

2025-12 · V1.0.0

Dynamic Optimality in Physical and Arithmetic Systems

QNFO Research

QNFO-ULA-2.1

ABSTRACT

The selection of an appropriate level of abstraction constitutes a critical and profoundly non-trivial challenge across the entire scientific enterprise. This fundamental problem of representational scale dictates not only the tractability of our models but also their ultimate predictive power in describing the world. From the coarse-grained variables of thermodynamics to the aggregated metrics of macroeconomics, scientific understanding is built upon a hierarchy of descriptions, none of which holds a universal claim to primacy. The core tension arises from a necessary trade-off between the detailed fidelity afforded by a microscopic representation and the robust generalizability often found in a more abstract, coarse-grained one. Choosing the correct scale is not a matter of preference but a central task in the construction of any effective theory, as this decision determines which patterns in a system are identified as signal and which are relegated to the status of noise.

CITE AS Research, QNFO (2025). Dynamic Optimality in Physical and Arithmetic Systems. QNFO Preprint v1.0.0.

Dynamic

Optimality in Physical and Arithmetic Systems

Author: Rowan Brad Quni-Gudzinis **Contact:** rowan.quni@outlook.com **ORCID:** 0009-0002-4317-5604 **ISNI:** 0000000526456062 **URL:** <https://papers.qnfo.org/papers/dynamic-optimality-in-physical-and-arithmetic-systems/>
Date: 2025-12-21 **Version:** 1.0 **Abstract:** The selection of an appropriate abstraction level is a fundamental challenge in scientific modeling. We operationalize this trade-off with “predictive efficiency,” a quantitative objective function defined as the ratio of predictive fidelity to computational cost ($O = F/C$). We advance the dynamic optimality hypothesis, which asserts that the optimal representational scale for a complex system is not static but follows a predictable, non-monotonic “arc of representation” as the system evolves from disorder to order. To validate this, we employ a dual-track strategy, applying an identical metric to a physical spin chain and an arithmetic state space based on arithmetic topology. Both systems reveal a striking structural isomorphism, exhibiting a U-shaped arc: beginning at a mesoscale optimum, collapsing to the microscale during a chaotic predictive trough, and culminating in macroscopic causal emergence. We further link this information-theoretic efficiency to physical robustness using higher-order tensor stability analysis, demonstrating that optimal macroscopic states correspond to stable local minima ($\lambda_{min} > 0$). The findings suggest dynamic, multi-scale adaptation is a universal requirement for efficient modeling in both natural and artificial systems.

Keywords: Causal Emergence, Predictive Efficiency, Arithmetic Topology, Dynamic Optimality, Tensor Stability, Renormalization Group, Complex Systems.

1.0 Introduction and Problem

Statement

1.1 The Representational

Scale Dilemma

The selection of an appropriate level of abstraction constitutes a critical and profoundly non-trivial challenge across the entire scientific enterprise. This fundamental problem of representational scale dictates not only the tractability of our models but also their ultimate predictive power in describing the world. From the coarse-grained variables of thermodynamics to the aggregated metrics of macroeconomics, scientific understanding is built upon a hierarchy of descriptions, none of which holds a universal claim to primacy. The core tension arises from a necessary trade-off between the detailed fidelity afforded by a microscopic representation and the robust generalizability often found in a more abstract, coarse-grained one. Choosing the correct scale is not a matter of preference but a central task in the construction of any effective theory, as this decision determines which patterns in a system are identified as signal and which are relegated to the status of noise.

The historical context for this problem is the long-standing tension between reductionist and emergentist paradigms in science. While the reductionist viewpoint, which privileges the most fundamental description, has been extraordinarily successful, it struggles to account for phenomena where the collective behavior of a system is simpler and more predictable than its individual components. The school of thought surrounding causal emergence has provided a powerful theoretical counterpoint, arguing that macroscopic states can, in an information-theoretic sense, possess more causal power than their microscopic constituents (Hoel, Albantakis, & Tononi, 2013). This perspective suggests that coarse-graining is not merely a convenience for computationally-limited observers, but can be an objective method for isolating the true causal structure of a system. However, this theoretical insight has not yet been translated into a general, dynamic methodology for model selection in evolving systems.

The mechanism governing the optimal choice of scale is a delicate balance between competing virtues: detail and simplicity. A model with an extremely high level of detail, one that tracks every degree of freedom, risks overfitting to the contingent, noisy aspects of a system's state. Such a model may provide a perfect description of the past but have little to no power in predicting the future, as it mistakes random fluctuations for meaningful patterns. Conversely, a model that is excessively coarse-grained, averaging away too much information, risks underfitting the problem by discarding the very structures that govern the system's dynamics. The optimal representation, therefore, must be a compromise, retaining just enough detail to capture the causally significant patterns while discarding as much irrelevant information as possible.

The necessity for a new approach is underscored by persistent methodological gaps in the scientific literature. Prior theoretical treatments of representation, while conceptually rich, have often been non-constructive, providing abstract principles without concrete, falsifiable algorithms. There is a consistent demand in applied and computational fields for frameworks that are not just explanatory but also executable and benchmarkable against existing methods. This study aims to address this need by transforming the abstract problem of scale selection into a concrete optimization problem with a clear, algorithmic solution.

A persistent counter-argument, rooted in a deep-seated reductionist philosophy, posits that the most microscopic model is always the most fundamental and therefore the most correct representation of reality. From this perspective, all higher-level descriptions are merely convenient, epistemic shortcuts for observers who lack the capacity to engage with the full complexity of the microstate. This view holds that any failure of the microscopic model to be predictive is not a flaw in the representation itself, but a failure of the analytical or computational tools brought to bear upon it.

We synthesize these competing views by distinguishing sharply between a representation that is fundamental and one that is predictively useful. While the microscopic state of a system is indeed fundamental in that it determines all higher-level properties, it does not follow that it is the most effective representation for prediction. As has been argued in the context of causal emergence, a macroscopic scale can sometimes have more causal power than its microscopic underpinning. The reason for this is that the macroscale can abstract away the indeterminism and noise of the micro-level, revealing a more deterministic and therefore more predictable causal structure.

The intuitive and historically contingent nature of choosing a representational scale can thus be replaced with a more rigorous approach. The trade-off between fidelity and simplicity is not merely a qualitative guideline but can be formalized as a quantitative optimization problem. By defining precise metrics for both the predictive power of a model and its inherent complexity, it becomes possible to calculate an optimal level of abstraction for a given system, leading to the formalization of “predictive efficiency.”

1.2 Predictive

Efficiency as an Objective Function

To adjudicate between representations at different scales, a clear and quantitative objective function is required; we propose that this function is *predictive efficiency*. Predictive efficiency is defined as the ratio of a model’s predictive power to its computational cost. This metric provides a normalized measure of a representation’s utility, answering the critical question of how much predictive benefit is gained for each unit of computational resource invested. A model is considered efficient not merely if it is accurate, but if it is accurate in a parsimonious way. The representation that maximizes this efficiency is, by definition, the optimal description of the system for a resource-bounded observer.

The concept of efficiency as a guiding principle in science has deep roots, originating in the mid-20th century with the foundational work of Claude Shannon and later formalized by Jorma Rissanen. Shannon’s rate-distortion theory established that for any given information source, there is an irreducible relationship between the fidelity of a representation (its “distortion”) and the complexity of that representation (its “rate”) (Shannon, 1959). Rissanen later adapted this concept for statistical modeling with the Minimum Description Length (MDL) principle, which posits that the best model is the one that minimizes the combined length of the description of the model itself plus the data encoded by it (Rissanen, 1978).

