INTERNATIONAL JOURNAL OF INNOVATIVE COMPUTING
ISSN 2180-4370
Journal Homepage : https://ijic.utm.my/

Theoretical Foundations and Applications of Space Complexity in Distributed and Geometric Computation

Liu Shiqian[1] & Yusliza Yusoff[2*]
Faculty of Computing
Universiti Teknologi Malaysia
81310 UTM Johor Bahru, Johor, Malaysia
Email: liushiqian@graduate.utm.my[1]; yusliza@utm.my[2]

Submitted: 17/6/2025. Revised edition: 14/12/2025. Accepted: 5/1/2026. Published online: 10/6/2026
DOI: https://doi.org/10.11113/ijic.v16n1.557

***Abstract*—Space Complexity. This paper presents a formal examination of space complexity across distributed consensus, geometric clustering, decision-tree inference, and functional convolution. Drawing on five recent research contributions, the study unifies theoretical models and structural bounds to characterize minimal spatial requirements under diverse computational frameworks. The analysis employs indistinguishability and valency arguments, coreset-based compression theory, functional-space metrics, and convolutional operations. The results establish several tight lower bounds and structural equivalences, providing methodological insights for the design of algorithms operating under stringent memory constraints.**

***Keywords*—Space complexity, resource-constrained algorithmic design, indistinguishability, coreset compression theory**

I. INTRODUCTION

Space complexity is a fundamental metric in theoretical computer science, capturing the minimum amount of memory or storage required to solve a computational problem as a function of the input size. It provides essential insight into the efficiency and feasibility of algorithms, particularly in environments constrained by memory resources. Although time complexity has long dominated algorithmic analysis, space complexity has gained increasing prominence in recent years, driven in part by the growing need for scalable computation over high-dimensional data, distributed systems, and knowledge-based decision models.

Classical computational frameworks offer a natural foundation for investigating questions related to space usage. These include the read/write shared-memory model in distributed computing, clustering formulations in geometric learning, and tree-based representations in decision systems. Each framework introduces its own structural and operational constraints that shape how space complexity arises and how it may be bounded or optimized. Meanwhile, contemporary theoretical advances—such as indistinguishability arguments, covering-based lower-bound techniques, and quantization-aware compression methods—have deepened our understanding of minimal space requirements across these domains.

This study synthesizes and formalizes insights from five recent and influential research contributions that examine space complexity from multiple perspectives. The first set of results concerns obstruction-free consensus in distributed systems. Here, the computational power of shared-memory objects—such as swap objects and readable-swap variants—is analyzed, with tight lower bounds established through valency arguments and covering configurations. These findings illuminate fundamental limits on the memory necessary to coordinate decisions among asynchronous processes in the absence of failures.

The second line of research focuses on the Euclidean ((k, z))-clustering problem, a central task in high-dimensional geometric learning. We examine the concept of an ε-sketch, which provides a succinct approximation of clustering cost functions, and analyze its bit-level storage requirements. By combining coreset constructions with terminal-embedding techniques, the results demonstrate that even approximate settings demand substantial space, and that dimensionality reduction offers only limited relief.

The complexity of decision-making is examined through deterministic and nondeterministic decision-tree models. These models are assessed in terms of their depth and node growth, yielding a taxonomy of complexity classes defined by functional growth bounds. This classification supports more nuanced space–time tradeoff analyses and illuminates the inherent costs associated with interpretability and completeness in rule-based systems.

We further analyze convolution operations in complexity-function spaces as mechanisms for aggregating and optimizing algorithmic complexity. When formulated over weighted quasi-metric structures, these operations give rise to topological and algebraic insights that inform the modular construction of algorithms and their corresponding resource profiles.

Although diverse in formulation, these domains converge on several shared theoretical concerns: identifying space-optimal representations, establishing tight lower bounds, and characterizing the underlying algebraic structure of complexity under resource constraints. By systematically integrating these research directions, this paper seeks to present a unified perspective on space complexity as both a limiting factor and a guiding design principle across modern computational paradigms.

