

# Quantum Approximation for Phylogenetic Tree Construction

Yeuko Manganyi<sup>1</sup>, Darren Martin<sup>1</sup>

<sup>1</sup>Department of Computational Biology and Bioinformatics, University of Cape Town, Cape Town, South Africa



## Rationale

Phylogenetic tree construction is a fundamental problem in evolutionary biology and bioinformatics.

However, the number of possible tree topologies grows super-exponentially with the number of taxa, making exhaustive search infeasible.

Many phylogenetic inference criteria, correspond to **NP-hard optimization problems**, motivating the exploration of alternative computational paradigms.

## Problem Statement

- Phylogenetic tree inference requires the use of convoluted search algorithms
- Finding the optimal tree is computationally intractable for large datasets
- Classical methods rely on heuristics without guarantees of global optimality
- The applicability of quantum optimization algorithms to this problem remains largely unexplored

## Research Aim

To evaluate whether a QAOA-based optimization approach can effectively solve the phylogenetic inference problem, starting with maximum parsimony, and how it compares to classical parsimony heuristics.

## Expected Outcomes

- A Hamiltonian/QUBO formulation of maximum parsimony inference
- A working QAOA-based proof-of-concept implementation
- Empirical comparison with classical parsimony methods
- Insights into the **practical limits and potential** of QAOA for phylogenetics

## Methodology

### Maximum Parsimony as Optimization

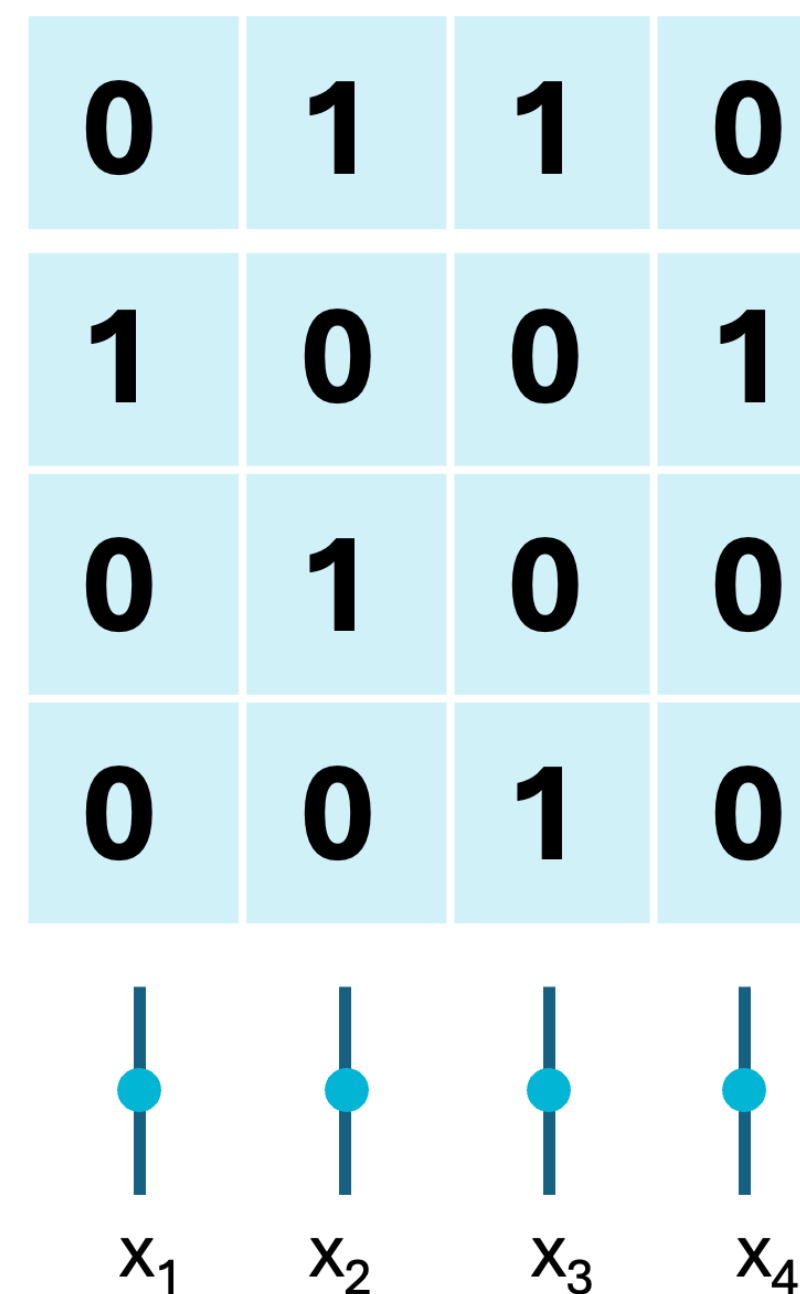
$$S(T) = \sum_{c=1}^C \min_{A_c} \sum_{e \in T} \mathbb{I}(A_c(u_e) \neq A_c(v_e))$$