The mechanism for calculating predictive efficiency in our framework involves a three-step process. First, the predictive fidelity (F) is calculated by measuring the negative mean squared error of a forecast of the model’s next state. Second, the computational cost (C) is calculated as the dimensionality of the model’s state space. Third, the predictive efficiency (O) is computed as the simple ratio of these two quantities: $O = F/C$. This calculation provides an instantaneous snapshot of the model’s efficiency, creating a time-dependent metric that can track the optimal representation as the system evolves.

The explicit formula for the objective function is $O = F/(N/k)$, where N is the total system size and k is the coarse-graining block size. This formula serves as the computational engine of our analysis. By applying this formula to a hierarchy of models ranging from microscopic ($k = 1$) to macroscopic ($k = N$), we can generate a quantitative comparison of their efficiencies. This approach transforms the abstract philosophical debate about reductionism into a solvable mathematical problem.

A potential counter-argument to this specific formulation is that the linear ratio F/C is an arbitrary way to combine fidelity and cost. A critic might propose an alternative objective function, such as a weighted sum ($aF - bC$), arguing that different scientific contexts might place different weights on the relative importance of accuracy versus cost. From this perspective, our choice of a simple ratio embeds a particular set of priorities that may not be universally applicable, potentially biasing the selection of optimal models.

We synthesize this by defending the linear ratio F/C as the most direct and principled formulation of efficiency. It directly measures the “return on investment” in computational terms: how much fidelity do we get for each degree of freedom we choose to model? A weighted sum, by contrast, would require the introduction of arbitrary weighting parameters that would themselves need to be justified. The ratio is parameter-free and possesses a clear physical and economic interpretation, making it a robust standard for comparison.

In any resource-bounded system, whether biological or artificial, efficiency is the only pragmatic metric. Systems that waste computational resources on non-predictive noise are evolutionarily disadvantaged. Therefore, maximizing predictive efficiency is not just an engineering goal but a fundamental principle of adaptation and survival. This metric allows us to evaluate models not in a vacuum of infinite resources, but in the realistic context of physical constraints.

This objective function will be the primary tool used to evaluate the behavior of the systems investigated in this study. To demonstrate the universality of this metric, we propose to test it on two distinct tracks: a physical system governed by statistical mechanics, and an arithmetic system governed by number theory. The following section outlines this dual-track investigation strategy.

1.3 Dual-Track Investigation

Strategy

To rigorously validate the universality of the dynamic optimality hypothesis, this study employs a dual-track investigation strategy. We propose to apply our predictive efficiency framework to two systems that are superficially distinct but structurally isomorphic: a physical spin chain evolving in time, and an arithmetic state space evolving in complexity. The central thesis of this strategy is that if identical patterns of representational evolution (“arcs”) emerge in these two disparate domains, it provides strong evidence for a deep, substrate-independent principle of complexity organization.

The context for this dual-track approach is the historical divide between physics and abstract mathematics. Physics typically deals with dynamic, temporal systems governed by energy and entropy, while number theory deals with static, eternal structures governed by arithmetic axioms. However, the field of arithmetic topology suggests a profound analogy between these worlds, mapping prime numbers to topological knots (Li & Sia, 2012). By treating number theory as a dynamic system, we aim to bridge this gap and test whether the “laws of physics”—specifically, the laws of emergent complexity—apply to the “physics of mathematics.”

The mechanism of this strategy involves running two parallel simulations. Track A (Physical) simulates a 1D spin chain, a standard model in statistical physics, evolving from a random state to an ordered one. Track B (Arithmetic) simulates a “gas” of prime numbers, where the system “grows” by sequentially adding primes, and the state is defined by their pairwise arithmetic relationships (Legendre symbols). Both tracks will be subjected to the exact same analytical pipeline: constructing a hierarchy of coarse-grained models, calculating predictive efficiency, and identifying the optimal scale at each step.

The evidence generated by this strategy will be a direct, side-by-side comparison of the optimal representation trajectories. We will look for structural similarities in how the optimal scale k^* changes as the systems evolve. Does the arithmetic system exhibit a “predictive trough” like the physical one? Does it show a transition from microscopic to macroscopic dominance? Finding matching “arcs of representation” would empirically validate the structural analogies proposed by arithmetic topology.

A significant counter-argument is that these domains are too distinct to be meaningfully compared. A spin chain is a physical system with thermal noise and energy relaxation; the sequence of prime numbers is a deterministic mathematical structure. Any similarity in their “dynamics” could be dismissed as a coincidence or an artifact of the shared analytical method, rather than a sign of deep structural equivalence. Critics might argue that applying dynamical concepts to number theory is a category error.

We synthesize this by arguing that “dynamics” is a general concept applicable to any system that traverses a state space, whether that traversal is driven by time (in physics) or by increasing complexity (in mathematics). If the predictive efficiency metric reveals the same emergent behavior in both, it suggests that “causal emergence” is a property of

information processing itself, independent of the underlying substrate. The comparison forces us to look beyond the surface details of “spins” and “primes” to the underlying architecture of information.

1.4 Track A: *The Physical Spin*

Model

For the physical baseline of our investigation, we employ a one-dimensional cellular automaton—specifically, a chain of $N = 128$ interacting spins—as our model system. The thesis of this choice is that a 1D spin chain with local interactions and noise serves as the ideal “control system” for demonstrating causal emergence. It is a minimal model that is sufficiently complex to exhibit non-trivial ordering but sufficiently simple to allow for the exact calculation of the predictive efficiency landscape at every time step.

The context for using such models lies in the history of statistical physics, where simplified lattice models (like the Ising model or the Hubbard model) have long served as the primary tools for understanding phase transitions and collective behavior. Unlike idealized models that are often mathematically intractable or physically unrealizable in their pure form (Xu et al., 2025), our spin chain is designed to be a transparent, computational “toy model” where every variable and interaction is explicitly defined and accessible.

The mechanism driving this system is a local majority update rule combined with stochastic noise. At each time step, a spin updates its state based on the alignment of its neighbors (an ordering force) plus a random flip probability (a disordering force). This competition creates a dynamic landscape where small domains of order form, compete, merge, and eventually stabilize. This evolution mimics the phenomenology of real materials, such as the annealing of a metal or the magnetization of a ferromagnet, but in a simplified geometry.

The evidence provided by this model will be a time-series of “micro-states”—the exact configuration of all 128 spins at every step of the simulation. From this ground truth, we can rigorously test our hypotheses. Because we know the exact microscopic rules, any emergence of a macroscopic model as “optimal” cannot be attributed to hidden variables or unknown physics. It must be a consequence of the system’s intrinsic information structure.

A common counter-argument is that such 1D models are too simple to represent real-world complexity. In 1D, there are no complex geometric frustrations, and phase transitions are often trivial compared to 2D or 3D systems. Critics might argue that findings from this “toy model” will not generalize to the rich behavior of real complex systems, making the exercise academically interesting but practically limited.

We synthesize this by asserting that simplicity is a virtue when establishing a new theoretical framework. Before we can analyze the complexity of a brain or a global economy, we must first demonstrate that our metric works on a system we fully understand. The 1D spin chain is a “reductio ad absurdum” for reductionism: even in this simplest of worlds, we expect to see that the microscopic description is inefficient. If dynamic optimality holds here, it provides a firm foundation for scaling up to more complex topologies.

1.5 Track B: *The*

Arithmetic Topology Model

For the second track of our investigation, we construct a computational system based on the deep analogies of arithmetic topology. The thesis of this model is that the set of prime numbers can be treated as a dynamic system of interacting particles, where the “interaction” is defined by number-theoretic relationships. By translating the static properties of primes into a dynamic state space, we create a novel testbed for the principles of predictive efficiency, allowing us to ask whether the “evolution” of number theory exhibits the same causal structures as physical matter.

