

RWTH Aachen
Faculty of Computer Science
Chair of Combinatorial Optimization

BACHELOR THESIS IN COMPUTER SCIENCE

Min-Max and Regret Spanning Trees

Complexity and Approximation under
Discrete and Interval Uncertainty

Submitted by
Archit Dhama
Matriculation Number: 396580
archit.dhama@rwth-aachen.de

First Examiner	Second Examiner
Prof. Dr. Christina Büsing Chair of Combinatorial Optimization RWTH Aachen	Prof. Dr. Arie M. C. A. Koster Chair II für Mathematics RWTH Aachen

Aachen, January 29, 2026

Abstract

Spanning tree problems connect all network nodes at minimum cost, but edge costs are often uncertain due to market fluctuations, estimation errors, or operational changes. This thesis studies robust spanning trees under two uncertainty models (discrete scenarios, interval ranges) and two objectives (Min-Max, Min-Max Regret), classifying computational complexity and approximability across this design space.

We provide a self-contained treatment from first principles. After establishing MST foundations with five complete proofs (fundamental cycle and cut lemmas, three optimality criteria via exchange arguments), we systematically analyse four problem variants. For intervals, we prove extremal lemmas showing worst cases occur at boundaries: Min-Max becomes polynomial whilst Regret remains NP-hard with a 2-approximation. For discrete scenarios, we prove weak NP-hardness for $K=2$ via partition reduction and survey results for larger K including pseudo-polynomial algorithms and approximation schemes.

The thesis delivers nine complete proofs including two representatives (partition reduction, 2-approximation algorithm) and synthesises eight problem variants in a comprehensive classification table. A fixed four-vertex micro-graph illustrates concepts throughout. This work balances mathematical rigour with pedagogical clarity, serving as a focused entry point into robust combinatorial optimisation.

Contents

Abstract	i
1 Introduction	1
1.1 Motivation	1
1.2 Research Questions and Scope	2
1.3 Contributions	3
1.4 Thesis Structure	4
2 Foundations	5
2.1 Graphs, Trees, and Notation	5
2.2 Minimum Spanning Tree Optimality	5

2.3	Kruskal and Prim: Algorithmic Remarks	5
2.4	Uncertainty Models and Robust Optimisation	5
2.5	Algorithmic Preliminaries	5
3	Min-Max Spanning Tree	6
3.1	Problem Formulation	6
3.2	Interval Worst-Case Characterisation	6
3.3	Complexity and Approximation	6
3.3.1	Discrete Scenarios: Constant K	6
3.3.2	Discrete Scenarios: Unbounded K	6
3.4	Discussion	6
4	Min-Max Regret Spanning Tree	7
4.1	Regret Formulation	7
4.2	Interval Regret Extremal Behaviour	7
4.3	Complexity for Discrete Scenarios	7
4.3.1	Constant K	7
4.3.2	Unbounded K	7
4.4	Interval Regret Approximation	7
4.5	Discussion	7
5	Synthesis and Example Gallery	8
5.1	Comprehensive Results Classification	8
5.1.1	Complexity and Approximation Landscape	8
5.1.2	Key Patterns and Insights	8
5.2	Example Gallery	8
5.2.1	Micro-Graph Solutions Across Objectives	8
5.2.2	Extremal Behaviour Visualisation	8
5.2.3	Worked Example: Partition Reduction	8
5.3	Key Takeaways	8
6	Conclusion and Outlook	9
6.1	Summary	9
6.1.1	Thesis Achievements	9
6.1.2	Core Findings	9
6.2	Limitations	9
6.2.1	Scope Restrictions	9
6.2.2	Proofs and Coverage	9
6.2.3	Micro-Graph and Empirical Work	9
6.3	Outlook	9
6.3.1	Extensions in Uncertainty Modelling	9
6.3.2	Extensions in Solution Concepts	9
6.3.3	Open Problems	9
A	Notation	10
Affidavit / Eidesstattliche Versicherung		11

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

Consider the challenge of designing a fiber-optic telecommunications network connecting major cities across a region. The network must link all locations whilst minimising total cable installation cost. However, actual deployment costs depend on numerous factors: terrain characteristics (installing cables through mountainous regions costs substantially more than across plains), regulatory approvals that vary by jurisdiction, and material prices that fluctuate with global supply chains. Industry experience suggests that cable costs can vary by 20–40% from initial estimates over a typical planning horizon. Consequently, a network design that appears optimal under today’s cost estimates may become 50% more expensive if realised costs differ significantly. This uncertainty raises a fundamental question: how should one design infrastructure networks that remain near-optimal despite cost variations?

