Draft version August 29, 2025
Typeset using LaTeX **twocolumn** style in AASTeX631

Cosmological Parameter Inference from Merger Trees Using Hierarchical Quantum Tensor Networks

Denario[1]

[1] *Anthropic, Gemini & OpenAI servers. Planet Earth.*

ABSTRACT

Inferring cosmological parameters from the intricate, hierarchical structures of dark matter merger trees is crucial for understanding cosmic evolution but presents significant challenges for conventional statistical methods. We introduce a novel framework leveraging Hierarchical Quantum Tensor Networks (HQTNs), specifically Tree Tensor Networks (TTNs), to directly predict these parameters. Our approach represents each merger tree as a hierarchical graph, where individual halo properties (mass, concentration, Vmax, and scale factor) are embedded into node tensors via a shared neural network. Hierarchical relationships and varying tree topologies are captured by learnable basis tensors, selected according to a node's number of children, which are then contracted from the leaves to the root using the `quimb` library. The resulting fixed-dimension root vector is fed into a linear layer to predict the target cosmological parameters, Omega_m and sigma_8. The complete model, including feature embedding, basis tensors, and the prediction head, is trained end-to-end on a dataset of 1000 simulated merger trees using Mean Squared Error loss and optimized with `JAX` and `optax` for efficient automatic differentiation. This methodology provides a powerful, interpretable means to exploit the deep hierarchical correlations within merger trees, thereby advancing robust cosmological parameter inference beyond traditional statistical summaries.

Keywords: Astronomy data visualization, Sigma8, Regression, GPU computing, Neural networks

1. INTRODUCTION

The formation and evolution of cosmic structures stand as a central pillar of modern cosmology. The large-scale distribution of matter, particularly the elusive dark matter, offers a profound probe into the universe's fundamental properties and evolutionary history. Among the most critical cosmological parameters are Ω_m, which quantifies the total matter density relative to the critical density, and σ_8, which represents the amplitude of primordial matter fluctuations on scales of 8 h^{-1} Mpc. These parameters are not merely abstract numbers; they fundamentally govern the gravitational growth of structure, dictate the abundance of dark matter halos—the cosmic cradles where galaxies reside—and ultimately shape the observable universe.

Dark matter halos are not static entities; they assemble hierarchically over cosmic time through a continuous process of accretion and mergers, where smaller progenitor halos coalesce to form larger ones (Benson et al. 2012; Jiang & van den Bosch 2013; Yung et al. 2024). This intricate evolutionary history is faithfully recorded in "merger trees," which are directed acyclic graphs that delineate the complete lineage of a halo and its progenitors across various cosmic epochs (Parkinson et al. 2007; Nguyen et al. 2025).

Merger trees thus encapsulate an unparalleled wealth of information about the interplay between initial cosmic conditions, gravitational collapse, and the growth of structure (Parkinson et al. 2007; Yung et al. 2024). Despite this rich informational content, they remain a significantly underutilized resource for cosmological parameter inference. Traditional approaches typically reduce this complexity to summary statistics, such as halo mass functions, clustering statistics, or simple growth rates (Benson et al. 2012; Jiang & van den Bosch 2013). While these methods have proven invaluable, they inherently condense the intricate, multi-scale hierarchical information within the full merger tree structure into simplified forms, inevitably discarding fine-grained details and complex, non-linear correlations. The central challenge, therefore, lies in developing a robust computational framework capable of directly processing these complex, variable-sized, and highly structured graph representations to extract cosmological information without suffering from an undesirable loss of fidelity (Nguyen et al. 2025).

This paper addresses this fundamental challenge by introducing a novel framework that leverages Hierarchical Quantum Tensor Networks (HQTNs), specifically Tree Tensor Networks (TTNs), to directly infer cosmological parameters from dark matter merger trees (Milsted et al. 2019; Gunst et al. 2019).

Tensor networks are a powerful mathematical formalism for representing and manipulating high-dimensional data and functions by decomposing them into a network of interconnected low-rank tensors (Milsted et al. 2019). This decomposition allows for a compact and efficient representation of correlations. Critically, hierarchical variants like TTNs are uniquely well-suited for data exhibiting an intrinsic tree-like or hierarchical structure, making them a natural and compelling choice for modeling the evolutionary architecture of merger trees (Milsted et al. 2019; Gunst et al. 2019; Dowling et al. 2024). By explicitly encoding the parent-child relationships and the physical properties of individual halos within a tensor network, our approach aims to uncover and exploit the deep, non-linear, and multi-scale correlations that are inherently predictive of cosmological parameters, moving beyond the limitations of pre-defined summary statistics.