II. BACKGROUND

Before undertaking a rigorous analysis of space complexity, it is important to clarify its manifestations and foundational challenges across different computational paradigms. Space complexity—defined as the minimum amount of memory required to solve a computational problem—serves not only as a key indicator of algorithmic efficiency but also as a primary bottleneck in many classical settings, particularly in distributed computing, combinatorial decision systems, and high-dimensional data processing. This section reviews the current state of research and major theoretical developments in three representative domains: distributed consensus and historyless objects, where memory constraints shape the feasibility of coordination; decision trees and infinite binary information systems, which illuminate structural limits in query-based computation; and Euclidean clustering and compression bounds, which highlight the tension between geometric accuracy and memory efficiency in high-dimensional spaces.

A. Distributed Consensus and Historyless Objects

The consensus problem is a central challenge in the theory of distributed systems. It requires a collection of asynchronous processes to agree on a common value despite delays, reordering, or crash failures. This challenge underlies numerous practical applications, including fault-tolerant database replication, blockchain consensus mechanisms, and distributed transaction processing. The seminal impossibility result of Fischer, Lynch, and Paterson (1985) established that deterministic wait-free consensus is impossible for two or more processes in a purely asynchronous system, thereby catalyzing extensive research on relaxed progress conditions.

To circumvent the FLP impossibility, models such as randomized consensus and obstruction-free consensus have been introduced. These models allow limited nondeterminism or impose conditional progress assumptions, creating settings in which consensus becomes solvable. Such relaxations naturally prompt a complementary line of inquiry: determining the minimal space requirements—specifically, the number and type of shared-memory objects—necessary to implement these weaker forms of consensus.

In this context, Ovens (2022; 2024) conducted a detailed investigation into consensus solvability using historyless objects. A historyless object is defined by the property that its state is fully determined by its most recent non-trivial operation (write, test-and-set, or swap), independent of the sequence of preceding operations. Representative examples include read/write registers, swap objects, and test-and-set objects.

Using a combination of covering arguments and indistinguishability techniques, Ovens established the following results: Any obstruction-free consensus protocol for (n) processes using swap objects requires at least (n - 1) distinct swap instances. When restricted to readable binary swap objects—swap objects with a binary domain augmented with a read operation—the lower bound improves to (n - 2) objects.

Both lower bounds are matched by constructive algorithms, making the results asymptotically tight. These findings not only delineate the computational power of historyless objects but also provide baseline reference points for the spatial costs inherent in designing efficient consensus protocols.

B. Decision Trees and Infinite Binary Information Systems

Decision trees are widely used in classification, pattern recognition, and rule-based reasoning due to their conceptual simplicity and inherent interpretability. Classical algorithms such as ID3, C4.5, and CART construct decision trees by recursively partitioning the input space based on attribute values until a classification outcome is reached. However, as the attribute space grows—particularly in settings involving infinite or high-dimensional inputs—the time and space complexity of decision trees becomes increasingly significant.

Moshkov (2022) generalized traditional decision-tree models by introducing the notion of infinite binary information systems, in which the attribute set is countably infinite and each attribute assumes a binary value (0 or 1). Within this framework, the optimal decision tree for a given classification problem is analyzed along two dimensions: tree depth (the maximum length of any classification path) and node count (the total number of non-leaf nodes).

Key theoretical results include: In the worst case, the minimum depth of a deterministic decision tree grows either logarithmically or linearly with the number of attributes. The number of nodes required grows either polynomially or exponentially. These behavioral regimes collectively define five distinct complexity classes, forming a “complexity taxonomy” for information systems.

To formalize the corresponding trade-offs, Moshkov introduced the concept of a boundary (d)-pair—a pair of functions ($\varphi(n),\psi(n)$), where $\varphi(n)$ denotes the minimum achievable depth and $\psi(n)$ the minimum number of nodes for a problem defined over (n) attributes. This formulation captures the time–space optimality frontier for problems within a given information system. The resulting framework provides analytical tools for algorithmic optimization, as well as theoretical foundations for model compression and structural complexity control in machine learning.

