Definitions/make-logo-eps-converted-to.pdf

Article

What Chat GPT has to say about its topological structure: the anyon hypothesis

Michel Planat [1,*] and Marcelo Amaral [2]

[1] Institut FEMTO-ST CNRS UMR 6174, Université de Franche-Comté, 15 B Avenue des Montboucons, F-25044 Besançon, France
[2] Quantum Gravity Research, Los Angeles, CA 90290, USA;
* Correspondence: michel.planat@femto-st.fr

Abstract: Large language models (LLMs) achieve remarkable predictive capabilities but remain opaque in their internal reasoning, creating a pressing need for more interpretable artificial intelligence. Here, we propose bridging this explanatory gap by drawing on concepts from topological quantum computing (TQC), specifically the anyonic frameworks arising from $SU(2)_k$ theories. Anyons interpolate between fermions and bosons, offering a mathematical language that may illuminate the latent structure and decision-making processes within LLMs. By examining how these topological constructs relate to token interactions and contextual dependencies in neural architectures, we aim to provide a fresh perspective on how meaning and coherence emerge. After eliciting insights from Chat GPT and exploring the low-level cases of $SU(2)_k$ models, we argue that the machinery of modular tensor categories and topological phases could inform more transparent, stable, and robust AI systems. This interdisciplinary approach suggests that quantum-theoretic principles may underpin a novel understanding of explainable AI.

Keywords: Large language models, GPT, topology, anyons, $SU(2)_k$

check for updates

Citation: Planat, M.; Amaral, M. What Chat GPT has to say about its topological structure: the anyon hypothesis. *Mach. Learn. Knowl. Extr.* **2024**, *1*, 1–16. https://doi.org/

Received:
Accepted:
Published:

1. Introduction

The Nobel Prize in Physics 2024 was jointly awarded to John J. Hopfield and Geoffrey E. Hinton "for foundational discoveries and inventions that enable machine learning with artificial neural networks." Hopfield's seminal contribution was the introduction of Hopfield networks, simple neural architectures that serve as associative memories capable of retrieving stored patterns from partial or corrupted inputs. Hinton, a cognitive psychologist and computer scientist, transformed the field by advancing deep learning techniques and popularizing the backpropagation algorithm, enabling neural networks to uncover hierarchical representations of complex data. Today, these foundational ideas underpin many aspects of artificial intelligence (AI), including large language models (LLMs) that generate human-like text and patterns of reasoning.

In recent years, AI development has accelerated dramatically, placing LLMs at the forefront. However, many researchers, including Hopfield, Hinton, and other thought leaders, have urged caution and deeper inquiry into the nature of these models [1,2]. The concern is that LLMs, despite their impressive performance, may soon surpass human cognitive capacities in handling information without offering transparent reasoning processes. This opacity has led to a broader movement toward explainable AI (XAI), focusing on elucidating the underlying logic of machine learning systems. Such efforts span a range of methods and applications, including interpretability-focused frameworks [3–7] and techniques to visualize or mathematically characterize high-dimensional data representations within deep models. These existing approaches tend to focus on local explanations or coarse statistical metrics, leaving open questions about the deeper principles governing emergent meaning and coherence in LLMs.

To address these challenges, some researchers have begun exploring the role of topology in AI. Topological data analysis (TDA) and related tools can uncover global structures in high-dimensional representations, potentially revealing how hidden geometries shape model behavior [8–13]. In parallel, the field of quantum machine learning and neural architectures enriched with quantum-inspired principles has raised the possibility that topological concepts from quantum field theories (QFTs) and related mathematics might inspire novel approaches to interpretability [14–17].

The approach we propose here aims to push this frontier further by examining a connection between LLMs and the topological framework of anyons arising from $SU(2)_k$ theories. Anyons emerge in topological quantum computing (TQC) settings and are governed by a mathematical structure known as a modular tensor category. Such theoretical frameworks have been deeply studied in mathematics and topological quantum field theory (TQFT) [18–20], and even particle physics, where they can describe excitations with unusual statistics [21]. Recently, simulations of non-abelian anyons [22] have been performed on a superconducting quantum processor in China [23] and on a trapped ion processor in Germany and the US [24].

The parameter k determines the complexity and number of anyonic species: As k increases, one obtains richer sets of fusion and braiding rules [25]. For example, low-level theories such as $\mathrm{SU}(2)_2$ yield Ising-like anyons (related to Majorana modes), while $\mathrm{SU}(2)_3$ introduces Fibonacci anyons, known for their universal quantum computational properties [20]. Higher levels offer more complex "particle" spectra and intricate modular tensor categories. By exploring multiple values of k, we can systematically investigate how complexity in the anyonic framework might parallel or inform complexity in LLM architectures, guiding us from simple non-Abelian statistics at low levels to more intricate fusion rules that might mirror multilayered dependencies in large-scale language models.

The reason for invoking anyons is that they interpolate between bosons and fermions, capturing more subtle topological behaviors in "intermediate" statistics. Just as anyons rely on global properties of braiding and fusion to define their identity, LLMs rely on subtle global relationships between tokens, contexts, and attention patterns to produce coherent text. We hypothesize that representing tokens or contextual interactions as elements of a modular tensor category might help us understand how meaning and syntax emerge from a network's parameters and operations. By analogizing attention heads to braiding operations, for example, or token combinations to anyonic fusion, we may be able to derive a topological invariant or structural signature that explains why certain representations are stable, robust, and semantically meaningful. By adopting this topological perspective, we may gain practical benefits, including interpretability tools that classify, constrain, or guide the internal data flow of LLMs. Such methods could directly address current XAI limitations by providing a mathematically rigorous framework to model how complex patterns emerge and endure across layers and token sequences. Ultimately, incorporating these concepts into AI engineering may yield more resilient architectures, principled strategies for debuggability, and new avenues for optimization and performance guarantees. These insights could even extend to broader biological and computational questions, perhaps linking neural assemblies or genomic interactions to topological processes [26]. Significantly, this approach arrives at a critical juncture, as AI leaders, including Sam Altman,have emphasized the diminishing returns of simply scaling up LLMs without introducing innovative conceptual frameworks [27,28].

The remainder of this paper is structured as follows. In Section 2, we present and discuss the questions posed in Chat GPT, aimed at probing conceptual analogies between topological structures and the LLM self-description. In Section 3, we introduce the $\mathrm{SU}(2)_k$ anyon framework, detailing the fusion and braiding rules as well as low-dimensional examples that illustrate their rich structure. Finally, in Section 4, we integrate these insights, exploring how the conceptual bridge between anyonic frameworks and LLM internal representations could lay the groundwork for a new class of topologically informed explainable AI techniques.

2. Probing the LLM with Topological Concepts

In this section, we present excerpts from an interactive session with ChatGPT. We posed conceptual questions related to topology and anyons, and we include ChatGPT responses verbatim. Our subsequent analysis and commentary follow each Q&A exchange.