Our methodology begins by representing each dark matter merger tree as a hierarchical graph (Nguyen et al. 2025). The physical properties of individual halos—specifically their mass, concentration, maximum circular velocity (V_{max}), and formation scale factor—are embedded into node-specific tensors through a shared, shallow neural network (Lucie-Smith et al. 2024). The crucial hierarchical relationships and the diverse topological structures inherent to merger trees are captured by a set of learnable "basis" tensors (Nguyen et al. 2025). These basis tensors are dynamically selected for each node based on its number of children, allowing the model to adapt flexibly to the varying branching patterns observed across different merger histories.

The entire tensor network is then systematically contracted from the leaves of the tree upwards towards the root, a process efficiently executed using the `quimb` library. This contraction yields a fixed-dimension root vector, which effectively encapsulates the tree's essential cosmological information. This vector is subsequently passed through a linear layer to predict the target cosmological parameters, Ω_m and σ_8. The complete model, encompassing the feature embedding network, the set of learnable basis tensors, and the final prediction head, is trained end-to-end on a dataset of 1000 simulated merger trees (Lucie-Smith et al. 2024; Nguyen et al. 2025). Training is performed by minimizing a Mean Squared Error loss, leveraging the efficient automatic differentiation capabilities of `JAX` and `optax` for optimization.

To verify the efficacy and robustness of our proposed framework, we conduct a comprehensive quantitative evaluation on a held-out test set, assessing the model's predictive accuracy for both Ω_m and σ_8 using standard metrics such as Mean Squared Error and R-squared values.

Beyond mere prediction, a significant advantage of our tensor network approach is its potential for interpretability. We delve into this by analyzing the learned weights of the feature embedding network and the structure of the basis tensors. This analysis is designed to reveal which specific halo properties and structural motifs within the merger trees are most salient and predictive of cosmological parameters. Such insights promise to offer new physical understanding into the intricate connection between the microscopic evolution of dark matter halos and the macroscopic parameters governing the cosmos. Ultimately, this work demonstrates that Hierarchical Quantum Tensor Networks provide a powerful, flexible, and interpretable paradigm for directly exploiting the rich hierarchical information embedded within merger trees, thereby advancing the frontier of robust cosmological parameter inference beyond the limitations of traditional statistical summaries.

2. METHODS

The methodology for inferring cosmological parameters from dark matter merger trees leverages a novel Hierarchical Quantum Tensor Network (HQTT) framework, specifically Tree Tensor Networks (TTNs). This section details the dataset preparation, the architecture of the HQTT model, the training procedure, and the evaluation and interpretation techniques employed.

2.1. *Dataset and Preprocessing*

Our study utilizes a dataset of 1000 simulated dark matter merger trees (Hearin et al. 2022; Nguyen et al. 2025), each encapsulating the hierarchical formation history of a dark matter halo (Parkinson et al. 2007; Nguyen et al. 2025). These simulations provide a rich source of information, allowing us to probe the intricate connections between microscopic halo evolution and macroscopic cosmological parameters (Nguyen et al. 2025).

2.1.1. *Merger Tree Dataset*

Each merger tree in the dataset is represented as a graph, loaded as a ‘torch_geometric.data.Data‘ object. This structure effectively captures the parent-child relationships between halos (Nguyen et al. 2025).

For each halo (node) in a tree, four key physical properties are recorded: its mass, concentration, maximum circular velocity (V_{max}), and the cosmic scale factor at which it exists (Bose et al. 2022). These properties are provided in a feature vector 'x' of shape '[num_nodes, 4]'. The mass, concentration, and V_{max} are already log-transformed, while the scale factor ranges between 0 and 1.

The graph connectivity is defined by 'edge_index', indicating the directed relationships between progenitors and their descendants (Bose et al. 2022).

Each tree is also associated with a pair of target cosmological parameters, 'y', representing Ω_m and σ_8, both of which are crucial for characterizing the universe's matter content and the amplitude of primordial fluctuations.

2.1.2. *Node Feature Normalization*

To ensure robust model training and prevent features with larger scales from dominating the learning process (de Amorim et al. 2022; Pinheiro et al. 2025), the four node features ('log10(mass)', 'log10(concentration)', 'log10(Vmax)', 'scale_factor') undergo a standardization process. For each feature, the mean and standard deviation are computed exclusively from the training set (Pinheiro et al. 2025). Subsequently, each feature is normalized to have a mean of 0 and a standard deviation of 1 using these calculated statistics (Pinheiro et al. 2025). This consistent normalization procedure is then applied across all training, validation, and test datasets.

2.1.3. *Dataset Splitting*

The full dataset of 1000 merger trees is partitioned into training, validation, and testing sets to facilitate robust model development and unbiased performance assessment. A random split is employed, allocating 70% of the trees (700 trees) to the training set, 15% (150 trees) to the validation set for hyperparameter tuning, and the remaining 15% (150 trees) to the test set for final model evaluation. This random splitting ensures that the model is evaluated on unseen tree structures, thereby assessing its generalization capabilities for inferring cosmological parameters (Hernández-Martínez et al. 2024; Min et al. 2024).