C. Euclidean Clustering and Compression Bounds

The proliferation of high-dimensional data in machine learning and data mining has renewed interest in the space complexity of clustering algorithms. In particular, the ((k, z))-clustering problem seeks to partition (n) points in Rd into (k) clusters so as to minimize the sum of the (z)-th powers of the Euclidean distances between each point and its nearest cluster center. The well-known (k)-median and (k)-means problems correspond to the cases ($z = 1$) and ($z = 2$), respectively.

Zhu *et al*. (2025) initiated a rigorous investigation of the space complexity of clustering under compression constraints by introducing the notion of an ε-coreset—a weighted subset of the data that approximates the clustering cost within a factor of ($1\pm\varepsilon$) for all possible center configurations. Their key contributions can be summarized as follows: For a fixed approximation parameter ε, the size of the coreset is independent of the dataset size (n) and depends only on (k), ε, and (z). The use of ε-coresets substantially reduces storage overhead while preserving convergence behavior and stability in downstream clustering algorithms. Most critically, they establish a lower bound on the space complexity of clustering approximations: even when dimensionality reduction techniques such as the Johnson–Lindenstrauss (JL) embedding are employed, the coreset must incur a bit complexity of ($\Omega(|S| \cdot d)$), where ($|S|$) denotes the coreset size.

This lower bound reflects the fact that, although embeddings can reduce ambient dimensionality, the embedding map itself must preserve sufficient geometric information, which necessitates explicit storage. Consequently, in the common regime where (k) is constant and (d) is large, ε-coresets are shown to be theoretically optimal for spatial compression in clustering.

These findings indicate that traditional techniques such as naive dimensionality reduction or random sampling are insufficient for scalable clustering. Instead, compression strategies grounded in geometric preservation and approximation theory are required to achieve effective control of storage complexity in high-dimensional spaces.

III. METHODOLOGY

To systematically investigate the theoretical boundaries and constructive mechanisms of space complexity across diverse computational models, this study integrates model abstraction, lower-bound derivation, compression-structure analysis, and functional-space methods. The scope of investigation spans obstruction-free consensus protocols, Euclidean clustering, decision-tree models, and the convolutional structure of complexity spaces. Grounded in computational complexity theory, the study formalizes minimal spatial requirements under worst-case assumptions and identifies optimal or near-optimal representations.

In distributed computing, the focus is on determining the minimal number of historyless shared-memory objects required to solve binary obstruction-free consensus. Through indistinguishability configurations, covering arguments, and valency-based reasoning, the study formally establishes that at least ($n - 1$) swap objects are necessary to achieve obstruction-free consensus among (n) processes. For readable binary swap objects, the lower bound tightens to ($n - 2$), matching known upper bounds asymptotically. These results further extend to the (k)-set agreement problem: solving obstruction-free (k)-agreement requires at least ($\lceil n/k \rceil - 1$) swap objects, thereby generalizing classical consensus lower bounds to broader agreement settings.

In the domain of high-dimensional data processing, the space complexity of the Euclidean ((k, z))-clustering problem is examined under ε-approximate constraints using coreset constructions and terminal-embedding techniques. An ε-sketch is defined as a compact data structure that approximates clustering cost to within a ($1\pm\varepsilon$) factor. Under a discrete grid assumption $P \subseteq [\Delta]d$, geometric inequalities and quantization arguments demonstrate that any such structure must require $\Omega(kd\log\Delta)$ bits. Moreover, due to the inherent cost of encoding embedding mappings, dimensionality reduction techniques cannot circumvent this lower bound, thereby confirming that quantized coresets are asymptotically optimal under practical parameter regimes.

For decision-based inference models, this study analyzes deterministic and nondeterministic decision trees constructed over infinite binary information systems. Tree depth and node count serve as the respective metrics for time and space complexity. A complexity-function pair ($\varphi(n),\psi(n)$) is defined, where $\varphi(n)$ denotes the minimum achievable depth and $\psi(n)$ denotes the minimum number of nodes required to solve a problem involving (n) attributes. The resulting space of information systems is partitioned into five distinct classes, each characterized by asymptotic behaviors such as logarithmic, linear, polynomial, or exponential growth. These classifications support rigorous evaluation of worst-case spatial requirements and illuminate the intrinsic trade-offs between depth minimization and structural complexity in decision-based models.