- Tree inference formulated as a combinatorial optimization problem
- Parsimony score expressed as a cost function
- Tree validity enforced via constraint penalties

### QUBO Conversion

- Tree structure encoded using **binary variables**
- Parsimony score expressed as a quadratic cost function
- Constraints added to penalize invalid trees

Q

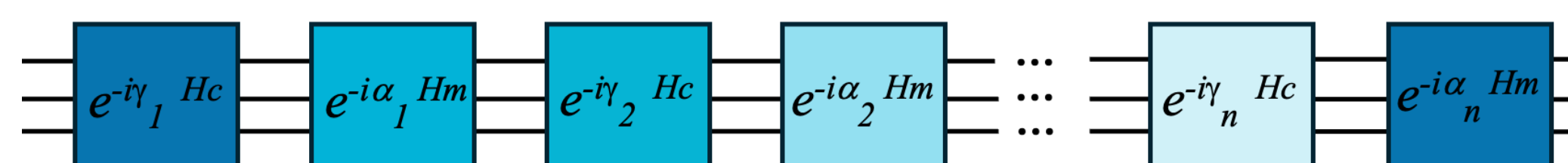


### Hamiltonian / QUBO Formulation

- Total cost = parsimony + constraint penalties
- Low energy = valid tree with few changes
- High energy = invalid or suboptimal tree