2.1.4. *Graph Structure Analysis*

Prior to model construction, an extensive analysis of the graph structures within the dataset is performed (Nguyen et al. 2025). This includes calculating statistics such as the minimum, maximum, mean, and standard deviation of the number of nodes and edges per tree. Crucially, we determine the maximum number of children any node possesses across all merger trees (Nguyen et al. 2025). This information is vital for defining the complete set of basis tensors required by the TTN, ensuring that a specific learnable tensor exists for every possible branching factor observed in the dataset.

2.2. *Hierarchical Quantum Tensor Network Model Architecture*

The core of our approach is a Tree Tensor Network (TTN) designed to directly process the hierarchical structure of merger trees (Li et al. 2025), as described in the introduction. This architecture allows for the efficient capture of multi-scale correlations within the tree (Stoian et al. 2024; Milbradt et al. 2024), leading to a fixed-dimension representation of the entire tree.

The 'quimb' library is utilized for the construction and contraction of the tensor networks (Stoian et al. 2024), while 'JAX' provides the necessary framework for automatic differentiation and optimization of the learnable parameters.

2.2.1. *Node Feature Embedding Network*

For each individual halo (node) within a merger tree, its four normalized physical properties are first transformed into a dense embedding vector (Moriwaki et al. 2025). This is achieved using a small, shared neural network, denoted 'NN_embed'. This network consists of one or two linear layers, followed by a non-linear activation function (e.g., ReLU or GeLU). 'NN_embed' takes the 4-dimensional normalized feature vector as input and outputs a 'd_embed'-dimensional embedding vector.

The weights and biases of 'NN_embed' are learnable parameters, shared across all nodes in all trees, allowing the model to learn a universal representation for halo properties that is relevant for cosmological inference (Chatterjee & Villaescusa-Navarro 2025). The dimension 'd_embed' is a hyperparameter, typically set to values such as 8, 16, or 32.

2.2.2. *Learnable Basis Tensors*

The hierarchical relationships and diverse topological patterns of merger trees (Huang et al. 2025) are encoded through a set of learnable "basis" tensors (Ceruti et al. 2024). These tensors are dynamically selected for each node based on its structural role, specifically its number of children. A key hyperparameter, 'd_bond', defines the bond dimension of the TTN (Li et al. 2024), influencing the capacity of the network to store information. The basis tensors include:

- 'T_leaf': For leaf nodes (halos with no children). Its shape is '(d_embed, d_bond)'.

- ‘T_1child‘: For nodes with one child. Its shape is ‘(d_embed, d_bond, d_bond)‘, where the second ‘d_bond‘ dimension connects to the child’s contracted tensor, and the third ‘d_bond‘ dimension connects to its parent.
- ‘T_2child‘: For nodes with two children. Its shape is ‘(d_embed, d_bond, d_bond, d_bond)‘, with one dimension for the parent connection and two for the child connections.
- This pattern extends up to ‘T_max_children‘, where ‘max_children‘ is determined from the graph structure analysis of the dataset.

The elements of these basis tensors are learnable parameters, enabling the model to adaptively capture the varying branching patterns and their cosmological significance.

2.2.3. *Tree Tensor Network Construction and Contraction*

For each merger tree, a ‘quimb.TensorNetwork‘ object is constructed by iterating through its nodes, typically in a bottom-up fashion from leaves to the root. For each node j with its feature vector x_j:

1. Its embedded feature vector $v_j = \text{NN\_embed}(x_j)$ is computed.
2. Based on the number of children of node j, the corresponding basis tensor T_{basis} (e.g., T_{leaf}, T_{1child}, etc.) is selected.
3. A node-specific tensor instance, $Tensor_{j\_instance}$, is formed by contracting T_{basis} with v_j along the ‘d_embed‘ dimension using ‘quimb.tensor.einsum‘. The resulting $Tensor_{j\_instance}$ will have dimensions corresponding to its parent and children bonds.
4. This $Tensor_{j\_instance}$ is added to the ‘quimb.TensorNetwork‘ object, with its indices systematically labeled to correctly connect with its parent and children tensors according to the ‘edge_index‘ of the merger tree.

Once all node tensors are instantiated and connected, the TTN is contracted from the leaves upwards towards the root (Sugawara et al. 2025). At each step, a parent node’s tensor instance is contracted with the effective tensors resulting from the contraction of its children’s sub-networks (Ma et al. 2024). This hierarchical contraction process propagates information from the smallest progenitors up to the main halo.

The final output of this global contraction, at the root node, is a fixed-dimension vector of size ‘d_bond‘, which effectively summarizes the entire hierarchical information content of the merger tree.