The study further examines convolution operations in the complexity space (C) and its dual C*, treating them as mechanisms for complexity composition and functional optimization. Within the framework of weighted quasi-metric spaces, convolution is shown to preserve quasi-metric continuity and order consistency. A class of improver functionals is constructed via convolution, enabling space-efficient optimization of complexity measures. Additionally, in C*, convolution under three distinct topological structures—including the quasi-metric topology—is proven to yield a topological monoid under appropriate conditions. This result establishes the algebraic and topological foundations for the structured composition and optimization of complexity functions.

Together, these methodological components formalize space-complexity analysis across diverse computational domains through a unified theoretical lens. In consensus protocols, covering configurations and solo-termination arguments establish tight lower bounds; in clustering, ε-coreset theory and embedding-based quantization characterize compression limits; in decision logic, function-pair–based classification enables stratified estimation of spatial requirements; and in complexity-function spaces, convolution and topological structure inform higher-level optimization. Collectively, these methods provide foundational insights into the space-efficient realization of computation under constrained resources.

IV. RESULT AND ANALYSIS

Based on the aforementioned three types of models, this section respectively analyzes the theoretical lower bound, upper bound of space complexity and its compactness conclusion in the scenarios of consistency, decision tree modeling and Euclidean clustering.

A. The Spatial Lower Bound of the Consistency Problem

The consensus problem is a classical challenge in the theory of distributed systems. Its objective is to ensure that, in a completely asynchronous and clock-free environment, all non-faulty processes eventually agree on a common value, even in the presence of process delays or crash failures. The absence of global timing guarantees and synchronization mechanisms makes reliable agreement inherently difficult. The seminal FLP impossibility theorem (Fischer, Lynch, and Paterson, 1985) established that in such an environment, no deterministic wait-free consensus algorithm can be implemented using only basic read/write registers.

To circumvent this impossibility, researchers have introduced several relaxed models, including randomized algorithms and progress conditions such as non-blocking or obstruction-freedom. While these models weaken the guarantees of progress, they also raise a new theoretical question: under these more permissive semantics, what is the minimum amount of space required to achieve consensus. In particular, when the system is restricted to specific types of shared objects—such as swap or test-and-set objects—determining the precise limits of their expressive power has become a central topic in distributed computability.

Ovens (2022; 2024) conducted an in-depth formal investigation of this question, focusing on historyless objects—shared objects whose state depends solely on the most recent non-trivial operation and does not store any prior execution history. Typical examples include swap objects, which atomically exchange their stored value with an input value but do not support reading the current value. Within a shared-memory model constructed from such objects, Ovens applied a combination of constructive counterexamples and covering-based analysis to derive the following results: Implementing obstruction-free consensus among (n) processes requires at least (n - 1) swap objects. When using readable binary swap objects with domain ({0, 1}), the lower bound improves to (n - 2).

Both lower bounds are matched by corresponding upper-bound constructions—that is, explicit algorithms exist that achieve consensus using exactly (n - 1) or (n - 2) objects. The results are therefore asymptotically tight.

More importantly, Ovens does not rely on enumeration or ad hoc arguments but develops a reusable framework for proving lower bounds. The core tools include:

Indistinguishability Argument: This method is grounded in the observation that processes in certain global states cannot distinguish between different execution histories because the information they have observed is identical. From this indistinguishability, one can deduce that certain decision paths cannot be uniquely determined, implying that agreement cannot be reached under specific conditions.

Valency Analysis: This technique classifies system states as either univalent—leading inevitably to a single decision value—or bivalent—capable of leading to multiple possible outcomes. By analyzing how a system transitions from a bivalent to a univalent state, Ovens demonstrates that when the number of shared objects is insufficient, the system inevitably contains a “stuck” bivalent state, preventing some processes from reaching a safe and final decision.