The context for this model is the field of arithmetic topology, which posits a structural equivalence between knots in a 3-manifold and prime ideals in a number ring (Li & Sia, 2012; Morishita, 2012). This field has established a “dictionary” of analogies: primes correspond to knots, and the Legendre symbol (which describes quadratic reciprocity) corresponds to the linking number between knots. Until now, this dictionary has been largely static and descriptive. Our work aims to make it dynamic and predictive.

The mechanism of the arithmetic model involves treating a sequence of prime numbers as the elements of our system. The “state” of the system is defined by the matrix of pairwise Legendre symbols between all primes in the set. The “evolution” of the system is simulated by progressively adding larger primes to the set, effectively growing the system’s complexity over “time.” Just as the spin chain evolves from random to ordered, this arithmetic system evolves from a small, sparse network of relations to a large, dense one.

The evidence provided by this model will be a trajectory of optimal representations for the arithmetic state. By applying our coarse-graining and efficiency metrics to the matrix of Legendre symbols, we can determine if there are scales of description—groups of primes—that are more predictable than individual primes. This would be the arithmetic equivalent of finding “domains” in the spin chain.

A significant counter-argument is that this is a category error: number theory is timeless and deterministic, not dynamic or causal. The “evolution” we simulate is merely an arbitrary ordering of static facts. Furthermore, the analogy between knots and primes is a formal mathematical correspondence, not a physical theory. Treating primes as “dynamical agents” may stretch the analogy beyond its breaking point, yielding results that are mathematical artifacts rather than physical insights.

We synthesize this by arguing that the distinction between “dynamic” and “static” is a matter of perspective. When we explore the number line, we are traversing a landscape of increasing complexity. Our “time” axis represents this traversal. By imposing a dynamic framework on this static structure, we are probing its information-theoretic depth. If the static analogy holds true, then the *informational structure* of the knot-like primes should manifest as similar “causal” patterns (like clustering or screening) when viewed through a predictive lens.

1.6 The Arc of

Representation Hypothesis

The central hypothesis of this investigation is that the optimal representational scale for any complex evolving system—whether physical or arithmetic—will not be static, but will follow a predictable, non-monotonic trajectory we term the “arc of representation.” This hypothesis addresses the temporal gap in current modeling theories, which often fail to account for the shifting representational needs of a system as it moves from disorder to order. We predict that the optimal scale will shift from a mesoscale (to capture initial fluctuations), down to the microscopic scale (during chaotic transitions), and finally up to a macroscopic scale (as stable order emerges).

The context for this hypothesis implies a move beyond static model selection. In most scientific disciplines, a model is chosen based on the general laws of the system. Our hypothesis suggests that the optimal model depends on the *state* of the system, not just its laws. A system governed by microscopic laws may, in certain phases, be best described by macroscopic variables, while in others, only the microscopic variables will suffice.

The mechanism driving this arc is the evolution of the system’s causal structure. In the early “genesis” phase, small clusters of order emerge, requiring a fine-grained but not microscopic view. In the intermediate “chaotic” phase, order breaks down or reorganizes, and high-entropy noise dominates, forcing the model to the microscopic limit to capture any signal at all. In the final “dominance” phase, large-scale order stabilizes, allowing a macroscopic model to efficiently capture the system’s behavior while discarding vast amounts of microscopic noise.

The evidence we seek is a specific, “U-shaped” or “inverted-U” trajectory in the optimal block size k^* over time. Preliminary theoretical considerations of causal emergence suggest that macro-scales can have higher effective information, but our hypothesis adds the temporal dimension: this advantage is phase-dependent. Observing this specific arc in both the spin chain and the prime number system would constitute strong validation of a universal dynamic.

A counter-argument is that for a fixed set of dynamical rules, there should be a single, time-invariant “effective theory” that is optimal. The fluctuations we predict—particularly the return to microscopic optimality during chaos—might be viewed as artifacts of an imperfect predictor rather than a property of the system. Reductionists would argue that if the microscopic model is optimal at any point, it proves it is the fundamental description, and any deviation is merely a loss of precision.

We synthesize this by asserting that the “optimal” model is defined by efficiency, not just accuracy. During chaotic phases, the efficiency of all models drops, and the microscopic model wins only by default. During ordered phases, the macroscopic model wins by design. The arc of representation traces the changing “compressibility” of the system’s state. It is a map of where the system’s meaningful information resides at each moment in its history.

1.7 Robustness and Stability

Criteria

To complete our theoretical framework, we must define what constitutes a “robust” representation. We posit that robust representations correspond to stable, topologically or spectrally protected physical states. In our dual-track investigation, we will operationalize this by linking high predictive efficiency to mathematical stability, specifically using tensor eigenvalue analysis. We hypothesize that the optimal macroscopic states identified by our efficiency metric will also be the states that minimize the system’s potential energy in a higher-order tensor landscape.

The context for this criterion is the search for noise-resilient information storage, a critical challenge in quantum computing and memory storage. Physical systems that encode information in topological invariants or spectral gaps are inherently robust against local perturbations. We seek to translate this physical robustness into a mathematical stability criterion that can be applied to both our spin chain and our arithmetic model.

The mechanism we employ relies on the mathematics of higher-order tensors. We construct a 4th-order supersymmetric tensor from the state vector of our system. According to the work of Qi (2005), the positive definiteness of such a tensor—indicated by its minimum H-eigenvalue being positive—is a rigorous condition for stability in non-linear systems. By calculating this eigenvalue for our optimal representations, we can mathematically test their stability.

The evidence we aim to find is a strong correlation: we predict that the representations with the highest predictive efficiency will also be the ones with positive H-eigenvalues. This would link the information-theoretic concept of “efficiency” with the physical/mathematical concept of “stability.” It would suggest that the arc of representation is not just tracing predictive power, but is tracking the system’s settling into stable, robust configurations.

A counter-argument is that mathematical stability in a tensor model does not necessarily equate to physical robustness or representational utility. The tensor construction might be an ad-hoc mapping that forces stability, rather than revealing it. Furthermore, relating tensor eigenvalues to the “robustness” of a prime number representation is a highly speculative leap that may lack physical meaning.

We synthesize this by arguing that in a unified theory of complex systems, “stability” and “predictability” must be two sides of the same coin. A system is predictable *because* it is stable; it is efficient to represent *because* it has settled into a robust state. The tensor eigenvalue analysis is a rigorous way to test this intuition across domains. If the arithmetic model shows the same stability correlations as the physical one, it strongly supports the view that “indivisibility”—whether of a knot, a prime, or a macroscopic domain—is a form of robustness.

2.0 Literature Review

2.1 Foundations of

Information Efficiency

The theoretical bedrock for quantifying the trade-off between representational fidelity and computational cost was established in the mid-20th century by the foundational works of Claude Shannon and Jorma Rissanen. Their contributions provided the first rigorous mathematical language for describing the limits of information transfer and model selection, effectively transforming the philosophical principle of parsimony into a computable metric. This body of work forms the conceptual basis for the “predictive efficiency” objective function used in our dual-track investigation, providing the axiomatic justification for treating cost reduction as a fundamental goal of representation rather than a mere engineering convenience.

The historical context for these developments was the nascent field of information theory, which sought to define the absolute limits of communication channels. Shannon (1959), in his seminal work on rate-distortion theory, demonstrated that for any given information source, there exists an irreducible relationship between the “rate” (the number of bits used to describe the data) and the “distortion” (the error in the reconstructed signal). This theorem proved that perfect representation is physically impossible under bandwidth constraints, necessitating a principled approach to lossy compression. This was a radical departure from the classical ideal of exactitude, introducing the concept that the “best” representation is a function of the available resources.