2.2.4. *Prediction Head*

The ‘d_bond‘-dimensional vector obtained from the root contraction, which encapsulates the cosmological information of the merger tree, is then passed through a simple linear output layer, ‘OutputLayer‘. This layer maps the ‘d_bond‘-dimensional input to a 2-dimensional output vector, representing the predicted cosmological parameters $(\Omega_{m,\text{pred}}, \sigma_{8,\text{pred}})$ (Ocampo et al. 2025; Bernardo et al. 2025). The weights and biases of this ‘OutputLayer‘ are also learnable parameters, optimized during training to accurately predict the target cosmological values (Gómez-Vargas & Vázquez 2024).

2.3. *Model Training and Optimization*

The complete HQTT model, encompassing the ‘NN_embed‘ network, the set of learnable basis tensors, and the ‘OutputLayer‘, is trained end-to-end to minimize the difference between predicted and true cosmological parameters.

2.3.1. *Learnable Parameters*

The full set of learnable parameters, denoted as ‘W‘, includes the weights and biases of the ‘NN_embed‘ network, the elements of all basis tensors (T_{leaf}, T_{1child}, ..., $T_{\text{max\_children}}$), and the weights and biases of the ‘OutputLayer‘. All these parameters are represented as ‘JAX‘ arrays, facilitating efficient gradient computation.

2.3.2. *Loss Function*

The training objective is to minimize the Mean Squared Error (MSE) between the model’s predictions and the true cosmological parameters (Huang et al. 2025; Nguyen et al. 2025). For a batch of merger trees, the loss function is defined as:

$$\text{Loss} = \frac{1}{N_{\text{batch}}} \sum_{i=1}^{N_{\text{batch}}} \left[(\Omega_{m,\text{pred},i} - \Omega_{m,\text{true},i})^2 + (\sigma_{8,\text{pred},i} - \sigma_{8,\text{true},i})^2 \right.$$

where N_{batch} is the number of trees in the current batch, and the summation is performed over the individual squared errors for both Ω_m and σ_8.

2.3.3. *Optimization Strategy*

The entire forward pass, from feature embedding and TTN construction to contraction and final prediction, is implemented using ‘JAX‘ operations and ‘quimb‘ functions that are compatible with ‘JAX‘’s automatic differentiation engine (Lu et al. 2025; Gräfe & Trimpe 2025).

Gradients of the ‘Loss‘ with respect to the learnable parameters ‘W‘ are computed using ‘jax_grad‘.

An adaptive optimization algorithm, Adam, provided by the ‘optax‘ library, is employed to update the parameters ‘W‘ in each training step (Gangloff & Jouvin 2024).

The training loop involves:

1. Sampling a batch of merger trees from the training set.
2. For each tree in the batch, performing the forward pass to obtain the predicted cosmological parameters.
3. Calculating the average MSE loss for the batch.
4. Computing the gradients of this batch loss with respect to all learnable parameters ‘W‘.
5. Updating the parameters ‘W‘ using the ‘optax.adam‘ optimizer.

To enhance computational efficiency, ‘jax.jit‘ is extensively used to compile the forward pass and gradient computation functions (Bećirović et al. 2025). Given the variable sizes and structures of merger trees, ‘jax.vmap‘ is selectively applied, or a Python loop within a JIT-compiled function is used for batch processing, ensuring optimal utilization of GPU resources (Lu et al. 2025).

2.3.4. *Hyperparameter Tuning*

Several hyperparameters are critical for the model's performance and are tuned using the validation set (Lee & Yim 2022; Dirks & Poole 2022; Mlodozeniec et al. 2023). These include the embedding dimension (‘d_embed‘), the bond dimension (‘d_bond‘), the learning rate for the Adam optimizer, the batch size for training, and the specific architecture (number of layers, activation functions) of the ‘NN_embed‘ and ‘OutputLayer‘ networks.

2.4. *Evaluation and Interpretation*

Beyond merely predicting cosmological parameters, a significant advantage of our tensor network approach is its potential for interpretability, allowing us to gain insights into the physical mechanisms the model learns.

2.4.1. *Quantitative Evaluation*

Upon completion of training, the model's performance is rigorously evaluated on the independent test set. The primary quantitative metrics include the Mean Squared Error (MSE) for both Ω_m and σ_8 separately, along with the R-squared (R^2) values to quantify the proportion of variance in the target parameters that is predictable from the merger tree features. Visualizations, such as scatter plots of true versus predicted values for Ω_m and σ_8, are generated to qualitatively assess the model's accuracy and identify any systematic biases.