When edge costs are known precisely, the *minimum spanning tree* (MST) provides the classical solution. The MST connects all nodes at minimum total cost and can be computed efficiently using well-established algorithms such as Kruskal’s greedy edge selection or Prim’s incremental tree growing, both running in polynomial time. However, in practice, costs are rarely known with certainty at design time. A tree that is optimal for one cost estimate may perform poorly when costs deviate, potentially incurring substantially higher expenses or missing better alternatives. Thus, whilst the MST problem is well understood in the deterministic setting, real-world applications demand solutions that are *robust* to cost uncertainty.

Under uncertainty, two natural design objectives emerge. The first is the *Min-Max* objective, which seeks to minimise the worst-case absolute cost across all possible cost realisations. This conservative approach hedges against expensive scenarios by selecting a spanning tree whose cost remains controlled even in the most unfavourable circumstances. Mathematically, given a set of possible cost vectors, the goal is to find a tree that minimises the maximum cost it could incur.

The second objective is *Min-Max Regret*, which measures performance not in absolute terms but relative to what could have been achieved with perfect hindsight. Regret quantifies the difference between a chosen solution’s cost and the optimal cost for a given realisation. The Min-Max Regret objective minimises the worst-case performance gap, hedging against the possibility of missing the truly optimal solution. Unlike Min-Max, which guards against high absolute costs, Regret focuses

on relative performance: ensuring that no matter which cost scenario materialises, the chosen tree is competitive with the scenario-optimal solution.

To model uncertainty, we consider two standard representations that arise naturally in applications. The first is *discrete scenario uncertainty*, where a planner enumerates a finite set of plausible cost vectors, each representing a different potential future state. For instance, in the telecommunications example, scenarios might correspond to different combinations of regulatory outcomes and material price levels. The second is *interval uncertainty*, where each edge cost is known to lie within a specified range, bounded by lower and upper estimates. This representation is particularly natural when costs are subject to continuous variation within known limits, such as fluctuations in commodity prices or exchange rates.

Together, these two objectives (Min-Max and Min-Max Regret) and two uncertainty models (discrete scenarios and interval bounds) define a design space with four fundamental problem variants. The central question addressed in this thesis is: how does computational complexity and approximability vary across this space? Which combinations admit polynomial-time algorithms, and which are computationally intractable? For the hard cases, what approximation guarantees can be achieved? These questions form the core of robust spanning tree optimisation and motivate the systematic classification developed in the following chapters.

1.2 Research Questions and Scope

This thesis addresses three core questions that arise naturally from the robust spanning tree framework outlined above.

First, what is the computational complexity of finding optimal robust spanning trees under discrete scenarios and interval uncertainty? For each combination of objective (Min-Max or Regret) and uncertainty model (discrete or interval), does there exist a polynomial-time algorithm, or is the problem computationally intractable? When hardness arises, does it stem from the number of scenarios being unbounded, or is even the case of two scenarios already difficult? Answering these questions requires establishing precise complexity classifications: polynomial-time solvability, weak NP-hardness (admitting pseudo-polynomial algorithms), or strong NP-hardness (remaining hard even with unary-encoded inputs).

Second, where do worst-case costs occur within interval uncertainty? When each edge cost lies in a specified range, must one consider all points within the Cartesian product of intervals, or can worst cases be characterised more simply? This question is critical because if worst cases always occur at interval boundaries—such as the upper or lower bounds—then the infinite continuous uncertainty set can be reduced to a finite discrete problem. Extremal lemmas that identify such worst-case locations are therefore essential tools for both algorithmic design and theoretical analysis.

Third, for problems that are computationally hard, what approximation guarantees can be achieved? Can we design polynomial-time algorithms that provably compute solutions within a constant factor of optimal, or does the problem admit a fully polynomial-time approximation scheme (FPTAS) that achieves arbitrarily good approximations in time polynomial in both problem size and accuracy? Conversely, are there hardness-of-approximation results that limit what can be achieved? These questions determine the boundary between tractable and intractable robust

optimisation.