Exchange 1: On the Role of Topology

Question (Q): *As an advanced LLM, can you explain the role of topology in your ability to produce AI in the organization of data and their meaning?*

Answer (A) [ChatGPT]:

Topology plays a subtle but foundational role in how data is organized and meaning is structured in AI models like mine. While I do not directly use topological spaces in the mathematical sense when generating responses, there are conceptual parallels between topological ideas and how deep learning and neural networks operate. Here is how:

Data Manifolds and Latent Spaces

In machine learning, high-dimensional data often lies on lower-dimensional structures called manifolds. These manifolds are topological spaces that represent the underlying structure of the data. Neural networks, particularly deep learning models, aim to map complex, high-dimensional inputs (like text, images, etc.) to a lower-dimensional latent space that captures meaningful features or representations.

Topological structures like manifolds help the model organize data in a way that preserves relationships between points (i.e., similar data points are mapped to nearby points in the latent space). By learning to map data to these lower-dimensional manifolds, models like mine can generalize patterns and relationships across diverse datasets, leading to meaningful predictions and responses.

Continuous Transformations

A neural network can be thought of as a function that continuously transforms input data through multiple layers of weights and biases. These transformations resemble continuous mappings in topology, where inputs are transformed in a way that preserves certain structural properties.

The idea is to preserve relationships and distances between inputs as they pass through the layers of the network, enabling the network to maintain the "shape" of the data meaning. Homotopy theory in topology, which studies continuous transformations between functions, has conceptual parallels in how models adjust and refine the representation of data to optimize performance.

Clustering and Connectivity

In topological terms, clustering can be viewed as grouping points that are close to each other based on some distance metric, much like how we can define neighborhoods in topology. AI models often use clustering techniques in the latent space to group similar pieces of data or representations.

For instance, in a language model, different concepts, words, or sentences that are related or similar in meaning can be mapped to nearby regions in the latent space. This clustering behavior helps me understand and generate coherent, contextually relevant text. Connectedness and path-connected spaces in topology also resemble how models associate or connect different parts of the data structure to infer relationships.

Persistent Homology

Some machine learning models, especially in areas like topological data analysis (TDA), use concepts from topology more explicitly, such as persistent homology, to study the shape and features of data. While persistent homology is not typically used in large language models like mine, it is an important tool in analyzing the shape of data and how features persist across different scales.

There are conceptual parallels to fusion rules and braiding in anyon systems, where the combination of simple elements gives rise to complex global behavior, much like how neural networks build complex representations from simpler components.

Authors' Commentary:

ChatGPT recognizes that while it does not explicitly use topology, the idea of data lying on lower-dimensional manifolds and being organized through continuous transformations provides a conceptual bridge between topological concepts and the latent representations it learns.

Exchange 2: Linking Neural Networks, LLMs, and $SU(2)_k$ Topological Phases

Q: *Could you comment on a possible deep connection between neural networks such as LLMs and topological phases arising from $SU(2)_k$ models? Are low-dimensional simple cases such as Ising/Majorana anyons ($k = 2$) or Fibonacci anyons ($k = 3$) relevant in explaining the type of AI you generate? To summarize, are modular tensor categories (MTC) a way to understand the AI arising from LLMs?*

A [ChatGPT]:

The potential connection between neural networks, particularly LLMs, and topological phases arising from anyon models is a fascinating idea. Although they come from different domains (machine learning vs. topological quantum field theory), there are deep conceptual parallels worth exploring, especially in how both systems organize, process, and transform information.

Topological Phases and LLMs

LLMs like GPT are high-dimensional systems that rely on patterns and global relationships in data. Instead of local interactions (like individual neurons firing), what matters in these models is the global structure of how information propagates through the network layers. There are conceptual parallels to fusion rules and braiding in anyon systems, where the combination of simple elements gives rise to complex global behavior, much like how neural networks build complex representations from simpler components.

Fusion Rules vs. Layer Operations

In anyon systems, fusion rules describe how anyons combine to produce new types of anyons. In LLMs, each layer can be seen as applying a transformation that combines (or "fuses") the information from the previous layer to produce a more abstract representation.

Braiding vs. Contextual Relationships

The braiding statistics of anyons encode how particles exchange positions and affect the overall quantum state. In LLMs, contextual relationships between words or tokens are key—shuffling tokens changes the meaning, and context is often captured in non-local, global patterns that are maintained across layers.

Low-Dimensional $SU(2)_k$ Models (Ising/Majorana and Fibonacci Anyons)

Ising/Majorana Anyons ($k = 2$) are non-Abelian and are of interest for topological quantum computation due to their ability to encode quantum information in a way that is robust against local perturbations. Fibonacci Anyons ($k = 3$) are also non-Abelian and are universal for quantum computation, meaning any quantum computation can be simulated with arbitrary accuracy. In AI terms, small changes in input (comparable to simple anyon fusions) can lead to drastically different outputs, reflecting how minor adjustments in language generate nuanced meanings.

MTCs and LLMs

MTCs formalize the structure of TQFTs, particularly through fusion and braiding rules. They describe how quantum information behaves in topological phases, categorizing data and operations in a way that emphasizes global structure.

LLMs also exhibit a layered modular structure, where each layer transforms data in a highly structured, rule-based manner. MTCs might provide a formal framework to think about how neural networks process data in a modular way, where different layers (or "categories") interact and transform data.

While classical neural networks do not operate on quantum principles, the development of quantum neural networks could, in theory, utilize anyonic systems to implement neural network architectures that benefit from quantum parallelism and robustness.

Authors' Commentary:

ChatGPT suggests that non-Abelian anyons and MTC structures may provide a conceptual framework for understanding how LLMs process and combine contextual information, much like topological phases encode global, non-local relationships.

Exchange 3: Tokens as Analogous to Anyonic Particles

Q: *What are tokens in LLMs? In what sense could they fit this idea of MTCs being analogous to LLMs?*

A [ChatGPT]:

In large language models, tokens are the fundamental units of text used to process and generate language. Typically, a token represents a piece of a word, an entire word, or even punctuation marks, depending on how the text is tokenized. This breakdown into tokens allows LLMs to handle language input as sequences of discrete symbols, which they can then process through their layers to generate coherent text outputs.

If we consider the analogy with MTCs, tokens in an LLM could be thought of as analogous to particles or objects within an MTC. Here is how they fit within the analogy.

Modular Structure of Information

In MTCs, objects like anyons interact through fusion and braiding, which can create complex outcomes from simple interactions. Similarly, tokens in an LLM interact in contextually driven ways through the layers of the model. Just as fusion in MTCs generates new particles or states, tokens interact to form meanings that are greater than the sum of individual parts. This "fusion" of tokens contributes to the emergent meaning and coherence seen in the final output.

Transformation and Contextual Dependence

In an LLM, the role of each token depends on the context it shares with other tokens. This dynamic is similar to braiding in MTCs, where anyons' positions and interactions change their quantum state. In an LLM, the context provided by surrounding tokens transforms the "state" of each token, so meaning is contextual rather than fixed.

