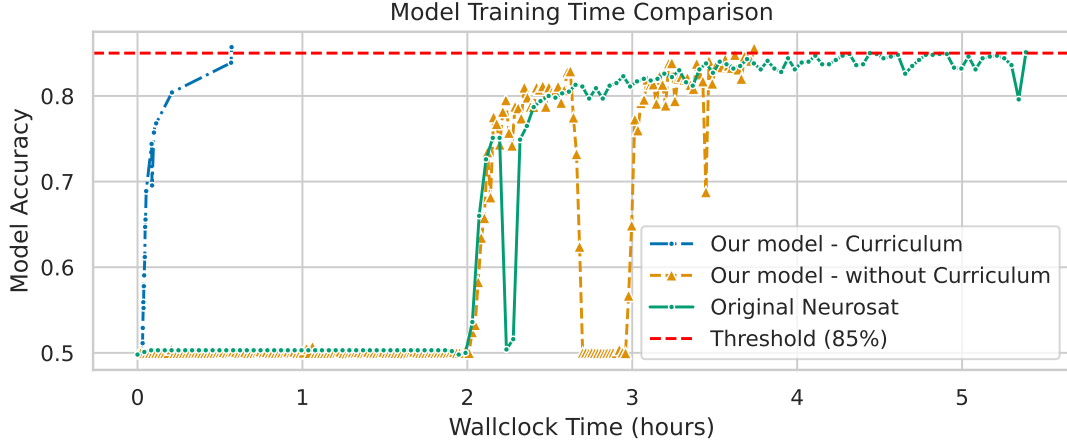


Figure 6: Validation accuracy during training. Our model with a curriculum achieves reaches 85% in approximately 30 minutes, whereas the original NeuroSAT implementation needs over 5 hours. For comparison, we also add our implementation trained on the same data, but without a curriculum. The training of each model stops once it achieves an accuracy of 85% on a validation set.

A Appendix

A.1 Training Tricks and Information

A.1.1 Curriculum Learning

We implement a curriculum learning strategy to improve training efficiency and generalization. The key insight is that starting with simpler (smaller) formulas and gradually increasing complexity allows the model to learn basic logical reasoning patterns before tackling more complex instances.

Our curriculum proceeds as follows:

1. Start training with small formulas (5 variables)
2. Set a validation accuracy threshold for each formula size (starting at 65% for the smallest size and increasing to 85% for the largest)
3. Once the model reaches the threshold accuracy on the current size or reaches a maximum number of epochs (100), increase the formula size by 2 variables
4. When introducing a new size, include formulas from the four previous sizes to prevent catastrophic forgetting
5. Continue until reaching the maximum formula size (40 variables)

This curriculum approach significantly accelerates training compared to starting with the full distribution of formula sizes. Our previous experiments showed that the curriculum-trained model reaches 85% validation accuracy in approximately 30 minutes, compared to over 5 hours for the non-curriculum approach.

A.1.2 Exponential Moving Average (EMA)

We employ Exponential Moving Average (EMA) for model parameter updates during training. EMA maintains a shadow copy of the model parameters that is updated after each training

batch:

$$\theta_{\text{EMA}} \leftarrow \beta \theta_{\text{EMA}} + (1 - \beta) \theta_{\text{current}} \quad (13)$$

where β is the decay rate (we use $\beta = 0.999$).

During validation and testing, we use the EMA parameters instead of the current parameters. This technique significantly stabilizes training and improves generalization, especially in the early stages of training. Our experiments show that EMA provides a smooth validation accuracy curve, while the validation accuracy of the non-EMA model exhibits high variance and jumps of up to 10%.