Furthermore, this study broadens the applicability of classical lower-bound techniques. Under Ovens’ framework, even if the scheduling discipline satisfies obstruction-freedom—that is, any process running in isolation must eventually make progress—the previously established lower bounds remain unbreakable. This shows that even under significantly weakened progress requirements, there exists an intrinsic and unavoidable limit to the spatial expressive power of historyless objects in implementing consistent decision logic.

These results not only strengthen the theoretical foundations of space-complexity analysis but also carry practical implications. For example, in the design of resource-constrained distributed databases, lightweight consensus modules, or blockchain smart-contract runtimes, Ovens’ findings provide concrete guidance for architectural resource allocation, clearly identifying that—given specific correctness conditions—the minimum number of shared objects required constitutes a fundamental, non-compressible limit.

Overall, Ovens’ work advances the study of the space complexity of consensus from a focus on “upper-bound constructions” to a stage characterized by the precise characterization of tight lower bounds, marking a significant step forward in the theoretical development of distributed computability.

B. Decision Tree Complexity Classification

In information-system analysis and data-classification tasks, decision trees constitute a widely used and highly interpretable model class. Their classification process proceeds by recursively partitioning the attribute space, ultimately guiding each input object to a leaf node that determines the final decision. Compared with alternative models, decision trees possess the notable advantages of structural transparency and explicit decision paths. However, in practical applications, their construction often entails substantial space and time costs. This challenge becomes especially pronounced when the attribute set is large or the attribute domain exhibits high complexity, making rigorous complexity analysis both essential and urgent.

Moshkov (2022) introduced a theoretical framework for analyzing the time and space complexity of decision-tree models within infinite binary information systems. In this setting, the number of attributes may grow without bound—for example, in genomic sequences or sensor-stream data—while each attribute is constrained to take only a binary value (0 or 1). This restriction simplifies the combinatorial structure of the attribute space, yet simultaneously heightens the difficulty of achieving optimal structural design.

Within this framework, Moshkov defined four core functions to characterize the structural complexity of decision trees, with two being particularly central: Minimum tree depth φ(n): The minimum length of any decision path required to correctly classify all objects when (n) attributes are available. This measure corresponds to the time complexity of inference. Minimum number of nodes ψ(n): The minimum number of internal nodes required to achieve the same classification accuracy. This measure reflects the compactness of the decision-tree structure and serves as the primary indicator of space complexity.

Moshkov classified all possible growth behaviors of the pair (φ(n), ψ(n)) into five canonical complexity classes. These range from the favorable regime of logarithmic depth with polynomial node count to the extreme case of linear depth with exponential node count, corresponding to highly redundant structural representations. He further observed that different classification tasks and attribute-dependence structures naturally align with specific regions of this complexity spectrum. Consequently, this taxonomy provides a rigorous theoretical foundation for task analysis, model selection, and structural optimization in decision-based systems.

More importantly, Moshkov introduced a key theoretical tool: the boundary (d)-pair. This construct consists of a pair of functions (φ(n), ψ(n)) that jointly characterize the minimal time–space resources required by any discriminable information system with (n) attributes. The definition of boundary pairs is fully general, allowing the framework to encompass all possible information-system types and to quantify the optimal complexity attainable under combined resource constraints. This modeling approach not only enables rigorous specification of the worst-case complexity of classification systems but also provides a theoretical lower-bound foundation for practical optimization problems such as model compression and constrained deep learning.

Another significant strength of this framework lies in its scalability. Whether attributes are independent or exhibit interdependence, whether classification labels contain noise, or whether the system allows attribute-merging operations, Moshkov's analytical method remains applicable. This adaptability establishes a robust theoretical basis for extending the framework to domains such as expert-system construction, rule-based reasoning compression, and automated feature selection.

Overall, Moshkov's work on the complexity classification of decision trees deepens our understanding of the structural essence of decision-based inference models and offers precise structural boundaries and optimization guidelines for designing classification systems that balance interpretability with low computational overhead.