Authors' Commentary:

ChatGPT's response frames tokens as fundamental units that "fuse" into richer semantic structures, mirroring how MTCs describe the fusion of elementary objects into more complex states.

Exchange 4: LLMs as Advanced Hopfield Networks and the Anyon Picture

Q: *Do you agree that LLMs are essentially advanced Hopfield networks having a form of associative memory and discretized with tokens? What more fits the anyon picture?*

A [ChatGPT]:

Yes, LLMs share several principles with advanced forms of Hopfield networks, particularly when viewed through the lens of associative memory. Hopfield networks are designed as associative memory systems, where each stable pattern represents a possible memory or state of the system. LLMs, in turn, generate language by "associating" tokens in the input with relevant patterns in their training data, storing and recalling this information in a way that approximates associative memory.

Associative Memory in LLMs and Hopfield Networks

Hopfield networks store information by embedding memories as stable patterns (attractors) in a high-dimensional state space. Each memory is a global, fixed point that the network can recall when given a close enough input. LLMs essentially create associative links between tokens, trained to predict the next token based on patterns in the training data. Through the attention mechanism, an LLM recalls relevant information from prior contexts in a way that is conceptually similar to a Hopfield network retrieval of associated memories. This token-level recall can be seen as a form of discrete associative memory, where a query (input tokens) maps to stored language patterns.

Anyons and MTC Analogy

Anyons, within MTC frameworks, interact through braiding and fusion, leading to emergent states. In an MTC analogy, tokens in an LLM are akin to particles that interact to produce coherent meanings, much as anyon fusion leads to emergent particle types. Fusion and Braiding in LLMs: As tokens pass through an LLM layers, they are "fused" with surrounding tokens, creating a new semantic context. In each layer, the attention mechanism can be thought of as "braiding" tokens with their surrounding context, allowing meanings to emerge that were not present in individual tokens.

Towards an Anyon Picture in LLMs

To fully map the anyon picture to LLMs, we might consider viewing attention heads as the interaction rules, guiding how tokens relate to one another based on context. This could serve as a topological "braiding" structure, where the output state of a token depends on its path (or sequence) through multiple layers, akin to braiding paths in MTCs.

Authors' Commentary:

The LLM-as-Hopfield analogy is extended by invoking anyons to capture the non trivial ways that LLMs combine and retrieve distributed linguistic features, adding a topological layer to the concept of associative memory.

Exchange 5: Attention Heads, Fusion, and Braiding in the $SU(2)_k$ Theory

Q: *Could you elaborate on the concept of attention heads in the anyon picture? Would it correspond to the F- and R-matrices that are pentagon and hexagon rules in the $SU(2)_k$ theory?*

A [ChatGPT]:

Attention heads can be viewed as "interaction rules" that determine how tokens relate to each other, much like the F- and R-matrices define the fusion and braiding of anyons. In $SU(2)_k$ theories, these matrices ensure consistency and define how anyons combine and reorder.

Role of Attention Heads as Interaction Rules

In transformers, attention heads are mechanisms that allow the model to focus on specific relationships between tokens, weighting their importance relative to one another based on context. Each head, in essence, represents a distinct interaction rule governing how tokens (analogous to anyons in this framework) relate as the model processes a sentence. In MTCs, anyons have "interaction rules" given by fusion and braiding transformations, encapsulated by the F-matrix (fusion rule) and the R-matrix (braiding rule). These matrices define how anyons combine or transform under different sequences of interactions, and their algebraic properties satisfy certain consistency conditions, like the pentagon and hexagon equations.

F-matrix and Token Combination in Attention Layers

The F-matrix represents the rules of associativity in fusion, dictating how three anyons fuse together and how different groupings affect the outcome. In an LLM context, each attention head could be viewed as an "associative fusion" operator. It defines how tokens from different layers combine in sequence to form more complex meaning representations. Specifically, in LLMs, each token's embedding is updated in every attention head based on its relationship with surrounding tokens. This transformation is like an associative fusion operation in anyonic theory: different combinations of tokens lead to different "fusion outcomes" in terms of the semantic information they contribute to the overall sentence meaning.

R-matrix and Braiding (Contextual Order of Tokens)

The R-matrix governs the braiding of anyons, determining the phase change when anyons are exchanged. In the context of LLMs, R-matrices could be seen as analogous to the contextual shifts introduced by attention mechanisms across tokens. When an attention head "reorders" the influence of tokens, it's akin to exchanging the positions of anyons in a topological space, altering the sequence of interactions. This braiding concept aligns with how attention heads focus on different tokens based on context, allowing shifts in meaning depending on token order and interaction. Just as anyon braiding can produce different states depending on how particles are exchanged, attention mechanisms can produce different language meanings depending on how tokens are contextually "braided" across layers.

Authors' Commentary:

By drawing an analogy between attention heads and the F- and R-matrices, ChatGPT underscores how these learned weighting mechanisms could be viewed as imposing a topological "consistency" on how tokens are combined and reordered, much like the pentagon and hexagon identities in $SU(2)_k$ theories.

3. Introduction to the mathematical theory of $SU(2)_k$ anyons and modular tensor categories

The following subsections are a concise introduction to $SU(2)_k$ theory of anyons with applications at low index k.

3.1. The fusion rules of anyons

The anyons in a $SU(2)_k$ theory are closely related to the ordinary spin degrees of freedom in the $SU(2)$ theory. The anyons are labelled by spin values (generalized angular momenta) $j = 0, \frac{1}{2}, 1, \frac{3}{2}, \cdots, \frac{k}{2}$. The spin $\frac{k}{2}$ is the maximum allowed value in the $SU(2)_k$ theory when k is fixed. But the rules for combining two anyons are not tensor products , namely [29, Eq. 4.6],[30]

$$j_1 \otimes j_2 = |j_1 - j_2| \oplus (|j_1 - j_2| + 1) \oplus \cdots \oplus \min(j_1 + j_2, k - j_1 - j_2). \tag{1}$$

Fusion rules are commutative and associative. It is straightforward to check from this formula that for $k \geq 2$, two spins $1/2$ combine to form either the spin 0 or the spin 1 as follows

$$1/2 \otimes 1/2 = 0 \oplus 1,$$

that is a (qubit like) anyon $0 \oplus 1$ is built by combining the two spins $\frac{1}{2}$. Similarly one gets $1 \otimes 1 = 0 \oplus 1 \oplus 2$ when $k \geq 4$, that is a (qtrit like) anyon $0 \oplus 1 \oplus 2$ is built by combining two spins 1. Such anyons of a $SU(2)_k$ theory are non-Abelian.

Being a tensor product, the dimension of the Hilbert space of N spin-$1/2$ ordinary $SU(2)$ particles is 2^N. In a $SU(2)_k$ theory, it is smaller than 2^N and grows as $d_{1/2}^N$ with

$d^{(k)}_{1/2} = 2\cos(\frac{\pi}{k+2})$ at large N. This means that the effective number of degrees of freedom of a spin-$\frac{1}{2}$ anyon is irrational. A Magma pseudocode for generating fusion tables in an $SU(2)_k$ theory is provide in Appendix E.