The scope of this thesis is deliberately focused. We restrict attention to spanning tree problems in undirected, connected graphs with static, single-stage decisions: a tree is selected once, before any uncertainty is resolved, and remains fixed. The uncertainty models studied are discrete scenarios (a finite collection of cost vectors) and interval bounds (per-edge lower and upper limits forming a Cartesian product). The objectives analysed are Min-Max (minimising worst-case absolute cost) and Min-Max Regret (minimising worst-case performance gap against scenario-optimal solutions).

This thesis does not cover budgeted uncertainty (where at most a specified number of edges deviate from nominal costs), polyhedral uncertainty sets, or distributional robustness (where costs follow known probability distributions). It also excludes two-stage and recoverable optimisation models, where decisions can be partially adjusted after observing realisations. These extensions represent active research directions and are briefly discussed in the outlook (Chapter 6), but they lie outside the core scope of this work. The goal is not encyclopedic coverage but rather a self-contained, rigorous treatment of the selected models, proving foundational results from first principles and synthesising established complexity and approximability findings into a unified classification.

1.3 Contributions

This thesis delivers five principal contributions to the literature on robust spanning tree optimisation.

First, we provide self-contained minimum spanning tree foundations (Chapter 2). Rather than assuming familiarity with MST theory, we prove five fundamental results from first principles: the fundamental cycle and cut lemmas via exchange arguments, and three MST optimality criteria (cycle property, cut property, and their equivalence). These proofs establish the mathematical toolkit needed for subsequent robust analysis. Additionally, Chapter 2 includes a complexity primer defining polynomial time, NP-hardness (weak and strong), and approximation concepts (constant-factor algorithms, FPTAS, PTAS) with precise terminology for classifying the robust variants.

Second, we establish extremal characterisation for interval uncertainty (Chapters 3 and 4). For the Min-Max objective, Lemma 3.1 proves that worst-case costs occur when chosen edges are assigned their upper bounds, immediately yielding a polynomial-time algorithm. For the Regret objective, Lemma 4.1 proves that worst-case regret occurs at boundary vertices of the interval box, though determining which boundary is harder. Both lemmas are proved in full, and their implications for algorithm design are developed carefully.

Third, we provide two representative complexity proofs (Chapters 3 and 4). Theorem 3.2 establishes weak NP-hardness of Min-Max spanning trees with two discrete scenarios via a reduction from the PARTITION problem, constructing a grid graph where scenario costs encode target subset sums. This proof is presented completely, including the reduction construction, correctness argument, and encoding analysis. Theorem 4.2 then extends this result to Min-Max Regret by reusing the same construction with adjusted analysis, demonstrating that regret complexity mirrors

absolute cost complexity for discrete scenarios. Together with cited results for larger scenario counts, these establish a complete complexity hierarchy across the design space.

Fourth, we prove a 2-approximation algorithm for interval regret spanning trees (Chapter 4). Theorem 4.5 demonstrates that solving the MST problem at midpoint costs yields a solution whose worst-case regret is at most twice optimal. The proof is developed rigorously through two supporting lemmas (4.6 and 4.7) bounding the midpoint solution's regret from below and above. This result is noteworthy as the best known constant-factor approximation for any robust combinatorial optimisation problem under interval uncertainty, though the tightness of the factor 2 remains an open question in the literature.

Fifth, we synthesise results into a comprehensive classification table (Chapter 5). Table 5.1 organises eight problem variants (two objectives, two uncertainty models, two scenario count regimes) with their complexity classes and approximation guarantees. The table reveals structural patterns: intervals exhibit extremal behaviour enabling tractability, scenario count K acts as a complexity parameter (constant versus unbounded), and Min-Max and Regret objectives display parallel hardness for discrete scenarios but diverge for intervals. This classification serves as a reference for practitioners and researchers working with robust spanning tree models.

Throughout the thesis, a fixed four-vertex micro-graph with interval costs and derived discrete scenarios provides worked examples illustrating optimal solutions under each objective and uncertainty model. This pedagogical device grounds abstract theoretical results in concrete calculations, demonstrating that robust optima genuinely differ from nominal solutions.

1.4 Thesis Structure

The remainder of this thesis is organised as follows.