C. The Spatial Complexity of Euclidean Clustering

With the widespread use of high-dimensional data in areas such as intelligent manufacturing, biomedicine, and image recognition, achieving effective data compression without degrading clustering quality has become a central challenge in machine learning and algorithm design. Traditional clustering algorithms—such as (k)-means and (k)-median—exhibit rapid growth in both time and space complexity when applied to large-scale, high-dimensional datasets, thereby limiting their scalability in practical systems. As a result, researchers have increasingly turned to sketching and coreset-based approaches, seeking to compress input data efficiently while approximately preserving the original clustering structure.

Zhu *et al*. (2025) conducted a systematic analysis of the space-complexity limits of the ((k, z))-clustering problem in Euclidean space under data-compression constraints. Here, (k) denotes the number of cluster centers, and (z) specifies the error norm (e.g., (z = 2) corresponds to the classical (k)-means problem). Their central question is: for given data dimension (d), dataset size (n), and allowable approximation error, does there exist a compact data subset or compression structure that approximates the original clustering cost within the prescribed error tolerance.

Under this framework, their study first establishes a fundamental lower bound on space complexity. For a dataset in (d)-dimensional Euclidean space defined over an integer grid with input range Δ, any compression structure designed to serve as an ε-sketch must satisfy a bit-size lower bound proportional to Ω(kdlogΔ). This result is general, applying across different distance measures, data distributions, and approximation parameters ε. From an information-theoretic perspective, the bound indicates that any compressed representation must retain the essential geometric information governing the clustering cost function; thus, dependence on the data dimension is unavoidable.

To complement this lower bound, Zhu *et al*. proposed a coreset-based compression strategy, in which a carefully selected weighted subset (S) of the original data retains an approximate representation of the clustering cost across all possible center configurations. They further demonstrated that the space required by this method matches the near-optimal upper bound O(Ψ(n)·d), where Ψ(n) denotes the coreset size, which depends on (k), ε, the data dimension, and the chosen clustering error metric. Notably, this construction is not only theoretically optimal but also exhibits strong convergence properties and empirical stability across real-world datasets, making it one of the most practical and effective clustering-compression techniques available today.

Furthermore, the study provides a detailed examination of whether dimensionality-reduction techniques—most notably the Johnson–Lindenstrauss (JL) embedding—can reduce the space overhead of compressed clustering structures. Although JL projections can reduce dimensionality while approximately preserving pairwise distances, Zhu *et al*. demonstrate that such embeddings do not fundamentally break the established lower bound on space complexity. The reasons are twofold. First, JL embeddings require explicitly or implicitly storing the projection matrix, and when the original data dimension is large, the storage cost of this matrix becomes significant. Second, for certain input instances, the embedded data must still preserve intricate geometric and topological relationships of the original point set, making it impossible to eliminate the inherent dependence on the ambient dimension.

Based on these observations, Zhu *et al*. reach a clear conclusion: in practical scenarios where (k) is constant, the dimension (d) is large, and strict clustering-accuracy requirements are imposed, coresets represent the only theoretically feasible approach to achieving optimal spatial compression. This insight not only establishes a rigorous foundation for spatial optimization in high-dimensional clustering algorithms but also provides concrete guidance for system designers selecting data-compression mechanisms in machine learning workflows.

The applicability of this framework extends well beyond clustering alone. Its techniques can be adapted to other forms of structure-preserving compression—such as principal-component preservation, outlier detection, and geometric sketching—highlighting its broad theoretical and practical significance.

Table I presents a comparative overview of six recent studies related to space complexity across different research domains. The selected references cover topics including consensus algorithms, decision trees, Euclidean clustering, convolution complexity, numerical solvers, and ecological complexity analysis. Although these studies originate from diverse fields such as distributed computing, computational complexity theory, mathematics, scientific computing, and remote sensing, they all emphasize the importance of analyzing computational or structural complexity in space and time dimensions. The table also highlights the main contribution of each work, demonstrating the broad applicability and interdisciplinary significance of space complexity research.