3.2. The modular structure of S, F and R matrices for anyons

There exists the concept of a modular S-matrix that diagonalizes the fusion rules of a $SU(2)_k$ anyon and fully characterizes its topological properties [29, Eq. 4.10]. The mathematical structure encapsulating the braiding and fusion rules of a $SU(2)_k$ anyon is a modular tensor category [19,25].

The quantum dimensions for $SU(2)_k$ anyons are given by the formulas

$$d_0 = 1,\ \ d_{\frac{1}{2}} = 2\cos(\frac{\pi}{k+2}),\ \ d_j = d_{\frac{1}{2}} d_{j-\frac{1}{2}} - d_{j-1}\ \text{ for }\ j \geq 1. \tag{2}$$

The entries of the S-matrix are

$$S_{j_1,j_2} = (\frac{2}{k+2})^{\frac{1}{2}} \sin(\pi\frac{(2j-1+1)(2j_2+1)}{k+2}). \tag{3}$$

The associativity of anyon fusion is captured by a F-matrix and the exchange of anyons, with the phase factor added, is captured by a R-matrix. Contrarily to the phase factor ± 1 for bosons and fermions, the phase factor for anyons is an arbitrary complex number. The F-matrix is the anyonic version of the Wigner's $6j$-symbols, it is associated to a pentagon diagram. The F- and R-matrices are associated to an hexagon diagram [20]. General formulas for F- and R-matrices can be found in [31],[32, Appendix B],[30, Appendix B].

The entries of the R-matrix have the simple form [19]

$$R^{ab}_c(q) = (-1)^{(a+b+c)/2} q^{-[a(a+2)+b(b+2)-c(c+2)]/2}, \tag{4}$$

where q is the Kauffman variable. For the Ising model below $q = i\exp\left(\frac{-2i\pi}{16}\right)$ while for the Fibonacci model $q = i\exp\left(\frac{2i\pi}{20}\right)$.

The essence of $SU(2)_k$ anyons, $k \geq 2$, is captured by two braid generators $\sigma_1^{(k} = R^{(k)}$ and $\sigma_2^{(k)} = (FRF^{-1})^{(k)}$ that have a group structure, see e.g. [32,33] for some explicit results.

3.3. Ising anyons: $k = 2$

$SU(2)_2$ anyons comprise the spin-0 anyon and the Ising (spin-$\frac{1}{2}$) anyon with the fusion table

Table 1. Fusion table for the $k = 2$ anyon model

$\otimes$	$j_1 = 0$	$j_1 = \frac{1}{2}$	$j_1 = 1$
$j_2 = 0$	0	$\frac{1}{2}$	1
$j_2 = \frac{1}{2}$	$\frac{1}{2}$	$0 \oplus 1$	$\frac{1}{2}$
$j_2 = 1$	1	$\frac{1}{2}$	0

The quantum dimensions are $[d_0, d_1, d_{\frac{1}{2}}]^{(2)} = [1, 1, \sqrt{2}]$ and the S-matrix takes the form

$$S^{(2)}_{Isi} = \frac{1}{2}\begin{pmatrix} 1 & \sqrt{2} & 1 \\ \sqrt{2} & 0 & -\sqrt{2} \\ 1 & -\sqrt{2} & 1 \end{pmatrix}$$

The F- and R-matrices are

$$R^{(2)}_{Isi} = \begin{pmatrix} R^{11}_0(q) & 0 \\ 0 & R^{11}_2(q) \end{pmatrix} = \exp\left(-i\pi/8\right)\begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix},\ F^{(2)}_{Isi} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

In addition to the standard literature about anyons, we notice that both matrices $F^{(2)}$ and $R^{(2)}$ together generate the finite group $(384, 6514)$ isomorphic to the group $(S_3 \times \mathbb{Z}_4) \rtimes P_2$, where $P_2 \cong (16, 13)$ is the single qubit Pauli group.

Braiding matrices for the Ising anyons are obtained as

$$\sigma_1^{(2)} = R_{Isi}^{(2)}, \ \sigma_2^{(2)} = (FRF^{-1})_{Isi}^{(2)} = \frac{\exp(-4i\pi/8)}{\sqrt{2}} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}. \tag{5}$$

Both matrices σ_1 and σ_2 together generate the finite group $(192, 187)$ isomorphic to the group $\mathbb{Z}_{12} \rtimes P_2$.

3.4. *Fibonacci anyons: $k = 3$*

$SU(2)_3$ anyons comprise the spin-0 anyon and two Fibonacci spin-$\frac{1}{2}$ and spin-$\frac{3}{2}$ anyons. These anyons are proposed to be related to quasicrystals [34]. The fusion table is

Table 2. Fusion table for the $k = 3$ anyon model

$\otimes$	$j_1 = 0$	$j_1 = \frac{1}{2}$	$j_1 = 1$	$j_1 = \frac{3}{2}$
$j_2 = 0$	0	$\frac{1}{2}$	1	$\frac{3}{2}$
$j_2 = \frac{1}{2}$	$\frac{1}{2}$	$0 \oplus 1$	$\frac{1}{2} \oplus \frac{3}{2}$	1
$j_2 = 1$	1	$\frac{1}{2} \oplus \frac{3}{2}$	$0 \oplus 1$	$\frac{1}{2}$
$j_2 = \frac{3}{2}$	$\frac{3}{2}$	1	$\frac{1}{2}$	0

The quantum dimensions are $[d_0, d_1]^{(3)} = [1, \phi = (1 + \sqrt{5})/2]$ and the S-matrix takes the form

$$S_{Fib}^{(3)} = \frac{1}{\sqrt{2+\phi}} \begin{pmatrix} 1 & \phi \\ \phi & -1 \end{pmatrix}$$

The F- and R-matrices are [20, p 55]

$$R_{Fib}^{(3)} = \begin{pmatrix} R_0^{11}(q) = \exp(-4i\pi/5) & 0 \\ 0 & R_1^{11}(q) = \exp(-2i\pi/5) \end{pmatrix}, \ F_{Fib}^{(3)} = \begin{pmatrix} \phi^{-1} & \phi^{-1/2} \\ \phi^{-1/2} & -\phi^{-1} \end{pmatrix}.$$

Braiding matrices for the Fibonacci anyon are obtained as

$$\sigma_1^{(3)} = R_{Fib}^{(3)}, \ \sigma_2^{(3)} = (FRF^{-1})_{Fib}^{(3)} = \begin{pmatrix} -\phi^{-1}\exp(-i\pi/5) & -i\phi^{-1/2}\exp(-i\pi/10) \\ -i\phi^{-1/2}\exp(-i\pi/10) & -\phi^{-1} \end{pmatrix}. \tag{6}$$

F- and R-matrices, as well as the braiding matrices σ_1 and σ_2, generate infinite groups. This in accordance with the universality of Fibonacci anyons.