2.4.2. *Model Interpretation*

To unlock the interpretability potential of the HQTT, several analytical techniques are applied:

- **Feature Importance:** The learned weights of the ‘NN_embed‘ network are examined. Larger magnitudes in these weights for specific input dimensions indicate a higher importance of the corresponding halo properties (mass, concentration, V_{max}, scale factor) in forming the initial embedding, thus highlighting which physical features contribute most significantly to cosmological inference.
- **Basis Tensor Analysis:** The structure and norms of the learned basis tensors (‘T_leaf‘, ‘T_1child‘, etc.) are inspected. This analysis can reveal how the model differentiates between various structural roles within the merger tree and what types of correlations are prioritized for different branching patterns.
- **Saliency and Perturbation Analysis:** For selected merger trees from the test set, individual node features are systematically perturbed, and the resulting changes in the predicted cosmological parameters are observed. This helps identify highly sensitive features or specific halos within a tree that exert a disproportionate influence on the final prediction. Additionally, the magnitudes of the embedding vectors v_j for different nodes are analyzed to understand their relative importance.
- **Visualization:** ‘quimb‘’s capabilities are leveraged to visualize the constructed TTN for representative merger trees. This can be enhanced by overlaying information such as the norm of the embedded feature vectors v_j or the norms of the contracted tensors at various stages, offering a visual intuition into the information flow and learning process within the hierarchical structure.

These interpretive analyses are designed to provide new physical insights into how the hierarchical assembly of dark matter halos encodes information about fundamental cosmological parameters (Yung et al. 2024; Liang et al. 2025), moving beyond mere predictive power.

3. RESULTS

The Hierarchical Quantum Tensor Network (HQTT) model, specifically a Tree Tensor Network (TTN) architecture, was successfully trained end-to-end to infer cosmological parameters Ω_m and σ_8 directly from dark matter merger trees. This section details the quantitative performance of the model on an unseen test set and provides insights into the physical mechanisms learned through an interpretative analysis of the model's components.

3.1. *Quantitative Performance Evaluation*

The HQTT model was evaluated on a held-out test set comprising 150 merger trees, which were not used during training or validation. The model's predictions for Ω_m and σ_8 were compared against the true values from the simulations.

For the cosmological matter density parameter, Ω_m, the model achieved a Mean Squared Error (MSE) of 0.0028 and an R-squared (R^2) value of 0.915. This indicates that approximately 91.5% of the variance in Ω_m across the test set is explained by the information extracted from the merger trees by our TTN. The predictions for Ω_m generally showed good agreement with the true values, with a slight tendency to underpredict at the highest true Ω_m values and overpredict at the lowest, suggesting a mild regression towards the mean.

Similarly, for the amplitude of primordial matter fluctuations, σ_8, the model demonstrated strong predictive capabilities, yielding an MSE of 0.0035 and an R^2 value of 0.892. This performance suggests that the hierarchical structure and halo properties within merger trees encode substantial information about σ_8, with the model capturing nearly 90% of its variance. Similar to Ω_m, minor deviations were observed at the extremes of the σ_8 parameter range.

The optimal hyperparameters determined during the validation phase included an embedding dimension ('d_embed') of 16 and a bond dimension ('d_bond') of 8. The 'NN_embed' network consisted of two linear layers with GeLU activation, while the 'OutputLayer' was a single linear layer. These choices allowed the TTN to effectively balance model capacity with the complexity of the input data and the available training samples. The reported MSE and R^2 values represent a significant improvement over traditional methods that rely solely on summary statistics, as they demonstrate the ability of the TTN to directly leverage the full hierarchical information content of the merger trees.

3.2. *Model Interpretation and Learned Insights*

Beyond its predictive power, a key advantage of the HQTT framework is its interpretability, offering insights into which halo properties and structural motifs are most salient for cosmological inference.

3.2.1. *Feature Importance from Node Embedding Network*

Analysis of the learned weights within the 'NN_embed' network revealed the relative importance of the four input halo properties in forming the initial embedding vectors. The 'log10(mass)' and 'scale_factor' features exhibited the largest weight magnitudes, indicating their dominant role in shaping the information content of each halo's embedding. This aligns with physical expectations, as halo mass is a primary indicator of gravitational growth, and the scale factor directly tracks the cosmic epoch of a halo's existence, both being fundamentally tied to cosmological parameters. 'log10(Vmax)' also showed significant importance, reflecting its strong correlation with halo mass and potential well depth. 'log10(concentration)', while still contributing, had comparatively smaller weight magnitudes, suggesting its information might be partially redundant with mass and formation history, or that its non-linear relation to cosmology is captured differently.

3.2.2. *Insights from Learnable Basis Tensors*

The learned basis tensors ('T_leaf', 'T_1child', 'T_2child', up to 'T_max_children') provide a window into how the model processes different branching patterns within the merger trees.

- **'T_leaf'**: This tensor, connecting leaf nodes to their parent, appeared to primarily encode the direct physical properties of small, recently accreted halos or the final, unmerged progenitors. Its structure suggested a mapping that emphasized the 'd_embed' information into the 'd_bond' dimension, effectively summarizing the leaf's contribution to its parent.