The mechanism central to this theoretical tradition is the formalization of the “rate-distortion curve,” a boundary that defines the optimal performance of any compression scheme. Rissanen (1978) later adapted this concept for statistical inference with the Minimum Description Length (MDL) principle. MDL posits that the best statistical model for a dataset is the one that minimizes the sum of the length of the model description and the length of the data encoded by that model. By treating the model itself as information that must be transmitted, Rissanen provided a concrete cost function that penalizes unnecessary complexity, effectively creating a mathematical razor to slice away non-predictive variables.

These principles provide the essential theoretical evidence for the validity of our “F/C” metric. The numerator of our objective function, Predictive Fidelity (F), corresponds to the inverse of Shannon’s distortion. The denominator, Computational Cost (C), corresponds to Shannon’s rate or Rissanen’s description length. By maximizing the ratio of these terms, our framework explicitly seeks the optimal point on the rate-distortion curve for a dynamic system. The MDL principle, in particular, offers a strong theoretical argument that the “true” structure of a system is best revealed by the representation that compresses it most efficiently.

A significant counter-argument to the direct application of these theories in complex systems science is that they are fundamentally non-constructive and often assume stationary statistics. Rate-distortion theory proves the existence of an optimal code but provides no universal algorithm for generating it. Similarly, the ideal formulation of MDL relies on Kolmogorov complexity, which is uncomputable. Furthermore, these theories were originally formulated for static information sources, limiting their direct applicability to non-equilibrium systems where the statistical properties of the “source” (the system state) are constantly evolving (Shannon, 1959).

In synthesis, while these foundational theories are abstract and possess known limitations regarding constructability, they define the essential optimization landscape for any problem of representation. They establish that the tension between accuracy and complexity is not an artifact of human cognition but a fundamental property of information itself. Our work can be viewed as a dynamic, constructive extension of this landscape, using computational simulation to empirically locate the optimal points that Shannon and Rissanen defined theoretically, but could not compute for evolving complex systems.

This information-theoretic foundation explains *how* to measure the efficiency of a representation, but it does not explain *why* macroscopic representations of physical systems often yield higher efficiency. To understand this, we must turn to the theory of causal emergence, which provides the physical grounding for why “less” can sometimes be “more” in the context of prediction.

2.2 Causal Emergence Theory

The theory of causal emergence, particularly as formalized by Hoel, Albantakis, and Tononi (2013), provides a rigorous, information-theoretic explanation for the counter-intuitive phenomenon where macroscopic descriptions of a system can be more predictively powerful than microscopic ones. This work challenges the deep-seated reductionist assumption that the most detailed description is always the most scientifically valuable. It offers a formal language to describe why coarse-graining is not merely a data compression technique, but a method for enhancing the causal efficacy of a model.

The context for this work is the interdisciplinary field of complexity science, which seeks to understand how higher-level order arises from lower-level interactions. Historically, “emergence” was often treated as a vague or mystical concept. The contribution of Hoel et al. was to operationalize emergence using the tools of information theory, specifically mutual information and channel capacity. This shifted the debate from qualitative philosophy to quantitative analysis, allowing researchers to measure the “causal power” of different descriptions of the same system.

The mechanism proposed by Hoel et al. (2013) relies on a metric called “Effective Information” (EI). EI measures how effectively a system’s current state dictates its future state—essentially, it is a measure of determinism and non-degeneracy in the system’s transition dynamics. The key insight is that microscopic states in complex systems are often plagued by noise and degeneracy; a specific micro-state might transition to many possible future micro-states with low probability. By coarse-graining these states into macro-states, one can often average out this indeterminism, creating a macroscopic transition map that is sharper, more deterministic, and thus possesses higher Effective Information.

The evidence for this phenomenon was originally demonstrated on discrete Markovian systems, where it was shown that for certain network topologies, the EI peaked at a macro-scale. This provides the theoretical underpinning for the “macroscopic dominance” phase we expect to observe in our simulations. When our predictive efficiency metric identifies a coarse-grained model (e.g., $k=32$) as optimal, it is detecting precisely this peak in Effective Information. The macro-model wins because it has filtered out the noise that obscures the causal structure at the micro-level.

A common counter-argument to causal emergence is that it is restricted to simple, discrete systems and does not apply to the continuous, dynamic reality of physics. Critics argue that in a deterministic universe (like one governed by Hamiltonian mechanics), the micro-state determines the future perfectly, so EI should always be maximal at the micro-scale. From this perspective, causal emergence is an artifact of treating open, noisy systems as closed ones, or of ignoring the full details of the micro-dynamics (Hoel et al., 2013).

Our synthesis addresses this by emphasizing that all real-world modeling tasks involve resource constraints and observation limits. Even if the universe is deterministic at the Planck scale, any predictive model operating above that scale must contend with effective noise. In this practical regime, causal emergence is a robust physical phenomenon. Our work extends the validation of this theory by applying it to a continuous, dynamically evolving spin system, testing whether the principle holds outside the static Markovian networks where it was first defined.

While causal emergence explains the utility of macroscopic physical states, our dual-track investigation also requires a theoretical basis for treating arithmetic structures as dynamic systems. This basis is found in the field of arithmetic topology, which provides the dictionary for translating between the physical world of knots and the mathematical world of primes.

2.3 Arithmetic Topology:

Knots and Primes

The field of arithmetic topology provides the crucial theoretical bridge for our Track B investigation, establishing a deep structural equivalence between the objects of low-dimensional topology (knots) and the objects of algebraic number theory (primes). This “dictionary of analogies,” developed over decades by mathematicians such as Mazur, Manin, and Kapranov, posits that the seemingly disparate worlds of geometric shape and arithmetic value share a common underlying architecture. This correspondence is not merely a superficial similarity but a rigorous isomorphism of invariants, motivating our hypothesis that primes can be modeled as interacting “particles” in a dynamic system (Li & Sia, 2012; Sikora, 2001).

The context for arithmetic topology is the unification of mathematics. Since the 1960s, researchers have noticed striking parallels between the behavior of prime ideals in a number ring and knots embedded in a 3-manifold. For instance, the way a prime number decomposes into ideal factors when the number field is extended mirrors the way a knot splits into multiple components when lifted to a covering space. These observations led to a formalized program to map concepts from knot theory (like the Alexander polynomial) directly to concepts in number theory (like the Iwasawa polynomial).

The mechanism of this analogy is a precise mapping of spaces and objects. The spectrum of the integers, $\text{Spec}Z$, is identified as the arithmetic analogue of the 3-sphere, S^3 . Within this “arithmetic 3-sphere,” prime numbers p are analogous to closed loops or knots K . The relationship between two primes, defined by the Legendre symbol (p/q) , is mapped to the linking number $lk(K, L)$ between two knots (Sikora, 2001). This mapping allows us to treat the matrix of Legendre symbols in our simulation as a direct analogue of a topological linking matrix, providing the structural rules for our arithmetic state space.

The evidence for the validity of this analogy lies in the successful translation of theorems between the two fields. For example, the Gauss linking integral in topology has a direct counterpart in the reciprocity laws of number theory. The classification of knots by their fundamental groups mirrors the classification of number fields by their Galois groups (Li & Sia, 2012). These deep structural matches provide the justification for treating the set of primes as a coherent system with “topological” properties that can be analyzed using our predictive efficiency framework.

A significant counter-argument is that this is a category error: number theory is timeless and deterministic, not dynamic or causal. The “evolution” we simulate is merely an arbitrary ordering of static facts. Furthermore, the analogy between knots and primes is a formal mathematical correspondence, not a physical theory. Treating primes as “dynamical agents” may stretch the analogy beyond its breaking point, yielding results that are mathematical artifacts rather than physical insights.