A.1.3 Learning Rate Schedule

We implement a custom learning rate schedule that combines cosine annealing for the first half of training and a constant minimum learning rate for the second half:

$$\eta(t) = \begin{cases} \eta_{\min} + (\eta_0 - \eta_{\min}) \frac{1 + \cos(\pi t / t_{\text{half}})}{2} & \text{if } t < t_{\text{half}} \\ \eta_{\min} & \text{otherwise} \end{cases} \quad (14)$$

where η_0 is the initial learning rate, η_{\min} is the minimum learning rate (set to 10^{-5}), t is the current epoch, and t_{half} is half of the maximum number of epochs.

This schedule helps the model converge to a good solution in the first half of training and then fine-tune in the second half without disrupting the learned representations.

A.1.4 Impact of hidden dimension on GNN Performance

The dimensionality of the hidden representations, here denoted as `d_model`, specifies the size of the embedding vectors of variables and the hidden state dimension used during the message passing and update phases within the GNN architecture.

The choice of `d_model` directly influences the model’s capacity to learn complex patterns and relationships within the graph structure and node features. It also impacts computational resource requirements, such as memory usage and training time. Understanding how performance metrics vary with different `d_model` values is therefore crucial for effective model design and hyperparameter tuning.

Our evaluation in table 8 generally shows that increasing the `d_model` leads to improved model performance, likely due to the enhanced representational capacity allowing the model to capture more intricate features. However, we observed that this trend exhibits diminishing returns; while significant performance gains are noticeable as the dimension increases up to 64, further increases yield smaller improvements in accuracy relative to the growing computational cost (e.g., peak accuracy at `d_model=256` came with significantly longer training time). This suggests that, considering the marginal benefits, a dimension around 64 still presents a practical optimum, offering a good balance between performance and model complexity/efficiency for this specific setup.

A.2 Diffusion Model Extensions

A.3 Using Unit Propagation for Problem Simplification in the Diffusion Process

In the diffusion process, each iteration provides a belief for every variable, which can be leveraged to continuously simplify the problem via a unit propagation algorithm until it converges to an empty problem, thereby obtaining a solution. The overall solving process is recursive, and its main steps are described as follows:

Table 8: Experimental results demonstrating the impact of hidden dimension size (`d_model`) on model performance and training duration. ‘Embedding Size (`d_model`)’ refers to the dimensionality of the hidden representations within the GNN. ‘Accuracy’ indicates the performance metric achieved by the model. ‘Time (hours)’ specifies the total time required to train the model for each corresponding dimension size.

Embedding Size (<code>d_model</code>)	Accuracy	Time (hours)
16	0.782	1.62
32	0.852	1.66
48	0.860	1.70
64	0.869	1.87
96	0.861	2.32
128	0.864	2.99
256	0.877	7.55

1. Partial Assignment Extraction and Local Unit Propagation

In each diffusion step, a belief value between 0 and 1 is calculated for every variable. A value closer to 1 indicates a stronger inclination toward being *true*, and vice versa. We set a threshold to select variables with high belief and assign them accordingly to obtain a partial assignment. This partial assignment is then used to perform unit propagation for clause simplification. The unit propagation algorithm works as follows:

- If a clause contains a literal that is already satisfied by the current assignment, the clause is marked as satisfied.
- If all literals in a clause have been assigned but none satisfy it, a conflict signal is returned.
- For clauses that are not fully assigned, the unassigned literals are retained to form a simplified clause set.
- For unit clauses obtained during the simplification process (i.e., clauses containing only a single literal), the corresponding unassigned variable is directly assigned the appropriate value, further advancing the local solving process.

2. Multi-Threshold Strategy and Recursive Solving

In the partial assignment extraction step, setting a lower threshold allows for the selection of as many assignments as possible at each step, thereby greatly simplifying the problem; however, it is more likely to select unreliable assignments that may lead to contradictions. To balance this, we adopt a multi-threshold list, starting from the lowest threshold. For each given threshold, if a new partial assignment is obtained, unit propagation is used to update the clauses and evaluate:

- If the simplified clause set becomes empty, all clauses are satisfied and the final solution is directly returned.
- If a conflict occurs or the recursive call at the next level fails, the threshold is raised; if the highest threshold is reached, the process moves to the next recursive level and performs another diffusion step.
- If unit propagation succeeds but the problem is not yet completely solved, the process recurses to the next level, performing a diffusion step on the updated clauses.
- If the recursion reaches a preset maximum depth and the clauses still cannot be satisfied, the recursion at that level fails.