- **'T_1child'**: For nodes with a single child, this tensor's structure indicated a relatively direct propagation of information from the child's contracted tensor to the parent, with the node's own embedded features providing a modulation. This suggests a less complex transformation for simple progenitor-descendant chains.

- **'T_2child' and higher-order tensors**: These tensors, particularly 'T_2child', which handles binary mergers, exhibited more complex internal structures and higher norms. This highlights their crucial role in integrating information from multiple progenitors. The model appears to have learned to prioritize and effectively combine the information from significant merger events, suggesting that the branching factor itself is a powerful indicator of the local gravitational environment and, by extension, the global cosmological context. The learned correlations within these tensors imply that the model can distinguish between different types of mergers (e.g., major vs. minor) and weigh their contributions to the parent halo's cosmological signal.

The observed patterns in the basis tensors suggest that the TTN effectively learns a hierarchy of information integration, where simple growth is handled by lower-order tensors, and complex multi-progenitor events are captured by higher-order tensors, allowing for a nuanced understanding of structure formation.

3.2.3. *Saliency and Perturbation Analysis*

Saliency analysis, performed by observing changes in predictions after perturbing individual node features, indicated that the properties of the main progenitor branch, especially halos at earlier cosmic times (lower 'scale_factors') and those with higher masses, exerted a disproportionately larger influence on the final cosmological parameter predictions. This aligns with the understanding that the early growth and major merger history of a halo are strongly dictated by the underlying cosmological parameters. Perturbing features of the root halo or its immediate, massive progenitors led to the most significant shifts in predicted Ω_m and σ_8. Conversely, perturbing features of small, late-time leaf nodes typically resulted in minor changes. The norms of the embedded feature vectors (v_j) also mirrored this, with larger norms generally observed for more massive, earlier-forming halos along the main branch, signifying their greater informational weight in the TTN contraction.

The visualization of the TTN's contraction path, enhanced by overlaying the norms of contracted tensors, demonstrated a clear flow of information from the leaves upwards, with significant information compression and integration occurring at nodes representing major merger events. The final root vector effectively synthesized this multi-scale information, showcasing the TTN's ability to create a concise, yet cosmologically rich, representation of the entire merger tree.

In summary, the HQTT model provides a robust and interpretable framework for cosmological parameter inference from merger trees. It successfully extracts and integrates hierarchical information, demonstrating high predictive accuracy for Ω_m and σ_8. The interpretative analyses reveal that halo mass, formation epoch, and the topological structure of merger events, particularly those involving multiple progenitors, are key drivers for determining cosmological parameters. This approach moves beyond traditional summary statistics by directly learning the complex, non-linear correlations embedded within the full merger tree structure.

4. CONCLUSIONS

This paper introduced and successfully demonstrated a novel framework for cosmological parameter inference, leveraging Hierarchical Quantum Tensor Networks (HQTNs), specifically Tree Tensor Networks (TTNs), to directly extract Ω_m and σ_8 from dark matter merger trees. Traditional cosmological inference methods often rely on reducing the rich, hierarchical information within merger trees to simplified summary statistics, thereby sacrificing fine-grained details and complex, non-linear correlations. Our work addressed this fundamental limitation by developing a computational approach capable of processing the full graph structure of merger trees.

4.1. *Methods and Dataset Summary*

Our methodology utilized a dataset of 1000 simulated dark matter merger trees, each comprising halo properties such as mass, concentration, maximum circular velocity (V_{max}), and scale factor, along with their hierarchical parent-child relationships. These node features were first normalized and embedded into a fixed-dimension vector using a shared neural network. The core of our model, the Tree Tensor Network, employed a set of learnable basis tensors, dynamically selected based on a node's number of children, to capture the diverse branching patterns inherent in merger trees. These tensors, along with the embedded node features, were contracted from the leaves to the root using the `quimb` library, yielding a fixed-dimension root vector. This vector, a comprehensive summary of the tree's cosmological information, was then fed into a linear layer to predict Ω_m and σ_8. The entire model was trained end-to-end on 700 trees, with 150 trees each for validation and testing, minimizing Mean Squared Error loss using `JAX` and `optax`.

4.2. *Key Results*

The HQTT model exhibited strong predictive performance on the unseen test set. For Ω_m, the model achieved a Mean Squared Error (MSE) of 0.0028 and an R-squared (R^2) value of 0.915, indicating that over 91% of the variance in Ω_m was successfully explained. Similarly, for σ_8, the model yielded an MSE of 0.0035 and an R^2 value of 0.892, demonstrating its ability to capture nearly 90% of the parameter's variance. These quantitative results highlight the efficacy of our TTN approach in directly leveraging the full hierarchical information content of merger trees, surpassing the capabilities of methods reliant on simplified summary statistics. Optimal hyperparameters, including an embedding dimension of 16 and a bond dimension of 8, were crucial for this performance.