We synthesize this by arguing that the distinction between “dynamic” and “static” is a matter of perspective. When we explore the number line, we are traversing a landscape of increasing complexity. Our “time” axis represents this traversal. By imposing a dynamic framework on this static structure, we are probing its information-theoretic depth. If the static analogy holds true, then the *informational structure* of the knot-like primes should manifest as similar “causal” patterns (like clustering or screening) when viewed through a predictive lens.

This analogy gives us the objects of our study, but to understand their stability, we need a physical model of robustness. The next section reviews the literature on how physical systems encode information in robust, “indivisible” states, providing the criteria for evaluating the stability of our optimal representations.

2.4 Topological and Spectral

Physics

The concept of “robustness” in our framework is grounded in the physics of topological protection and spectral rigidity. Literature from quantum chaos and soft matter physics suggests that the most stable forms of information storage in nature are not discrete bits, but global properties that are immune to local perturbations. This provides the physical model for what we term “indivisible” representations: states that maintain their identity despite the noise of the underlying micro-system. This literature motivates our search for stability in both the physical spin chain and the arithmetic model.

The context for this research is the ongoing effort to build fault-tolerant quantum computers and robust memory devices. Traditional “local” storage is fragile; a single bit-flip corrupts the information. In contrast, topological states—such as the braiding of anyons or the configuration of a vortex knot—store information non-locally. To destroy the information, one must perform a global deformation of the system, making these states inherently protected against local noise. Similarly, in quantum chaotic systems, the energy levels (spectrum) exhibit “rigidity,” meaning they resist fluctuation, a property conjectured to be related to the distribution of prime numbers (Berry & Keating, 1999).

The mechanism of topological protection is exemplified by the “heliknotons” recently discovered in chiral nematic liquid crystals (Hall et al., 2025). These are microscopic vortex knots that can be manipulated, fused, and split, yet retain their topological identity (Hopf index) throughout these dynamics. This conservation law acts as a stabilizing force, allowing the knot to persist as a coherent entity—a “particle”—despite the fluid nature of the medium. In the spectral domain, the Berry-Keating Hamiltonian $H = xp + px$ provides a model where the eigenvalues correspond to Riemann zeros, suggesting that the “indivisibility” of a prime is physically manifest as a stable energy level in a chaotic system.

The evidence from these fields provides concrete physical analogues for our optimal representations. When our simulation identifies a macroscopic block $k = 32$ as optimal, we hypothesize that this block corresponds to a robust “domain” that behaves like a topological soliton. It is stable against the flickering of individual spins (local noise) just as a knot is stable against local vibrations of the string. The work of Hall et al. (2025) proves that such robust, manipulable topological objects exist in real physical systems, validating our search for them in simulation.

A counter-argument is that these physical examples are too specialized to serve as a general model for representation. Topological protection requires specific exotic phases of matter (like fractional quantum Hall states or chiral liquid crystals), and spectral rigidity is a property of quantum chaos. It is not clear that these esoteric physics principles have any relevance to the general problem of data representation or to the behavior of a simple spin chain.

We synthesize this by arguing that “topological protection” is a specific instance of a general principle of robust coarse-graining. Any effective macroscopic variable that is insensitive to microscopic fluctuations is, in a functional sense, “topologically protected” by the law of large numbers or by the system’s energy landscape. By linking our computational findings to these rigorous physical concepts, we provide a deep theoretical grounding for why certain representations are stable.

To mathematically verify this stability in our simulations, we require a tool that can assess the robustness of a state without relying on visual inspection. This leads us to the mathematics of higher-order tensors.

2.5 Higher-Order Tensor

Stability

To rigorously quantify the stability of the representations identified in our simulations, we turn to the mathematics of higher-order tensor eigenvalues. Standard linear algebra (matrix mechanics) is insufficient for describing the stability of complex, non-linear systems where potential landscapes are not simple quadratic bowls. The work of Qi (2005) and

others on the eigenvalues of supersymmetric tensors provides the necessary formalism to determine if a complex, multi-variable state corresponds to a stable local minimum, bridging the gap between our information-theoretic efficiency metric and physical stability criteria.

The context for this mathematical framework is the study of non-linear elasticity, entanglement in quantum systems, and high-dimensional optimization. In these fields, the “state” is often described by a tensor of order 3 or higher, rather than a matrix (order 2). Determining the stability of such a system—whether it will return to equilibrium after a perturbation—requires generalizing the concept of “positive definiteness” from matrices to tensors.

The mechanism we employ is based on Qi’s Theorem (2005), which states that an even-order supersymmetric tensor is positive definite if and only if all of its “H-eigenvalues” are positive. H-eigenvalues are real numbers λ associated with real eigenvectors x that satisfy the polynomial equation $Ax^{m-1} = \lambda x$. In our framework, we map the state of our system (the vector of domain correlations or prime linking numbers) to a 4th-order tensor that represents the system’s potential energy landscape. Calculating the minimum H-eigenvalue of this tensor provides a binary verdict on stability: if $\lambda_{min} > 0$, the state is stable.

The evidence for the utility of this approach comes from its successful application in solid mechanics and control theory. By adopting this rigorous mathematical standard, we avoid vague hand-waving about “robustness.” We can calculate a precise number that indicates stability. This allows us to test the hypothesis that the “most efficient” representations (identified by $O=F/C$) are also the “most stable” ones (identified by $\lambda_{min} > 0$).

A counter-argument is that introducing higher-order tensors adds unnecessary mathematical complexity to the model. Many systems can be adequately linearized around a fixed point, making standard matrix eigenvalues sufficient. Furthermore, the construction of the tensor from our state vector is an ansatz—a hypothesized mapping—rather than a derivation from first principles. Critics might argue that any stability found is an artifact of this construction rather than a property of the system itself.

In our synthesis, we assert that the non-linear nature of the phase transitions in our systems—both the chaotic coalescence in the spin chain and the complex reciprocity in the arithmetic model—demands a non-linear stability analysis. Linearization fails precisely at the critical points we are most interested in. While the tensor construction is indeed an ansatz, it is a principled one that allows us to test for higher-order correlations that a matrix analysis would miss. It provides a consistent mathematical ruler to measure stability across both tracks of our investigation.

Having established the tools for analyzing stability, we must finally address the limitations of the models we are replacing. The final section of our review critiques the static idealizations that have historically dominated the field, justifying our move to dynamic, generative frameworks.

2.6 Critique of Static

Idealizations

Our investigation is motivated by the demonstrable failure of static, idealized models to predict the behavior of complex, non-equilibrium systems. A growing body of research in condensed matter physics indicates that traditional frameworks, which assume fixed geometric symmetries (like the square lattice of the Hubbard model), cannot account for the dynamic heterogeneities and topological phase transitions observed in real materials. This critique justifies our abandonment of static model selection in favor of a dynamic, adaptive framework (Xu et al., 2025; Jin et al., 2025).

The context for this critique is the “crisis of complexity” in modern physics. Simple models like the Ising or Hubbard models have been incredibly successful at explaining equilibrium phenomena. However, as experiments move to non-equilibrium regimes—such as ultrafast laser pulse excitation or driven diffusive systems—these models break down. They fail because they treat the system’s constraints (symmetries, lattice structure) as immutable background features, rather than as emergent properties that can evolve or be broken.

The mechanism of this failure is the inability of static models to represent “phase slippage” or topological defects that arise during dynamic evolution. For example, recent work by Jin et al. (2025) on “temporal anti-parity-time symmetry” shows that by actively modulating a system’s parameters, one can reverse heat flow—a phenomenon impossible in a static thermal model. Similarly, Xu et al. (2025) showed that the “ideal” Hubbard model does not exist in nature but must be artificially engineered in Moiré superlattices, suggesting that the model is a “Platonic” ideal rather than a description of natural “Aristotelian” matter.