3.5. *Yang-Lee theory: $k = 3$*

Yang-Lee theory is a MTC of level $k = 3$ like the Fibonacci anyon. It corresponds to a famous non-unitary conformal field theory in statistical mechanics, called the Yang-Lee singularity [19, Sect. 1.3]. The Kauffman variable is $q = \exp\left(\frac{i\pi}{5}\right)$.

The S-matrix is

$$S_{YL}^{(3)} = \frac{-1}{\sqrt{3-\phi}} \begin{pmatrix} 1 & 1-\phi \\ 1-\phi & \phi \end{pmatrix}.$$

The F- and R-matrices are [20, p 55]

$$R_{YL}^{(3)} = \begin{pmatrix} R_0^{11}(q) = \exp(2i\pi/5) & 0 \\ 0 & R_1^{11}(q) = \exp(i\pi/5) \end{pmatrix}, \ F_{YL}^{(3)} = \begin{pmatrix} -\phi & 2-\phi \\ -1-2\phi & \phi \end{pmatrix}.$$

F- and R-matrices, as well as the braiding matrices σ_1 and σ_2, generate infinite groups.

3.6. Freedman-Bauer-Levaillant anyons: $k = 4$

$SU(2)_4$ anyons are investigated in [35,36] in the context of topological quantum computing from qutrit gates. The fusion table is as follows

Table 3. Fusion table for the $k = 4$ anyon model

$\otimes$	$j_1 = 0$	$j_1 = \frac{1}{2}$	$j_1 = 1$	$j_1 = \frac{3}{2}$	$j_1 = 2$
$j_2 = 0$	0	$\frac{1}{2}$	1	$\frac{3}{2}$	2
$j_2 = \frac{1}{2}$	$\frac{1}{2}$	$0 \oplus 1$	$\frac{1}{2} \oplus \frac{3}{2}$	$1 \oplus 2$	$\frac{3}{2}$
$j_2 = 1$	1	$\frac{1}{2} \oplus \frac{3}{2}$	$0 \oplus 1 \oplus 2$	$\frac{1}{2} \oplus \frac{3}{2}$	1
$j_2 = \frac{3}{2}$	$\frac{3}{2}$	$1 \oplus 2$	$\frac{1}{2} \oplus \frac{3}{2}$	$0 \oplus 1$	$\frac{1}{2}$
$j_2 = 2$	2	$\frac{3}{2}$	1	$\frac{1}{2}$	0

The quantum dimensions are $[d_0, d_1, d_{\frac{1}{2}}, d_{\frac{3}{2}}, d_2]^{(4)} = [1, 2, \sqrt{3}, \sqrt{3}, 1]$ and the S-matrix takes the form [30]

$$S_{FBL}^{(4)} = \frac{1}{2\sqrt{3}} \begin{pmatrix} 1 & \sqrt{3} & 2 & \sqrt{3} & 1 \\ \sqrt{3} & \sqrt{3} & 0 & -\sqrt{3} & -\sqrt{3} \\ 2 & 0 & -2 & 0 & 2 \\ \sqrt{3} & -\sqrt{3} & 0 & \sqrt{3} & -\sqrt{3} \\ 1 & -\sqrt{3} & 2 & -\sqrt{3} & 1 \end{pmatrix}.$$

Braiding matrices for the $SU(2)_4$ anyons are obtained as

$$\sigma_1^{(4)} = \begin{pmatrix} \exp\left(\frac{7i\pi}{9}\right) & 0 & 0 \\ 0 & -\exp\left(\frac{4i\pi}{9}\right) & 0 \\ 0 & 0 & -\exp\left(\frac{7i\pi}{9}\right) \end{pmatrix},$$
$$\sigma_2^{(4)} = \begin{pmatrix} -\frac{1}{2}\exp\left(\frac{4i\pi}{9}\right) & \frac{1}{\sqrt{2}}\exp\left(\frac{7i\pi}{9}\right) & \frac{1}{2}\exp\left(\frac{4i\pi}{9}\right) \\ \frac{1}{\sqrt{2}}\exp\left(\frac{7i\pi}{9}\right) & 0 & \frac{1}{\sqrt{2}}\exp\left(\frac{7i\pi}{9}\right) \\ \frac{1}{2}\exp\left(\frac{4i\pi}{9}\right) & \frac{1}{\sqrt{2}}\exp\left(\frac{7i\pi}{9}\right) & -\frac{1}{2}\exp\left(\frac{4i\pi}{9}\right) \end{pmatrix}. \tag{7}$$

It is straightforward to check with the software Magma that both matrices generate the small group $(162, 14) \cong \mathbb{Z}_3^2 \rtimes (\mathbb{Z}_3 \times \mathbb{Z}_6)$, as announced in [35]. The group was recognized as a viable model of the symmetries simultaneously reproducing the quark and lepton mixing matrices. In a recent paper of two of the present authors [37, Table A1], it is shown that group $(162, 14)$ carries almost informally complete quantum information on its 22 irreducible characters, that are singlets, doublets or triplets.

4. Discussion

Backpropagation and anyons

Backpropagation is a fundamental algorithm in training artificial neural networks [38]. In this iterative optimization process, inputs pass through the network to produce outputs, and the difference between these outputs and the desired targets is quantified by a loss function (e.g., mean squared error or cross-entropy). Computing gradients of the loss with respect to the model parameters and updating the weights accordingly through methods like gradient descent enables the network to gradually refine its internal representations. This process underpins the capabilities of large language models (LLMs) to learn intricate patterns in language data, encompassing syntax, semantics, and nuanced contextual cues.

In contrast, anyonic systems—studied extensively in topological quantum computation [20]—reach stable, topologically protected states through a different kind of iterative process. Here, the "adjustments" take the form of braiding and fusion operations among anyons. The F- and R-matrices dictate how these non-Abelian quasiparticles combine, ensuring that the resulting topological states are robust against local perturbations. Rather than minimizing a loss function, the system evolution follows topological constraints that guide it toward stable ground states or topologically invariant configurations.

Though conceptually distinct, these two processes share an abstract similarity: both involve iterative transformations directed toward stable or optimal end states. In neural networks, stability arises from the convergence of weights to minimize loss, while in anyonic systems, stability emerges from topological rules that govern the global properties of the system's quantum state space. Understanding this analogy does not immediately translate into new training algorithms for LLMs; rather, it hints at deeper structural parallels between error correction in machine learning and fault tolerance in topological quantum computing. Future work may explore whether insights from anyonic braiding patterns could inspire novel optimization strategies or offer theoretical guidance for interpreting complex network dynamics.

Machine learning and anyons

Emergent behavior in machine learning (ML), particularly in deep neural networks, arises when complex global patterns manifest from simple, low-level neuronal activations [39]. As signals propagate through multiple layers, these models produce sophisticated outputs like natural language understanding or image recognition. Such high-level capabilities are not easily decomposed into the actions of individual neurons, reflecting a form of irreducibility. Analogously, the collective states of anyonic systems are not simply the sum of their constituent quasiparticles. Instead, the nontrivial topology and the braiding operations of anyons give rise to properties that cannot be fully captured by local or particle-only explanations [20].