Table 9: Performance on SR100 for Different Diffusion Step and Fixed GNN Step. Note that the Performance is no longer Significantly Improved when Diffusion Steps Larger than 8.

GNN Steps	Diffusion Steps	Avg. Gap	Accuracy (%)
25	4	0.991	69.9
25	5	0.901	71.2
25	6	0.798	72.7
25	8	0.705	73.0
25	10	0.728	72.1
25	20	0.662	73.3
25	30	0.676	72.3
25	40	0.655	73.3
25	50	0.663	73.0

Table 10: Performance on SR100 for Different GNN Step and Fixed Diffusion Step. Note that the Performance is no longer Significantly Improved when GNN Steps Larger than 50.

GNN Steps	Diffusion Steps	Avg. Gap	Accuracy (%)
10	10	2.028	55.0
20	10	0.846	68.6
30	10	0.622	74.4
40	10	0.578	75.5
50	10	0.533	77.6
60	10	0.518	77.2
70	10	0.500	78.6
80	10	0.521	77.9
90	10	0.512	77.6
100	10	0.522	77.4

Table 7 shows the performance and computational cost after applying unit propagation. We tested a fixed model under the settings: GNN steps = 25, diffusion steps = 10, and multi-threshold list = [0.6, 0.75, 0.9]. As shown, incorporating unit propagation improves accuracy by approximately 10% across various problem settings. The last three columns of the table list the number of recursive function calls during the recursion process. Since each call involves one diffusion step, the computational cost incurred by the multi-threshold strategy is directly reflected. We observed that for harder problems, the computational cost of the multi-threshold strategy is actually lower, as unit propagation on partial assignments is more likely to encounter conflicts, thereby reducing the number of recursive branches.

A.4 Influence of Number of Message-passing and Diffusion Steps

For completeness, we also report evaluations in which the number of diffusion steps is fixed and the number of message-passing steps is changing (Table 10) and vice versa (Table 9). We observe that the expansion of the number of iterative steps does not always bring benefits: when the number of one kind of step is fixed, further increasing the number of another kind of step beyond a certain threshold will not lead to performance improvement.

B SDP for MAX-2-SAT

Semidefinite programming (SDP) is a mathematical optimization technique primarily used for problems involving positive semidefinite matrices. In SDP, a linear objective function is optimized over a feasible region given by a *spectrahedron* (an intersection of a convex cone formed by positive semidefinite matrices and an affine subspace) Ramana and Goldman [1995]. Along with the broad scope of applications, SDP has been used to design approximation algorithms for discrete NP-hard problems Gärtner and Matousek [2012]. This is achieved by lifting variables of a problem to a vector space and optimizing a loss function expressed in terms of these vectors.

In this section, we provide a detailed derivation of the SDP relaxation for MAX-2-SAT. The goal is to write an objective function for 2-CNF formulae, which consist of clauses c_1, \dots, c_k over variables x_1, \dots, x_n with at most two literals per clause.

B.1 Derivation of the SDP Relaxation

For each Boolean variable x_i (where $i \in \{1, 2, \dots, n\}$), a new variable $y_i \in \{-1, 1\}$ is associated, and an additional variable $y_0 \in \{-1, 1\}$ is introduced. This additional variable is introduced to unambiguously assign the truth value in the original problem from values of the relaxed problem. It is not possible to just assign *True* (*False*) to x_i if $y_i = 1$ (-1) because quadratic terms cannot distinguish between $y_i \cdot y_j$ and $(-y_i) \cdot (-y_j)$. Instead, the truth value of x_i is assigned by comparing y_i with y_0 : x_i is *True* if and only if $y_i = y_0$, otherwise it is *False*. The assignment is therefore invariant to negating all variables.