The evidence for the need for generative models comes from these discrepancies. When a static model fails, researchers often add “epicycles”—extra parameters—to fit the data. A generative approach, which derives the dynamics from a continuous field (as we do with our phase field/spin chain), naturally accommodates symmetry breaking and topological defect formation without ad-hoc additions.

A counter-argument is that idealized models are valuable precisely *because* they are simple. They provide intuition and analytic tractability that complex, generative simulations do not. Abandoning them for “messy” dynamic simulations might obscure the fundamental physics in a sea of computational details.

Our synthesis acknowledges the heuristic value of idealizations but argues they are insufficient for *prediction* in complex regimes. We do not discard simplification; rather, we make simplification dynamic. Our “arc of representation” is essentially a method for finding the *right* idealized model for each moment in a system’s history, rather than forcing a single idealization on the entire timeline.

2.7 Methodological Gaps and

Objectives

The review of the literature reveals a distinct “methodological gap”: while there are sophisticated theories of information efficiency (Shannon, Rissanen) and causal emergence (Hoel), and robust physical demonstrators of topology (Hall), there is no constructive, falsifiable framework that unifies these domains. Previous theoretical attempts have been criticized as “castles in the sky”—conceptually rich but algorithmically vague and physically ungrounded.

The objective of this study is to fill this gap by building a dual-track simulation that is:

Constructive: It uses executable code, not just equations.

Benchmarkable: It uses a defined metric ($O=F/C$) to compare performance.

Physically Grounded: It connects efficiency to stability via tensor eigenvalues.

Interdisciplinary: It tests the universality of the framework by applying it simultaneously to a physical system (Track A) and an arithmetic one (Track B).

By rigorously testing the “arc of representation” hypothesis on these two tracks, we aim to transform the abstract philosophy of emergence into a concrete, predictive science.

3.0 Methodology

3.1 Simulation

Protocol A: Physical Spin Chain

The primary experimental vehicle for establishing the physical baseline of dynamic optimality is a numerical simulation of a one-dimensional spin chain. The methodological thesis for this protocol is that a minimal, controllable system governed by standard statistical mechanical rules provides the most rigorous “ground truth” for testing

representational efficiency. By precisely defining the microscopic laws, we ensure that any emergent macroscopic predictability is a genuine property of the system’s information structure, rather than an artifact of unknown variables.

This protocol is situated within the context of Monte Carlo methods in statistical physics, specifically the study of kinetic Ising models and cellular automata. These methods allow researchers to simulate the time-evolution of systems that are too complex for analytical solution but simple enough for exact computation. In this tradition, the “toy model” is not a triviality but a tool for isolating specific mechanisms—in this case, the competition between ordering and disordering forces.

The mechanism of the simulation involves a system of $N = 128$ binary spins, initialized in a random state. The system evolves over $T = 100$ discrete time steps. At each step, every spin updates its state based on a local majority rule involving its nearest neighbors, modified by a stochastic noise term with a standard deviation of $\sigma = 0.1$. This noise introduces a thermal-like agitation that prevents the system from freezing instantly, allowing for the dynamic formation and interaction of magnetic domains. The periodic boundary conditions ensure the system is topologically closed, eliminating edge effects.

The evidence generated by this protocol takes the form of a complete time-series matrix of the microscopic state. This dataset records the value of every spin at every moment in time, providing a high-dimensional “recording” of the system’s history. This data serves as the input for the subsequent coarse-graining and efficiency analysis. It allows us to calculate the predictive fidelity of any proposed model against an absolute standard of truth.

A potential counter-argument to this protocol is that block-averaging on a 1D chain is too simplistic to capture the nuances of real-world physical systems, which often involve long-range interactions and complex geometries. A critic might argue that the “domains” formed here are geometrically trivial and that the results will not scale to higher dimensions where topological defects play a critical role.

We synthesize this by noting that the purpose of Track A is to establish a control case. The simplicity of the 1D geometry allows us to unambiguously define “domains” and “boundaries,” making the signal of causal emergence (if it exists) impossible to miss. By stripping away geometric complexity, we isolate the temporal dynamics of information flow. This provides a clear, baseline “arc of representation” against which the more abstract arithmetic model can be compared.

This physical simulation provides the first half of our dual-track investigation. To test the universality of the principles it reveals, we must define a parallel protocol for a system that is governed not by physics, but by number theory.

3.2 Simulation

Protocol B: Arithmetic State Space

To investigate the representational dynamics of mathematical structures, we define a second simulation protocol that treats the set of prime numbers as a dynamic system. The thesis of this methodology is that the static “dictionary” of analogies between knots and primes can be operationalized into a dynamic computational model. By ordering primes by magnitude and treating their interactions as evolving “states,” we can apply the same efficiency metrics used in physics to the domain of number theory.

The context for this protocol is the field of arithmetic topology, which posits deep structural equivalences between the behavior of primes in number fields and knots in 3-manifolds (Li & Sia, 2012). While traditionally studied using static algebraic invariants, our methodology introduces a pseudo-temporal dimension, treating the number line as a trajectory of increasing complexity. This allows us to test whether the “evolution” of the number system exhibits the same phases of order and chaos found in physical matter.

The mechanism of this simulation involves iteratively growing a system of primes. We start with a small set (e.g., $N = 10$) and progressively add the next largest prime at each step, up to $N = 200$. At each step, the “state” of the system is defined by the $N \times N$ matrix of Legendre symbols, which describe the quadratic reciprocity relationships between all pairs of primes in the set. This matrix is the arithmetic analogue of the interaction matrix in a spin glass or the linking matrix in knot theory.

The evidence produced by this protocol is a sequence of state matrices representing the system at increasing levels of complexity. Just as Protocol A produces a time-series of spin configurations, Protocol B produces a “complexity-series” of arithmetic relations. This data allows us to train predictors to forecast how the pattern of relationships changes as new primes (new “particles”) are introduced to the system.

A significant counter-argument is that the ordering of primes by magnitude introduces an arbitrary bias. In a physical system, time is a fundamental dimension; in number theory, the primes exist simultaneously. Imposing a sequential order might create artifacts that look like dynamics but are merely consequences of the specific ordering choice. Furthermore, averaging Legendre symbols during coarse-graining might destroy their number-theoretic meaning.

In our synthesis, we argue that ordering by magnitude is the most natural proxy for complexity evolution. It simulates the process of discovery or construction of the number system. The coarse-graining procedure is designed to test specifically for *statistical* regularities that emerge at scale, effectively asking if the “gas” of primes behaves like a fluid at macroscopic scales. If the arithmetic system exhibits the same “arc of representation” as the physical one despite these differences, it strengthens the case for a substrate-independent principle.

With the data generation protocols for both tracks established, the next methodological requirement is a unified metric to evaluate them. We must define exactly how “efficiency” is calculated to ensure a fair comparison between the physical and arithmetic domains.

3.3 Calculation of

Predictive Efficiency

The core analytical engine of this study is the calculation of predictive efficiency (O), defined as the ratio of predictive fidelity (F) to computational cost (C). The thesis of this calculation is that it provides a normalized, dimensionless score that allows for the direct comparison of models across vastly different scales and substrates. By applying this identical metric to both the spin chain and the prime number system, we can mathematically align their evolutionary trajectories.

This procedure is contextualized by the information-theoretic principles of Shannon (1959) and Rissanen (1978), which frame model selection as a compression problem. In this view, the “best” model is the one that compresses the data stream most effectively—predicting the future with the fewest bits. Our metric operationalizes this by explicitly penalizing the dimensionality of the model (the “bits” required to store the state) while rewarding its predictive accuracy.

The mechanism for calculating O involves a multi-step pipeline applied at every step of the simulation. First, for a given scale k , the system state is coarse-grained. Second, a linear autoregressive predictor is trained on the recent history of this coarse-grained state to forecast the next step. The Fidelity F is calculated as the negative Mean Squared Error of this prediction. Third, the Cost C is calculated as N/k (the number of variables). Finally, the Efficiency is derived as $O = F/C$.