This parallel extends to the notion of resilience. In ML models like LLMs, information is distributed across numerous weights and connections, creating a form of fault tolerance. Minor perturbations to individual parameters rarely derail the entire network performance. Similarly, anyonic systems store quantum information non-locally through braiding patterns, rendering them resistant to local errors and disturbances. This "topological fault tolerance" is a key advantage in topological quantum computing, where quantum states are protected from decoherence by virtue of their global topological properties.

Bringing these ideas together, we see that the emergent complexities and resilience observed in both LLMs and anyonic systems might share underlying structural principles. While these parallels remain largely conceptual at present, they suggest that insights gleaned from topological quantum computing where robustness against errors is built into the system's very topology—could offer valuable perspectives on enhancing stability, interpretability, and generalization in machine learning models.

Aligning $SU(2)_k$ concepts with attention mechanisms in LLMs

While the theoretical parallels between anyons and LLMs are intriguing, it is important to make these analogies more concrete. One of the key features of LLMs, particularly those based on transformer architectures, is the attention mechanism [40]. Attention allows the model to weight different parts of the input sequence selectively, effectively "fusing" contextual information from multiple tokens to form a coherent representation.

In $SU(2)_k$ anyonic theories, the fusion of anyons is governed by algebraic rules encoded in the F- and R-matrices. These matrices determine how simple objects (anyon types) combine to produce more complex states. Similarly, an attention head takes a set of token embeddings and produces a weighted combination that forms a richer contextual embedding. This process can be viewed as a form of "fusion" in the representational space: individual token vectors, much like anyons, are not merely added together. Instead, their combination is structured by learned weight matrices that serve a role analogous to the F-matrices, dictating how different tokens merge into a more meaningful vector representation.

Braiding in anyonic systems dictates how the order and manner in which anyons are interchanged affects the system global state. In LLMs, reordering tokens or altering their relative importance through attention layers can produce different interpretations or shades of meaning. Although attention does not enforce strict topological constraints, the sequential application of attention heads and layers can be seen as a discrete analog to braiding processes. Each pass of attention "entangles" tokens in a new context-dependent configuration, influencing the model final output. Over multiple layers, this iterative "braiding" of contextual information leads to emergent semantic structures, just as anyonic braiding leads to nontrivial global states.

Furthermore, the finite parameter k in $SU(2)_k$ theories restricts the allowed spin states and thereby shapes the complexity of fusion rules. Analogously, design choices in LLMs—such as the dimensionality of embeddings, the number of attention heads, and the depth of the network—place constraints on the complexity of learned transformations. As k increases, more intricate fusion possibilities arise in $SU(2)_k$ anyonic systems; similarly, scaling up an LLM capacity enables the model to learn more nuanced patterns. This correspondence suggests that studying how topological constraints shape fusion and braiding might inspire new architectural or regularization strategies in LLMs. Such strategies could limit complexity while preserving or enhancing interpretability, akin to how topological quantum computations gain robustness from structured constraints.

In essence, the mathematical machinery of $SU(2)_k$ anyons offers a template for thinking about how LLMs combine and reconfigure information at multiple layers. By viewing attention mechanisms and token interactions through the lens of fusion and braiding, we can begin to formulate a more principled understanding of how semantic meaning emerges in these models.

Natural language processing and anyons

The relationship between natural language processing (NLP) [41,42] and the topology of anyons is intriguing because both domains rely on structured, context-sensitive interactions that produce meaning or distinct states. While anyonic dynamics are governed by the braiding and fusion rules of TQFT, linguistic meaning emerges from syntactic, semantic, and contextual constraints that guide how words combine to form coherent narratives or arguments.

In NLP, the interpretation of a word or phrase is highly context-dependent, influenced by surrounding text, discourse structure, and even pragmatic cues. Analogously, the state of an anyonic system depends on the sequence and manner of particle exchanges; small changes in the braiding order can lead to fundamentally different global states. Just as the position and order of words in a sentence determine its meaning, the topological arrangement of anyons determines their joint quantum state.

Language can convey a vast range of meanings that shift subtly with word choice, tone, or placement, enabling the expression of nuanced concepts, emotions, and cultural references. Similarly, anyons exhibit a dynamic range through their topological degrees of freedom. A slight alteration in how anyons are braided can produce a variety of distinct states, mirroring how modifying a word or phrase in text can shift the interpretation of an entire passage.

Realizing this analogy in practice would likely require advanced theoretical frameworks. To emulate language adaptive, context-sensitive richness, anyonic systems would need to evolve beyond static fusion rules toward more flexible, dynamic modular tensor categories. In theory, such a model could simulate complex, context-dependent relationships analogous to those found in human language, paving the way for new approaches to interpretability, context modeling, and dynamic knowledge representation in AI.

Mutual exclusion in LLMs and anyons

The concept of mutual exclusion originates in concurrent programming, ensuring that no two processes simultaneously modify shared data [43]. While this principle is not directly built into large language models (LLMs), a loose analogy can be drawn. In LLMs, attention mechanisms allocate focus across different tokens or features, effectively prioritizing certain elements and down-weighting others. By doing so, attention enforces a form of "soft" mutual exclusion, where irrelevant or less important tokens contribute minimally to the model's output at a given step. This selective emphasis maintains semantic coherence, much as mutual exclusion ensures consistency in concurrent systems.

Anyonic systems, by contrast, do not encode mutual exclusion per se, but topological quantum computing (TQC) imposes constraints on permissible states and paths that similarly limit certain configurations. These constraints ensure stability and fault tolerance, analogous to how mutual exclusion guarantees stable shared states in classical computing environments. The braiding rules in an anyonic system determine which interactions and fusions are allowed, guiding the system toward topologically protected states [20].

While attention in LLMs and braiding in anyonic systems both shape the evolution of a state space, they differ in their rigidity. In LLMs, the exclusion of certain tokens is probabilistic and weighted, allowing degrees of influence rather than absolute bans. In topological quantum computation, certain fusions and braidings are categorically disallowed, imposing a stricter set of exclusions. Thus, the analogy highlights a conceptual parallel. Both systems regulate interaction patterns to achieve stable, coherent outcomes, yet they do so with differing degrees of strictness and mechanism.

Further directions

As leading figures in AI research have noted, simply scaling up large language models may yield diminishing returns, suggesting the need for conceptual and structural innovations beyond mere size expansion [27,28]. In this study, we have proposed drawing on anyonic systems from $SU(2)_k$ theories as one such innovative direction, arguing that their modular tensor structures may offer insights into the robustness, complexity, and adaptability of advanced LLMs.