To determine the value of a formula, we sum the value of its clauses c which are given by the value function $v(c)$. Here are examples of the value function for different clauses:

$$v(x_i) = \frac{1 + y_0 \cdot y_i}{2} \quad (15)$$

$$v(\neg x_i) = 1 - v(x_i) = \frac{1 - y_0 \cdot y_i}{2} \quad (16)$$

$$v(x_i \vee \neg x_j) = 1 - v(\neg x_i \wedge x_j) \quad (17)$$

$$= 1 - \frac{1 - y_0 \cdot y_i}{2} \cdot \frac{1 + y_0 \cdot y_j}{2} \quad (18)$$

$$= \frac{1}{4}(1 + y_0 \cdot y_i) + \frac{1}{4}(1 - y_0 \cdot y_j) + \frac{1}{4}(1 + y_i \cdot y_j) \quad (19)$$

By summing over all clauses c in the Boolean formula, the following integer quadratic program for MAX-2-SAT is obtained:

$$\text{Maximize: } \sum_{c \in C} v(c) \quad (20)$$

$$\text{Subject to: } y_i \in \{-1, 1\} \text{ for all } i \in \{0, 1, \dots, n\} \quad (21)$$

This can be rewritten by collecting coefficients of $y_i \cdot y_j$ for $i, j \in \{0, 1, \dots, n\}$ and putting them symmetrically into a $(n + 1) \times (n + 1)$ coefficient matrix W . The terms $y_i \cdot y_j$ can be collected in a matrix Y with the same dimensions as W . The elements Y_{ij} correspond to $y_i \cdot y_j$ for $i, j \in \{0, 1, \dots, n\}$. Both matrices are symmetric, hence the sum of all elements in their element-wise product (which is the objective function) can be compactly expressed by using the trace operation. This leads to the following version of the same integer program:

$$\text{Maximize: } \text{Tr}(WY) \quad (22)$$

$$\text{Subject to: } Y_{ii} = 1 \text{ for all } i \in \{0, 1, \dots, n\} \quad (23)$$

$$Y_{ij} = y_i \cdot y_j \text{ for all } i, j \in \{0, 1, \dots, n\} \quad (24)$$

$$y_i \in \{-1, 1\} \text{ for all } i \in \{0, 1, \dots, n\} \quad (25)$$

B.2 Relaxation to Semidefinite Programming

To make the discrete program continuous, we first allow the value of the variables y_i to be any real number between -1 and 1 . However, semidefinite programming goes further and allows variables to be $(n + 1)$ -dimensional unit vectors $(y_0, \dots, y_n) \rightarrow (\mathbf{y}_0, \dots, \mathbf{y}_n)$, as schematically depicted in Figure 7. In this relaxation, the binary products $y_i \cdot y_j$ in the objective function are replaced by inner products $\langle \mathbf{y}_i, \mathbf{y}_j \rangle$.

This can be compactly represented in matrix form by substituting each inner product $\langle \mathbf{y}_i, \mathbf{y}_j \rangle$ with a scalar Y_{ij} of a matrix Y . The fact that these scalars correspond to inner products is encoded by the restriction to *positive-semidefinite* matrices $Y \succeq 0$. The SDP relaxation of MAX-2-SAT can thus be formulated as:

$$\text{Maximize: } \text{Tr}(WY) \quad (26)$$

$$\text{Subject to: } Y_{ii} = 1 \text{ for all } i \in \{0, 1, \dots, n\} \quad (27)$$

$$Y \succeq 0 \quad (28)$$

Positive semidefiniteness of matrix Y ensures that it can be uniquely factorized as $Y = Y^{\frac{1}{2}}(Y^{\frac{1}{2}})^T$. We can then obtain real unit vectors \mathbf{y}_i for all $i \in \{0, \dots, n\}$ such that $Y_{ij} = \langle \mathbf{y}_i, \mathbf{y}_j \rangle$ for all $i, j \in \{0, \dots, n\}$. The constraints $Y_{ii} = 1$ ensure that all vectors \mathbf{y}_i lie on an $(n + 1)$ -dimensional unit sphere.