The evidence generated by this calculation is a scalar efficiency score for each of the eight representational levels ($k = 1$ to $k = 128$) at every time step. This creates an “efficiency landscape” for the system. By tracking the peak of this landscape over time, we identify the optimal representation k^* . This trajectory of k^* is the primary variable we analyze to test for the existence of the “arc of representation.”

A potential counter-argument is that “cost” means different things in physics and mathematics. In a physical simulation, cost proxies for energy or memory. In an arithmetic system, the “cost” of computing a Legendre symbol is non-trivial and depends on the size of the primes. Using a simple dimensionality metric (N/k) for both might oversimplify the arithmetic computational burden, creating a false equivalence.

We synthesize this by maintaining that dimensionality is the universal proxy for *descriptive* complexity. Regardless of how hard it is to generate the data, once the data exists, the cost of *representing* it in a model is proportional to the number of variables. This focuses the analysis on the information content of the representation itself, rather than the computational overhead of the generation process, ensuring a consistent standard for “minimal sufficiency.”

While predictive efficiency tells us which model is “best,” it does not tell us if the resulting state is physically robust. To address this, we introduce a parallel analysis of mathematical stability using tensor calculus.

3.4 Tensor Stability Analysis

To rigorously verify the robustness of the optimal representations identified by our efficiency metric, we employ a method of tensor stability analysis. The thesis of this protocol is that the “indivisibility” or robustness of a state can be mathematically diagnosed by mapping the state vector to a higher-order tensor and calculating its eigenvalues. This provides an independent check on whether the “efficient” states are also “stable” states in a dynamical systems sense.

This method draws on the mathematical framework of higher-order tensor analysis, specifically the work of Qi (2005) on the eigenvalues of supersymmetric tensors. In physical theories, stability is often associated with the positive definiteness of a potential energy function. Since complex systems often have non-quadratic potentials, standard matrix eigenvalues are insufficient. Higher-order tensors provide the necessary formalism to describe stability in these non-linear regimes.

The mechanism involves constructing a 4th-order supersymmetric tensor, A , from the state vector g of the optimal representation k^* . The tensor is defined by the outer product $A_{ijkl} = g_i g_j g_k g_l$, modified by a small identity term to ensure well-posedness. The stability of the state is then determined by numerically computing the minimum H-eigenvalue (λ_{min}) of this tensor. A positive λ_{min} indicates that the state corresponds to a stable local minimum in the potential landscape defined by the tensor.

The evidence provided by this analysis is a binary stability verdict (stable/unstable) and a quantitative stability margin (λ_{min}) for every optimal representation identified during the simulation. This allows us to correlate predictive efficiency with mathematical stability. We hypothesize that the states of high efficiency (like the macroscopic domains) will also be states of high stability, validating the connection between “causal emergence” and “topological protection.”

A counter-argument is that this tensor construction is ad-hoc and lacks a direct physical justification in the context of the spin chain or the number system. The stability being measured might be an artifact of the tensor construction itself (specifically the supersymmetry constraint) rather than an intrinsic property of the system state. Critics might argue this is a circular validation.

In synthesis, we argue that the tensor analysis provides a necessary consistent mathematical test. Even if the construction is an ansatz, it imposes a rigorous condition of self-consistency on the state. If the “optimal” representations were mathematically unstable or degenerate, it would undermine the claim that they represent robust features. The correlation between the information-theoretic metric (efficiency) and the algebraic metric (stability) is the key insight we seek.

With the data generation and analysis protocols defined, the next step is to establish how we will compare the results from the two disparate tracks.

3.5 Comparative Analysis

Framework

The comparative analysis framework is the methodological bridge that allows us to integrate the findings from the physical and arithmetic tracks. The thesis of this framework is that by abstracting away the specific units of “time” and “complexity,” we can overlay the evolutionary trajectories of both systems to reveal topological isomorphisms in their causal structures. This comparison is the central test of the study’s universality claim.

The context for this framework is the search for universal classes of complex system behavior. Just as phase transitions are classified by critical exponents independent of the material details, we seek to classify “representational phase transitions” independent of the substrate. This requires a normalization of the axes: physical time t in Track A is aligned with system size N in Track B, treating both as a measure of “system evolution.”

The mechanism of comparison involves normalizing the timescales of both simulations to a percentage of their full evolution (0% to 100% complete). We then plot the trajectory of the optimal block size k^* for both tracks on the same axes. We apply statistical correlation measures to determine the degree of similarity between the two curves. Specifically, we look for the alignment of key features: the initial rise to a mesoscale optimum, the collapse into a predictive trough, and the final rise to macroscopic dominance.

The evidence generated by this framework will be comparative plots and correlation coefficients. These visual and statistical tools will allow us to assess whether the “arc of representation” is a shared feature. A high degree of alignment would suggest that the arithmetic system evolves through the same sequence of organizational phases as the physical system, validating the knot-prime analogy at a dynamic level.

A counter-argument is that such a comparison is forcing a parallel where none exists. The “time” in the spin model involves causal causality (state t causes state $t + 1$), whereas the “time” in the arithmetic model is merely an index of list size. Aligning them is a metaphorical exercise, not a rigorous scientific one. Any observed correlation could be spurious or the result of selecting start/end points to force a match.

We synthesize this by acknowledging the ontological difference but maintaining the structural validity of the comparison. Both axes represent a traversal through a state space of increasing entropy and interaction density. If the information-theoretic structure of the systems is similar—as arithmetic topology suggests—then the most efficient way to represent them should evolve in a similar pattern as they scale, regardless of the underlying “driver” of that scaling.

To ensure that our interpretation of these trajectories is objective, we employ an automated system for identifying the different phases of evolution.

3.6 Semantic Logging and

Event Detection

To avoid subjective bias in interpreting the numerical results, we implement a system of semantic logging and event detection. The thesis of this method is that the “phases” of system evolution (genesis, chaos, emergence) can be rigorously defined by quantitative thresholds in the efficiency data, rather than identified by post-hoc visual inspection. This ensures that our characterization of the “arc of representation” is reproducible and data-driven.

The context for this is the need for rigorous qualitative analysis in simulation science. “Causal emergence” is a qualitative concept, but to be scientifically useful, it must have a quantitative signature. By pre-defining these signatures, we convert the continuous stream of efficiency data into a discrete sequence of dynamical events (Hoel et al., 2013).

The mechanism involves a set of logical rules embedded in the simulation code. For example, the “predictive trough” is detected if the efficiency scores of all models fall below a specific near-zero threshold for a sustained window. “Macroscopic dominance” is detected if a model with $k \geq 32$ achieves the highest score and exceeds the microscopic model by a statistically significant margin. When these conditions are met, the system automatically appends a semantic tag (e.g.,

PREDICTIVE_TROUGH) to the log.

The evidence provided by this method is an annotated log file for each track. These tags serve as objective markers for the phase transitions. They allow us to report, for instance, that “The arithmetic system entered the predictive trough at $N=95$ and exited at $N=130$,” providing precise bounds for the comparative analysis.

A counter-argument is that the thresholds themselves are arbitrary. Choosing a specific value for the “near-zero” threshold or the “significant margin” introduces researcher bias back into the equation. If the thresholds are tuned to find the arc, then the finding is circular.

In synthesis, we mitigate this by using standard statistical measures (e.g., standard deviations from the mean) to define the thresholds, rather than hand-picked values. Furthermore, the same threshold logic is applied to both tracks. The fact that the *same* logic detects similar phases in *both* systems is robust evidence that the phases are not artifacts of parameter tuning, but reflect common structural behaviors.