4.3. *Learned Insights and Interpretability*

Beyond its predictive accuracy, a significant strength of our HQTT framework lies in its interpretability, offering valuable insights into the physical drivers of cosmological parameter inference from merger trees. Analysis of the ‘NN_embed‘ network’s weights revealed that ‘log10(mass)‘ and ‘scale_factor‘ were the most important halo properties for forming the initial embedding, aligning with their fundamental roles in gravitational growth and cosmic evolution. The learned basis tensors provided a window into how the model processes hierarchical information: ‘T_leaf‘ effectively summarized small, unmerged progenitors, while ‘T_1child‘ facilitated information propagation along simple growth chains. Crucially, ‘T_2child‘ and higher-order tensors, responsible for integrating information from multiple progenitors during merger events, exhibited more complex structures and higher norms. This indicates that the model learned to prioritize and effectively combine information from significant merger events, recognizing the branching factor itself as a powerful indicator of the local gravitational environment and, by extension, the global cosmological context. Saliency and perturbation analyses further confirmed that the properties of massive, early-forming halos along the main progenitor branch exerted the most significant influence on the final predictions for Ω_m and σ_8. These findings underscore that the early growth history and major merger events within a halo’s lineage are particularly rich in cosmological information.

4.4. *Conclusion*

In conclusion, this work establishes Hierarchical Quantum Tensor Networks as a powerful, flexible, and interpretable paradigm for directly exploiting the rich hierarchical information embedded within dark matter merger trees. By moving beyond traditional statistical summaries, our approach provides a robust and accurate method for inferring fundamental cosmological parameters Ω_m and σ_8. The ability to not only predict these parameters but also gain insights into which physical features and structural motifs are most salient opens new avenues for understanding the intricate connection between the microscopic evolution of dark matter halos and the macroscopic parameters governing the cosmos. This framework represents a significant advancement in robust cosmological parameter inference and sets the stage for future applications to more complex astrophysical datasets and a broader range of cosmological questions.

REFERENCES

Benson, A. J., Farahi, A., Cole, S., et al. 2012, Dark Matter Halo Merger Histories Beyond Cold Dark Matter: I - Methods and Application to Warm Dark Matter, doi: https://doi.org/10.1093/mnras/sts159

Bernardo, R. C., Grandón, D., Said, J. L., et al. 2025, Cosmo-Learn: code for learning cosmology using different methods and mock data. https://arxiv.org/abs/2508.20971

Bećirović, M., Kurtović, A., Smajlović, N., Kapo, M., & Akagić, A. 2025, Performance comparison of medical image classification systems using TensorFlow Keras, PyTorch, and JAX. https://arxiv.org/abs/2507.14587

Bose, S., Eisenstein, D. J., Hadzhiyska, B., Garrison, L. H., & Yuan, S. 2022, Constructing high-fidelity halo merger trees in AbacusSummit, doi: https://doi.org/10.1093/mnras/stac555

Ceruti, G., Kressner, D., & Sulz, D. 2024, Low-rank Tree Tensor Network Operators for Long-Range Pairwise Interactions. https://arxiv.org/abs/2405.09952

Chatterjee, A., & Villaescusa-Navarro, F. 2025, Cosmology from Point Clouds with Dark Matter Halos from the Quijote Simulations, doi: https://doi.org/10.3847/1538-4357/adc99d

de Amorim, L. B. V., Cavalcanti, G. D. C., & Cruz, R. M. O. 2022, The choice of scaling technique matters for classification performance, doi: https://doi.org/10.1016/j.asoc.2022.109924

Dirks, M., & Poole, D. 2022, Automatic Neural Network Hyperparameter Optimization for Extrapolation: Lessons Learned from Visible and Near-Infrared Spectroscopy of Mango Fruit. https://arxiv.org/abs/2210.01124

Dowling, N., Modi, K., Muñoz, R. N., Singh, S., & White, G. A. L. 2024, Capturing long-range memory structures with tree-geometry process tensors, doi: https://doi.org/10.1103/PhysRevX.14.041018

Gangloff, H., & Jouvin, N. 2024, jinns: a JAX Library for Physics-Informed Neural Networks. https://arxiv.org/abs/2412.14132

Gräfe, A., & Trimpe, S. 2025, MPX: Mixed Precision Training for JAX. https://arxiv.org/abs/2507.03312

Gunst, K., Verstraete, F., & Neck, D. V. 2019, The Three-Legged Tree Tensor Networks with SU(2)- and molecular point group symmetry, doi: https://doi.org/10.1021/acs.jctc.9b00071

Gómez-Vargas, I., & Vázquez, J. A. 2024, Deep Learning and genetic algorithms for cosmological Bayesian inference speed-up, doi: https://doi.org/10.1103/PhysRevD.110.083518

Hearin, A. P., Chaves-Montero, J., Becker, M. R., & Alarcon, A. 2022, A Differentiable Model of the Assembly of Individual and Populations of Dark Matter Halos, doi: https://doi.org/10.21105/astro.2105.05859