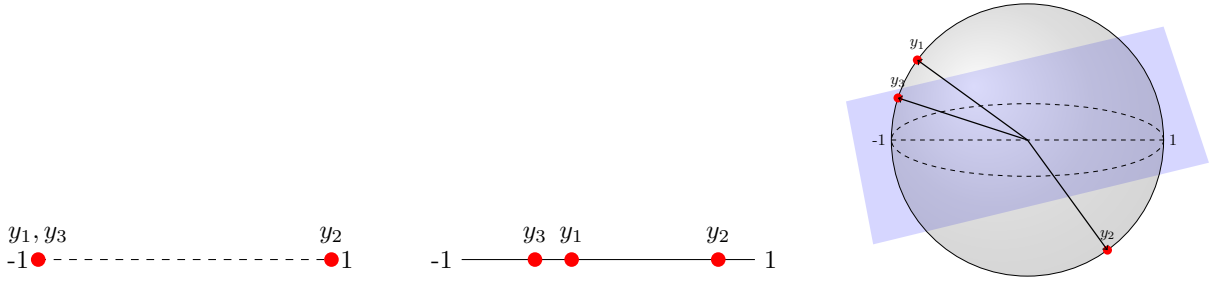


Figure 7: Lifting the variables to a higher dimension, demonstrated on variables y_1, y_2, y_3 . Initially, only integer values of -1 and 1 could be assigned to them (integer program). Next, constraints are relaxed, allowing variables to take any real value between -1 and 1 . Finally, it is permitted for them to be unit vectors in a high-dimensional space (here, 3 dimensions). The hyperplane in the last picture would be used for rounding the variables at the end. This hyperplane can be randomly selected, and truth values for variables y_1, y_2, y_3 are determined based on which side of the hyperplane they land after continuous optimization.

B.3 Interpretation and Rounding

The SDP solver optimizes the numbers in the matrix Y , but using the factorization, we can visualize what happens with the vectors \mathbf{y}_i . The process starts with random unit vectors that are continuously updated to maximize the objective function. If we fix the position of the vector \mathbf{y}_0 (corresponding to the value *true*), we would see that the vectors of variables that will be set to true in the final assignment get closer to the vector \mathbf{y}_0 , while the vectors \mathbf{y}_j of variables that will be set to false move away from it so that the inner product $\langle \mathbf{y}_0, \mathbf{y}_j \rangle$ is close to -1 .

If the formula is satisfiable, the objective function drives the vectors to form two well-separated clusters. However, if only a few clauses can be satisfied simultaneously, the vectors would end up being scattered.

A simple way to round the resulting vectors $(\mathbf{y}_1, \dots, \mathbf{y}_n)$ and get the assignment for the original Boolean variables is to compute the inner product $\langle \mathbf{y}_0, \mathbf{y}_i \rangle$ and assign the value according

to its sign. It is also possible to assign the values by picking a random separating hyperplane, and it can be shown that this rounding gives a 0.8785-approximation of the integer program optimum Goemans and Williamson [1995].

Note that the expressions of the clauses reach their maximum at 1 (when a clause is satisfied by the assignment). This means that the whole formula is satisfiable if the objective function achieves a value equal to the number of clauses in the formula. Another way to check satisfiability is to plug the obtained solution into the formula and verify whether it is satisfied. Therefore, we can obtain an incomplete SAT solver from this SDP formulation.

Similar SDPs can be obtained for different versions of MAX-SAT (with larger clauses). From empirical observation, the convergence threshold of the SDP solver needs to be decreased significantly compared to MAX-2-SAT in order to obtain a good approximation for these more complicated versions.