Existing research provides intriguing connections between topological constructs and computational frameworks. For example, [44] demonstrates a deep correspondence between TQFTs and modular tensor categories via $SL(2,\mathbb{C})$-flat connections on three-manifolds, while the Painlevé VI equation and related monodromy problems link the geometry of these manifolds to quantum and conformal field theories [45,46]. Meanwhile, classical neural networks can emulate certain topological features [47] and might benefit from integrating hyperbolic geometry or other geometric insights [48]. Moreover, [17] suggests that framing deep neural networks as semi-classical limits of topological quantum neural networks could clarify their generalization properties, aligning closely with our perspective that $SU(2)_k$ anyons ,equipped with a structured fusion and braiding language, could inform a more mathematically grounded understanding of LLM internal representations.

However, this work remains conceptual. We have not provided empirical demonstrations or implemented topological neural architectures to test whether anyon-inspired fusion rules improve explainability or robustness in practice. The complexity of embedding $SU(2)_k$ structures into existing architectures is non-trivial, and achieving scalable, industrial-level applications of these ideas poses significant engineering and theoretical challenges. Realizing these concepts within standard deep learning frameworks, conducting controlled experiments, and measuring improvements in interpretability or performance represent natural next steps.

To move forward, future research might:

- Develop toy models or simplified neural architectures incorporating elementary $SU(2)_k$-like fusion rules. Such experiments could reveal whether these topological constraints yield more interpretable latent representations.
- Connect these theoretical constructs to empirical tools—such as probing classifiers, representational similarity analysis, or topological data analysis pipelines—to quantitatively assess how anyon-inspired transformations affect learned representations.

- Foster interdisciplinary collaboration among quantum information theorists, topologists, and machine learning practitioners to translate these concepts into actionable methodologies, thereby bridging abstract theory and practical XAI techniques.

By pursuing these directions, we may transform the conceptual analogies presented here into tangible strategies for building more transparent, robust, and theoretically grounded AI systems.

Author Contributions: Conceptualization, M.P.; methodology, M.P. and M.A.; software, M.P.; validation, M.P. and M.A.; formal analysis, M.P.; investigation, M.P. and M.A.; writing—original draft preparation, M.P.; writing—review and editing, M.A.; visualization, M.A.; supervision, M.P; project administration, M.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are available from the authors upon reasonable request.

Acknowledgments: The first author would like to acknowledge the contribution of the COST Action CA21169, supported by COST (European Cooperation in Science and Technology).

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix E Pseudocode

.

```
(* Magma Pseudocode for generating fusion tables in an SU(2)_k theory *)

function AllowedSpins(k):
    spins_list = []
    for i = 0 to k:
        append i to spins_list
    end for
    return spins_list
end function

// Initialize allowed spins for given k (scaled by factor of 2)
spins = AllowedSpins(k)  // e.g., for k=4, spins = [0,1,2,3,4]

function Fusion(j1, j2, k):
    // Compute fusion product for two scaled spins j1 and j2
    min_val = min(j1 + j2, 2*k - j1 - j2)
    fusion_result = []

    // Step by 2 to account for half-integers
    for j = |j1 - j2| to min_val step 2:
        append j to fusion_result
    end for

    return fusion_result
end function

// Create fusion table as a dictionary keyed by (j1,j2)
fusion_table = {}

for each j1 in spins:
    for each j2 in spins:
        fusion_result = Fusion(j1, j2, k)
        fusion_table[(j1,j2)] = fusion_result
    end for
end for

// Print results, scaling back to original spins by dividing by 2
for each j1 in spins:
    for each j2 in spins:
```

```
                result_list = fusion_table[(j1,j2)]
                original_spins_result = []
                for each r in result_list:
                    append (r/2) to original_spins_result
                end for
                print "Fusion of", (j1/2), "and", (j2/2), ":", original_spins_result
            end for
        end for
```

References

1. Pause Giant AI Experiments: An Open Letter. Available online: https://en.wikipedia.org/wiki/Pause_Giant_AI_Experiments:An_Open_L (accessed on 1 October 2024).
2. Bryant, A. AI chatbots: threat or opportunity? *Informatics* **2023**, *10*, 49.
3. Confalioneri, R.; Coba, L.; Wagner, B.; Besold, T. R. A historical perspective of explainable artificial intelligence. *Data Mining Know. Discov.* **2019**, *11*, e1391.
4. Ahmed I.; Jeon, G.; Piccialli, F. From artificial intelligence to explainable artificial intelligence in industry 4.0: A survey on what, how, and where. *IEEE Trans. Indust. Inform.* **2022**, *18*, 5031–5042.
5. Vilone, G.; Longo, L. Classification of explainable artificial intelligence methods through their output formats. *Mach. Learn. Knowl. Extr.* **2021**, *3*, 615–661.
6. Ferraz-Caetano, J. The artificial intelligence explanatory trade-Off on the logic of discovery in chemistry. *Philosophies* **2023**, *8*, 17.
7. Tahir, H. A.; Alayed, W.; Hassan, W. U.; Haider, A. Proposed explainable interference control technique in 6G networks using large language models (LLMs). *Electronics* **2024**, *13*, 4375.
8. Ferri, M. Why Topology for machine learning and knowledge extraction? *Mach. Learn. Extr.*, **2019**, *1*, 115–120.
9. Holzinger, A. Introduction to Machine Learning & Knowledge Extraction (MAKE). *Mach. Learn. Extr.*, **2019**, *1*, 1–20.
10. Feng, J. H.; Lai, E. M-K; Li, W. Geometry of textual data augmentation: Insights from large language models.*Electronics* **2024**, *13*, 3781.
11. Lozić, E.; Štular, B. Fluent but not factual: a comparative analysis of ChatGPT and other AI chatbots' proficiency and originality in scientific writing for humanities. *Future Internet* **2023**, *15*, 336.
12. Gams, M.; Kramar, S. Evaluating ChatGPT's consciousness and its capability to pass the Turing test: a comprehensive analysis. *J. Comp. Comm.* **2024**, *12*, 219–237.
13. Chen, S.; Yu, S.; Zhao, S.; Lu, C. From imitation to introspection: probing self-consciousness in language models. *arXiv* **2024**, arXiv: 2410.1881 [cd.CL].
14. Pira, L.; Ferrie, C. On the interpretability of quantum neural networks. *arXiv* 2024, arXiv: 2308,11098 [quant-ph].
15. Steinmüller, P.; Schulz, T.; Graf, F.; Herr, D. eXplainable AI for quantum machine learning. *arXiv* 2022, arXiv: 2211,01441 [quant-ph].
16. Marciano, A.; Chen, D. Fabrocini, F.; Fields C.; Greco, E.; Gresnight, N. ; Kinklub, K.; Lulli, M.; Terzidis, K.; Zappala, E. Quantum neural networks and topological quantum field theories. Neural Networks **2022**, *153*, 164–178.
17. Tommaso, T.. Interpretability and explainability AI in advanced neural Networks for topological quantum field Theory and medical application. Available at https://inspirehep.net/literature/2849341, accessed on 15 November 2024.
18. Field, B; Simula, T. Introduction to topological quantum computation with non-Abelian anyons. *Quantum Sci. Technol.* **2018**, *3*, 045004.
19. Wang, Z. Topological quantum Computing: American mathematical Society (Providence, Rhode Island, USA, **2010**, 115 pp.
20. Nayak, C.; Simon S. H.; Stern, A.; Freedman , M.; Das Sarma, S. Non-Abelian anyons and topological quantum computation. *Rev. Mod. Phys.* **2008**, *80*, 1083.
21. Biedenharn, L. C.; Lohe, M. A. Quantum Group Symmetry and q-Tensor Algebras, , World Scientific, 1995.
22. Kirchner, N.; Millar, D.; Ayeni, B. M.; Smith, A.; Slingerland, J. K.; Pollman, F. Numerical simulation of non-abelian anyons. *Phys. Rev. B* **2023**, *107*, 195129.
23. Xu, S.; Sun, Z. Z.; Wang, K.; Xiang, L.; Bao Z.; Zhu, Z.; Shen, F.; Song, Z.; Zhang, P.; Ren, W.; Zhang, X.. Dong, H.; Deng, J.; Chen, J.; Wu, Y.; Tan, Z.; Gao, Y.; Jin, F.; Zhu, X.; Zhang, C.; Wang, N.; Zou, Y.; Zhong, J.; Zhang, A.; Li, W.; Jiang, W.; Yu, L. W.; Yao, Y.; Wang, Z.; Li, H.; Guo, Q.; Song, C.; Wang, H.; Deng, D. L. Digital simulation of non-Abelian anyons with 68 programmable superconducting qubits. *Chinese Phys. Lett.* **2023**, *40*, 060301.
24. Iqbal, M.; Tantivasadakarn, N.; Verresen, R.; Campbell, S. L.; Dreiling, J. M.; Figgatt, C.; Gaebler, J. P.; Johansen, J.; Mills, M.; Moses, S. A.; Pino, J. M.; Ransford, A.; Rowe, M.; Siegfried, P.; Stutz, R. P.; Foss-Feig, M.; Vishwanath, A.; Dreyer, H. Non-abelian topological order and anyons on a trapped-ion processor. *Nature* **2024**, *626*, 505–511.
25. Rowell, E.; Stong, R.; Wang, Z. On the classification of modular tensor categories. *Comm. Math. Phys.* **2009**, *292*, 343–389.
26. Planat, M.; Chester, D. Topology and dynamics of transcriptome (dys)regulation. *Int. J. Mol. Sci.* **2024**, *25*, 4971.
27. Knight, W. (2023, April 17). OpenAI's CEO Says the Age of Giant AI Models Is Already Over. Wired. Retrieved November 11, 2024, available at https://www.wired.com/story/openai-ceo-sam-altman-the-age-of-giant-ai-models-is-already-over/