TABLE I COMPARATIVE TABLE OF THE SIX REFERENCES

No.	First Author	Year	Research Topic	Domain	Main Contribution
1	S. Ovens	2022	Space complexity of consensus using swap objects	Distributed Computing	Establishes lower bounds on swap objects needed for randomized wait-free consensus.
2	M. Moshkov	2022	Time and space complexity of decision trees	Information Theory / Computational Complexity	Characterizes deterministic and nondeterministic decision-tree complexity and time–space tradeoffs.
3	X. Zhu	2025	Space complexity of Euclidean clustering	High-Dimensional Geometry / TCS	Provides tight space lower bounds for Euclidean clustering and limitations of dimensionality reduction.
4	J. M. Hernández-Morales	2024	Convolution in complexity space and its dual	Mathematics / Complexity Theory	Develops a generalized framework for complexity aggregation using convolution operations.
5	M. Skotniczny	2024	Complexity of solvers on space–time formulations	Numerical Analysis / Scientific Computing	Analyzes complexity of direct and iterative solvers on h-refined grids near singularities.
6	A. Rosen	2024	Tracking forest structural complexity over space and time	Ecology / Remote Sensing	Quantifies spatiotemporal changes in tropical forest structural complexity using ecological data.

V. CONCLUSION

This paper consolidates key theoretical developments in space complexity across distributed systems, high-dimensional clustering, decision structures, and functional models. By examining the intrinsic spatial limitations of each domain, the study identifies tight lower bounds and structural constraints that govern memory-efficient computation.

In distributed computing, the object-space requirements for consensus protocols are shown to be asymptotically tight through covering arguments, yielding precise criteria for designing minimal-memory coordination mechanisms. In high-dimensional clustering, ε-coreset constructions are demonstrated to be both practically effective and theoretically optimal under approximation guarantees, with rigorous evidence that dimensionality-reduction techniques cannot bypass fundamental space lower bounds.

In decision-tree analysis, the classification of infinite binary information systems into distinct complexity tiers provides a formal framework for assessing structural learning costs, supported by boundary-function characterizations that quantify optimal time–space tradeoffs. Functional convolution models extend these insights by generalizing complexity aggregation, offering new avenues for compositional and resource-aware optimization.

Collectively, these findings provide a unified perspective on space complexity, establishing a robust theoretical foundation for future work in approximation algorithms, resource-bounded computation, and the design of space-efficient computational systems across diverse paradigms.

ACKNOWLEDGMENT

I sincerely thank the course instructor and academic advisor for their insightful feedback and continuous guidance during the development of this article. The analysis of spatial complexity in distributed consensus, decision-making systems and high-dimensional clustering provides key perspectives and

inspirations. This work also benefited from discussions with peers, whose constructive suggestions helped enhance the clarity and depth of the final manuscript.

Acknowledgment to fundamental Research Grant Scheme (FRGS), Malaysia, FRGS/1/2024/ICT02/UTM/02/10. (R.J130000.7828.5F748).

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest regarding the publication of this paper.

REFERENCES

[1] Ovens, S. (2022). The space complexity of consensus from swap. *Proceedings of the ACM Symposium on Principles of Distributed Computing (PODC)*, 1–11.

[2] Moshkov, M. (2022). Time and space complexity of deterministic and nondeterministic decision trees. *Annals of Mathematics and Artificial Intelligence, 91*, 45–74.

[3] Zhu, X., Tian, Y., Huang, L., & Huang, Z. (2025). Space complexity of Euclidean clustering. *IEEE Transactions on Information Theory, 71*(6), 4515–4530.

[4] Hernández-Morales, J. M., Castañeda-Roldán, N. C., & Álvarez-Marín, L. C. (2024). Properties of the convolution operation in the complexity space and its dual. *Revista de la Unión Matemática Argentina, 67*(1), 281–299.

[5] Skotniczny, M., Paszyńska, A., Rojas, S., & Paszyński, M. (2024). Complexity of direct and iterative solvers on space–time formulations and time-marching schemes for h-refined grids towards singularities. *Journal of Computational Science, 76*, 102216.

[6] Rosen, A., Fischer, F. J., Coomes, D. A., Jackson, T. D., Asner, G. P., & Jucker, T. (2024). Tracking shifts in forest structural complexity through space and time in human-modified tropical landscapes. *Ecography, 2024*, e07377.