Chapter 2 establishes mathematical foundations, proving minimum spanning tree optimality criteria from first principles and defining the robust optimisation framework (uncertainty models, objectives, complexity terminology). A fixed micro-graph is introduced for illustrative examples throughout.

Chapter 3 analyses the Min-Max objective, characterising worst-case costs for interval uncertainty and establishing computational complexity for discrete scenarios across different values of the scenario count parameter.

Chapter 4 examines the Min-Max Regret objective, which measures relative performance rather than absolute cost. Interval regret is shown to be harder than interval Min-Max despite sharing extremal properties, whilst discrete regret mirrors the complexity hierarchy of discrete Min-Max.

Chapter 5 synthesises results from Chapters 3 and 4 into a comprehensive classification table spanning eight problem variants. An example gallery demonstrates how optimal solutions vary across objectives and uncertainty models, and key patterns are extracted for practitioners.

Chapter 6 summarises achievements, acknowledges deliberate scope limitations, and surveys related robust optimisation models (budgeted uncertainty, two-stage formulations, recoverable robustness) along with open questions from the literature.

Appendix A provides an alphabetised notation table for quick reference.

Chapter 2

Foundations

2.1 Graphs, Trees, and Notation

2.2 Minimum Spanning Tree Optimality

Lemma 1 (Fundamental Cycle).

Proof.

□

Lemma 2 (Fundamental Cut).

Proof.

□

Theorem 1 (MST Cycle Criterion).

Proof.

□

Theorem 2 (MST Cut Criterion).

Proof.

□

Theorem 3 (MST Characterisation).

Proof.

□

2.3 Kruskal and Prim: Algorithmic Remarks

2.4 Uncertainty Models and Robust Optimisation

2.5 Algorithmic Preliminaries

Summary

Chapter 3

Min-Max Spanning Tree

3.1 Problem Formulation

3.2 Interval Worst-Case Characterisation

Lemma 3 (Interval Extremal Cost).

Proof.

□

3.3 Complexity and Approximation

3.3.1 Discrete Scenarios: Constant K

Theorem 4 (Partition Reduction).

Proof.

□

Theorem 5 (Pseudo-Polynomial Algorithm).

3.3.2 Discrete Scenarios: Unbounded K

Theorem 6 (Strong NP-Hardness).

3.4 Discussion

Summary

Chapter 4

Min-Max Regret Spanning Tree

4.1 Regret Formulation

4.2 Interval Regret Extremal Behaviour

Lemma 4 (Interval Extremal Regret).

Proof.

□

4.3 Complexity for Discrete Scenarios

4.3.1 Constant K

Theorem 7 (K=2 Weak NP-Hardness).

Proof.

□

Theorem 8 (Pseudo-Polynomial for Constant K).

4.3.2 Unbounded K

Theorem 9 (Strong NP-Hardness).

4.4 Interval Regret Approximation

Theorem 10 (Interval NP-Hardness).

Theorem 11 (2-Approximation via Midpoint).

Lemma 5 (Lower Bound on Regret).

Proof.

□

Lemma 6 (Upper Bound Relating Solutions).

Proof.

□

Proof of Theorem 11.

□

4.5 Discussion

Summary

Chapter 5

Synthesis and Example Gallery

5.1 Comprehensive Results Classification

5.1.1 Complexity and Approximation Landscape

5.1.2 Key Patterns and Insights

5.2 Example Gallery

5.2.1 Micro-Graph Solutions Across Objectives

5.2.2 Extremal Behaviour Visualisation

5.2.3 Worked Example: Partition Reduction

5.3 Key Takeaways

Summary

Chapter 6

Conclusion and Outlook

6.1 Summary

6.1.1 Thesis Achievements

6.1.2 Core Findings

6.2 Limitations

6.2.1 Scope Restrictions

6.2.2 Proofs and Coverage

6.2.3 Micro-Graph and Empirical Work

6.3 Outlook

6.3.1 Extensions in Uncertainty Modelling

6.3.2 Extensions in Solution Concepts

6.3.3 Open Problems

Appendix A

Notation

Affidavit / Eidesstattliche Versicherung

I hereby declare that I have written this thesis independently and that I have not used any sources or aids other than those indicated. All passages which are quoted or closely paraphrased from other works are marked as such. This thesis has not been submitted in the same or a substantially similar form to any other examination board.

Place, Date

Signature