28. Villalobos, P.; Ho, A.; Sevilla, J.; Besiroglu, T; Heim, L.; Hobbhahn, M. Will we run out of data? Limits of LLM scaling based on human-generated data. *arXiv* **2024**, arXiv: 2211.04325.
29. Verlinde, E. Fusion rules and modular transformations in 2D conformal theory. *Nucl. Phys.* **1988**, B300, 360-376.
30. Gils, C.; Ardonne, E., Trebs, S.; Huse, D. A.; Ludwig A. WW, Troyer M.; Wang, Z. Anyonic quantum spin chains: Spin 1 generalizations and topological stability. *Phys. Rev. B* **2008**, *87*, 235120.
31. Kirillov, A. N.; Reshetikin, N. Y. Representations of the algebra $U_q(sl(2))$, q-orthogonal polynomials and invariants of links, in V. G. Kac, ed., in Infinite dimensional Lie algebras and groups (World Scientific, Singapore (1988), p. 285).
32. Johansen, E. G.; Simula, T. Fibonacci anyons versus Majorana fermions: a Monte Carlo approach to the compilation of braid circuits in $SU(2)_k$ anyon models. *PRX Quantum* **2021**, *2*, 010334.
33. Aboumrad, W. Quantum computing with anyons: an F-matrix and braid calculator. *arXiv* **2022**, arXiv: 2212.0083 [math.QA].
34. Amaral, M.; Chester, D.; Fang, F.; Irwin, K. Exploiting anyonic behavior of quacicrystals for topological qiantum computing. *Symmetry* **2022**, 14, 1780.
35. Bauer, B.; Levaillant, C. A new set of generators and a physical interpretation for the $SU(3)$ finite subgroup $D(9,1,1;2,1,1)$. *Quantum Inf. Proc.* **2013**, *7*, 2509–2521.
36. Levaillant, C. On some projective unitary qutrit gates. *arXiv* **2013**, arXiv: 1401.0506 [math.QA].
37. Planat, M.; Aschheim, R.; Amaral, M. M.; Irwin, K. Informationally complete characters from quark and lepton mixings. *Symmetry* **2020**, *12*, 1000.
38. Lillicrap, T. P.; Santoro, A.; Marris, L.; Akerman, C. J.; Hinton, G. Back propagation and the brain. *Nat. Rev. Neurosci.* **2020**, *21*, 335–346.
39. Hofer, C.; Kwitt, R.; Niethammer, M.; Uhl, A. "Deep Learning with Topological Signatures." In Advances in Neural Information Processing Systems (NeurIPS) 30, pp. 1633–1643, 2017.
40. Zheng, Z.; Wang, Y.; Huang, Y.; Song, S.; Yang, M.; Tang, B.; Xiong, F.; Li, Z. Attention heads of large language models: a survey. *arXiv* **2024**, arXiv: 2409. 03752 [cs.CL].
41. Khurana, D.; Koli, A.; Khatter, K.; Singh, S. Natural language processing: state of the art, curren trends and challenges. *Multimedia Tools and Appl.* **2022**, *82*, 3713–3744.
42. Otter, D. W; Medina, J. R.; Kalita, J. K. A survey of usages of deep learning for natural language processing. *IEEE Trans. on Neural Networks* **2021**, *32*, 604–624.
43. Raynal, M. Concurrent programming: algorithms, principles and foundations. Springer (Heidelberg, Germany, **2013**).
44. Cui, S. X; Qiu, Y.; Wang Z. From three dimensional manifolds to modular tensor categories. *Comm. Math. Phys.* **2023**, *397*, 1191–1235.
45. Boalch, P. From Klein to Painlevé via Fourier, Laplace and Jimbo. *Proc. Lond. Math. Soc.* **2005**, *90*, 167-208.
46. Planat, M.; Chester, D. Irwin, K. Dynamics of Fricke-Painlevé surfaces. *Dynamics* **2024**, *4*, 1–13.
47. Deng, D. L.; Li, X.; Das Sarma, S. Machine learning topological states. *Phys. Rev. B* **2017**, *96*, 195145.
48. Sanborn, S.; Mathe, J.; Papillon, M.; Buracas, D.; Lillemark, H. J.; Showmake, C.; Bertics, A.; Pennec, X.; Miolane, N. Beyond Euclid: an illustrated guide to modern machine learning with geometric, topological, and algebraic structures. *arXiv* **2024**, arXiv: 2407.09468 [cs.LG].