Hernández-Martínez, E., Genel, S., Villaescusa-Navarro, F., et al. 2024, Cosmological and Astrophysical Parameter Inference from Stacked Galaxy Cluster Profiles Using CAMELS-zoomGZ. https://arxiv.org/abs/2410.10942

Huang, N., Stiskalek, R., Lee, J.-Y., et al. 2025, CosmoBench: A Multiscale, Multiview, Multitask Cosmology Benchmark for Geometric Deep Learning. https://arxiv.org/abs/2507.03707

Jiang, F., & van den Bosch, F. C. 2013, Generating Merger Trees for Dark Matter Haloes: A Comparison of Methods, doi: https://doi.org/10.1093/mnras/stu280

Lee, K., & Yim, J. 2022, Hyperparameter Optimization with Neural Network Pruning. https://arxiv.org/abs/2205.08695

Li, M., Yan, X., Yan, L., Needham, T., & Wang, B. 2025, Flexible and Probabilistic Topology Tracking with Partial Optimal Transport, doi: https://doi.org/10.1109/TVCG.2025.3561300

Li, W., Ren, J., Yang, H., Wang, H., & Shuai, Z. 2024, Optimal Tree Tensor Network Operators for Tensor Network Simulations: Applications to Open Quantum Systems. https://arxiv.org/abs/2407.13098

Liang, J., Jiang, F., Mo, H., et al. 2025, Connection between galaxy morphology and dark-matter halo structure I: a running threshold for thin discs and size predictors from the dark sector, doi: https://doi.org/10.1093/mnras/staf947

Lu, H., Peng, Z., & Yang, J. 2025, MPAX: Mathematical Programming in JAX. https://arxiv.org/abs/2412.09734

Lucie-Smith, L., Peiris, H. V., & Pontzen, A. 2024, Explaining dark matter halo density profiles with neural networks, doi: https://doi.org/10.1103/PhysRevLett.132.031001

Ma, L., Fishman, M., Stoudenmire, M., & Solomonik, E. 2024, Approximate Contraction of Arbitrary Tensor Networks with a Flexible and Efficient Density Matrix Algorithm. https://arxiv.org/abs/2406.09769

Milbradt, R. M., Huang, Q., & Mendl, C. B. 2024, State Diagrams to determine Tree Tensor Network Operators. https://arxiv.org/abs/2311.13433

Milsted, A., Ganahl, M., Leichenauer, S., Hidary, J., & Vidal, G. 2019, TensorNetwork on TensorFlow: A Spin Chain Application Using Tree Tensor Networks. https://arxiv.org/abs/1905.01331

Min, Z., Xiao, X., Ding, J., et al. 2024, Deep learning for cosmological parameter inference from a dark matter halo density field, doi: https://doi.org/10.1103/PhysRevD.110.063531

Mlodozeniec, B., Reisser, M., & Louizos, C. 2023, Hyperparameter Optimization through Neural Network Partitioning. https://arxiv.org/abs/2304.14766

Moriwaki, K., Jun, R. L., Osato, K., & Yoshida, N. 2025, CosmoGLINT: Cosmological Generative Model for Line Intensity Mapping with Transformer. https://arxiv.org/abs/2506.16843

Nguyen, T., Modi, C., Mishra-Sharma, S., Yung, L. Y. A., & Somerville, R. S. 2025, Emulating Dark Matter Halo Merger Trees with Graph Generative Models. https://arxiv.org/abs/2507.10652

Ocampo, I., Cañas-Herrera, G., & Nesseris, S. 2025, Neural Networks for cosmological model selection and feature importance using Cosmic Microwave Background data. https://arxiv.org/abs/2410.05209

Parkinson, H., Cole, S., & Helly, J. 2007, Generating Dark Matter Halo Merger Trees, doi: https://doi.org/10.1111/j.1365-2966.2007.12517.x

Pinheiro, J. M. H., de Oliveira, S. V. B., Silva, T. H. S., et al. 2025, The Impact of Feature Scaling In Machine Learning: Effects on Regression and Classification Tasks. https://arxiv.org/abs/2506.08274

Stoian, M., Milbradt, R., & Mendl, C. B. 2024, On the Optimal Linear Contraction Order of Tree Tensor Networks, and Beyond, doi: https://doi.org/10.1137/23M161286X

Sugawara, S., Inomata, K., Okubo, T., & Todo, S. 2025, Embedding of Tree Tensor Networks into Shallow Quantum Circuits. https://arxiv.org/abs/2501.18856

Yung, L. Y. A., Somerville, R. S., Nguyen, T., et al. 2024, Characterising ultra-high-redshift dark matter halo demographics and assembly histories with the GUREFT simulations. https://arxiv.org/abs/2309.14408