$$H = H_{\text{parsimony}} + \lambda H_{\text{constraints}}$$

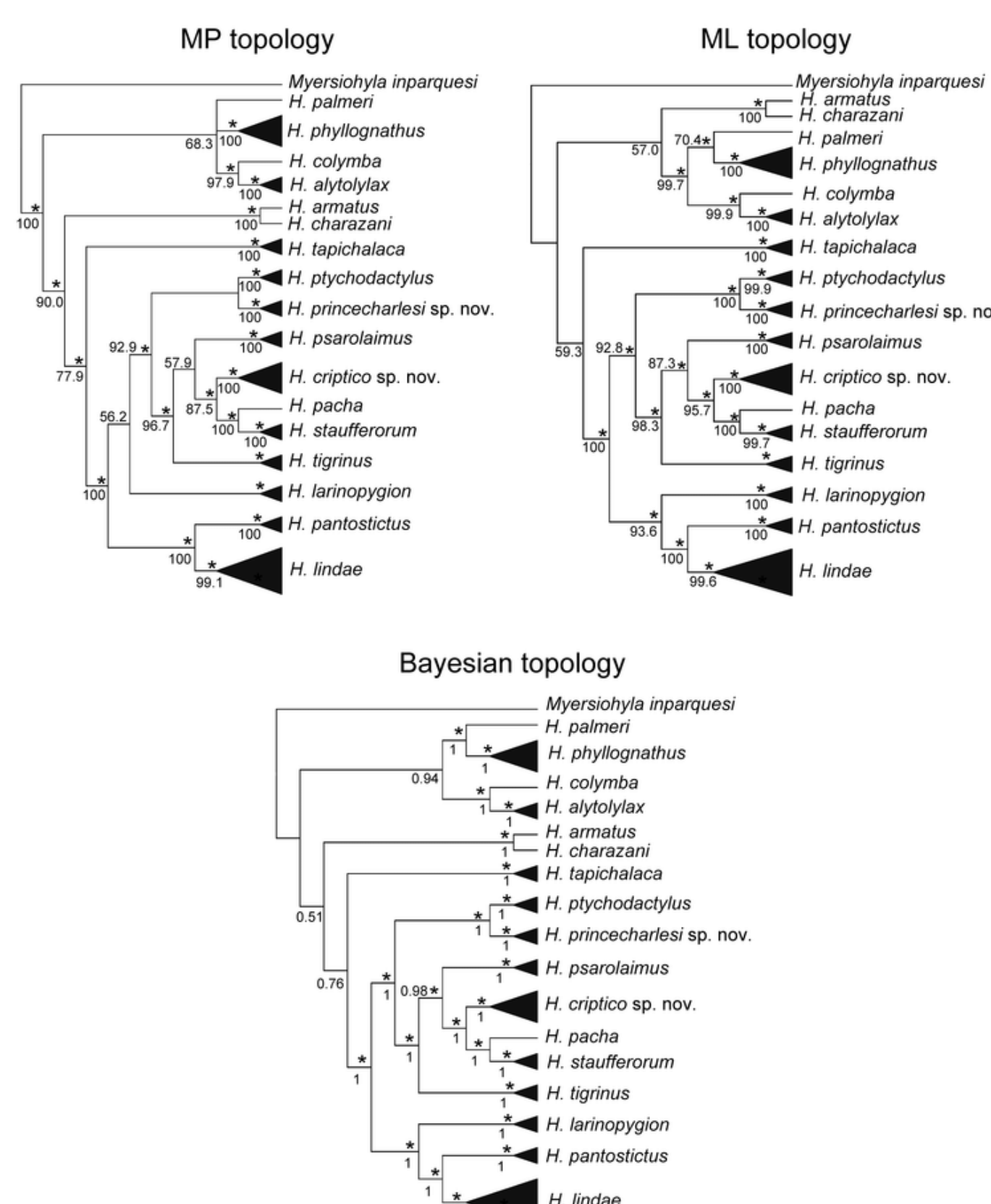
### Quantum Approximate Optimization Algorithm (QAOA)



- Alternates between two quantum operations
- $H_C$ : encodes the optimization objective (parsimony score)
- $H_M$ : explores the solution space
- Parameters are classically optimized

### Phylogenetic Tree Comparisons

- Trees inferred using QAOA and classical heuristics
- Computational performance metrics for the four methods
- Compared using parsimony score and topology
- Highlights similarities and differences in solutions



## References:

- Coloma LA, Carvajal-Endara S, Duenas JF, Paredes-Recalde A, Morales-Mite M, Almeida-Reinoso D, Tapia EE, Hutter CR, Toral E, Guayasamin JM. Molecular phylogenetics of stream treefrogs of the Hyloscirtus larinopygion group (Anura: Hylidae), and description of two new species from Ecuador. Zootaxa. 2012 Jul 4;3364(1):1-78
- Onodera W, Hara N, Aoki S, Asahi T, Sawamura N. Phylogenetic tree reconstruction via graph cut presented using a quantum-inspired computer. Molecular Phylogenetics and Evolution. 2023 Jan 1;178:107636.
- Glover F, Kochenberger G, Hennig R, Du Y. Quantum bridge analytics I: a tutorial on formulating and using QUBO models. Annals of Operations Research. 2022 Jul;314(1):141-83.
- Blekos K, Brand D, Ceschini A, Chou CH, Li RH, Pandya K, Summer A. A review on quantum approximate optimization algorithm and its variants. Physics Reports. 2024 Jun 2;1068:1-66.