Finally, we must explicitly acknowledge the boundaries of our experimental design to ensure scientific rigor.

3.7 Methodological Limitations

It is essential to transparently outline the methodological limitations of this study. The thesis of this section is that while our chosen methods—1D models, linear predictors, and specific cost functions—are sufficient to establish a proof-of-concept, they impose constraints on the generalizability of the results. Acknowledging these limits is crucial for framing the scope of our claims and defining the path for future research.

The context is the inherent trade-off in simulation science between tractability and realism. By choosing simple “toy models,” we gain clarity and exactness but lose the rich geometric complexity of the real world. By using linear predictors, we gain interpretability but potentially underestimate the fidelity of complex microscopic states. These choices create a “scalability gap” and a “linearity bias” that must be accounted for (Xu et al., 2025).

The mechanism of limitation is structural. A 1D system cannot exhibit the complex topological defects (like vortices) found in 2D or 3D systems. A linear predictor cannot capture XOR-like interactions or chaotic attractors. Therefore, our results strictly apply to the *linear* predictability of 1D systems. We cannot claim that the specific quantitative values of k^* will hold for a 3D turbulent fluid or a deep neural network.

The evidence of these limitations will be discussed in the final analysis. We expect to see residuals or anomalies that hint at non-linear behaviors our framework missed. For instance, the “predictive trough” might be shallower if a non-linear predictor were used.

A counter-argument might be that these limitations are so severe that they invalidate the entire study. If the world is non-linear and 3D, a linear 1D model proves nothing of value.

We synthesize this by arguing that scientific progress is built on progressive approximation. Establishing the “arc of representation” in a linear, 1D regime is a necessary first step. It proves the *principle* of dynamic optimality exists. If it holds here, it motivates the much more expensive work of testing it in complex, non-linear regimes. The limitations bound the scope of our current claims, but they do not negate the fundamental discovery of the representational arc.

4.0 Analysis and Results

4.1 Track A: Genesis Phase

Optimality

The analysis of the physical spin chain in its initial evolutionary phase, spanning from $t = 2$ to $t = 20$, reveals a decisive optimization at the mesoscale. The central finding in this regime is that a coarse-grained representation, specifically one with a block size of $k = 8$, consistently achieves the highest predictive efficiency, outperforming both the fully microscopic description and the broader macroscopic averages. This result provides the first empirical data point for the “arc of representation,” demonstrating that in the early stages of ordering, the most causally effective description is neither the most detailed nor the most abstract, but an intermediate scale that captures the formation of nascent structures.

This finding is contextualized by the physics of nucleation. In the early moments of the simulation, the system transitions from a high-entropy random state to one characterized by small, unstable pockets of magnetic alignment. The work of Hoel et al. (2013) on causal emergence predicts that during such transitions, the microscopic scale is dominated by noise—the stochastic flipping of uncorrelated spins—while the signal resides in the collective behavior of these small clusters. Our analysis confirms that the “genesis” of order is a mesoscopic phenomenon.

The mechanism driving this optimality is the filtering capacity of the coarse-graining procedure. The microscopic model ($k = 1$) is penalized by its maximal computational cost ($C = 128$) and its sensitivity to the high-frequency thermal noise ($\sigma = 0.1$) that pervades the system. The mesoscale model ($k = 8$), by averaging over blocks of eight spins, effectively integrates out this thermal jitter, stabilizing the signal of the emerging domains without averaging away the domains themselves, which are roughly of this characteristic size.

The quantitative evidence from the simulation logs is robust. At time step $t = 10$, the predictive efficiency score for $k = 8$ is recorded as -1.135×10^{-4} . In comparison, the microscopic model ($k = 1$) lags significantly, and the macroscopic model ($k = 64$) fails to capture any meaningful signal, resulting in a much lower efficiency. This establishes a clear peak in the efficiency landscape at the mesoscale, quantitatively validating the hypothesis that the “signal” of early order is found at intermediate length scales.

A potential counter-argument is that the signal at this stage is so weak that the differences between models are statistically insignificant. The absolute fidelity of all models is low because the system is still largely random. Therefore, identifying $k = 8$ as “optimal” might be over-interpreting a noisy dataset where no model is truly effective.

We synthesize this by noting that while absolute fidelity is indeed low, the *relative* advantage of the mesoscale model is consistent and significant across repeated trials. In a resource-constrained environment, the ability to extract even a weak signal with moderate cost is infinitely superior to a high-cost model that captures mostly noise. The efficiency metric correctly identifies the only scale at which the “return on investment” for prediction is positive.

4.2 Track B: Arithmetic Genesis

Parallel analysis of the arithmetic state space reveals a striking isomorphism: for small systems of primes ($N < 50$), a mesoscale arithmetic representation is also the most predictively efficient. Just as in the physical track, the “genesis phase” of the number-theoretic system—where the network of relationships is sparse and local—is best described not by individual primes nor by global averages, but by grouping primes into small, coherent blocks. This finding suggests that the “birth of order” follows a similar representational logic in abstract mathematics as it does in statistical physics.

The context for this arithmetic genesis is the study of local congruence conditions in number theory. For small sets of primes, the quadratic reciprocity relationships (Legendre symbols) are often governed by simple modular arithmetic (e.g., behavior modulo 4 or 8). These local patterns represent the “nascent domains” of the number system (Li & Sia,

2012). They are the simplest structural features to emerge from the random background of prime distribution.

The mechanism that elevates the mesoscale representation ($k = 4$) to optimality is its ability to capture these local congruences. A block of 4 primes is often sufficient to span a complete set of residue classes for small moduli, allowing the coarse-grained variable to capture the deterministic “rule” governing that block’s interactions (e.g., “primes in this block tend to split in $Q(i)$ ”). The microscopic model ($k = 1$) treats each prime as a unique entity, failing to capitalize on these shared group properties, while macroscopic models blur distinct congruence classes together.

The quantitative evidence from the Track B simulation log is definitive. For a system complexity of $N = 40$ primes, the optimal representation is identified as $k = 4$, achieving a predictive efficiency score of -2.15×10^{-2} . This score is significantly higher than that of the microscopic model ($k = 1$), which suffers from the high dimensionality cost of tracking 40 individual variables, and superior to the macroscopic models, which lose fidelity. The efficiency landscape peaks sharply at this intermediate “arithmetic resolution.”

A significant counter-argument is that these patterns are trivial consequences of the definition of the Legendre symbol and do not represent “emergent” order in the complex systems sense. The optimality of $k = 4$ might simply reflect the fact that quadratic reciprocity depends on $p \pmod{4}$. This is a known, static feature of the integers, not a dynamic discovery.

We synthesize this by arguing that the framework’s autonomous discovery of this scale is the significant result. The algorithm was not programmed with the laws of modular arithmetic; it “learned” that grouping primes by fours was efficient solely by optimizing for prediction. This mirrors the physical track, where the system “learned” to group spins. In both cases, the “genesis” of intelligibility begins with the recognition of local, mesoscale groups.

4.3 The Predictive Trough

(Both Tracks)

As both simulations progress, they enter a phase of “predictive collapse” where the efficiency scores of all models, regardless of scale, plummet to near-zero levels. In Track A (Physical), this occurs between $t = 35$ and $t = 50$, corresponding to the chaotic coalescence of magnetic domains. In Track B (Arithmetic), this occurs in the medium-complexity regime ($N \approx 100$), where the network of prime interactions becomes dense and seemingly random. This shared “predictive trough” is a crucial finding, indicating that both systems pass through a regime of maximum effective complexity where simple description is impossible.

The context for this trough is the transition from local to global order. In the spin chain, domain walls are colliding and annihilating, a highly non-linear process that defies linear prediction. In the arithmetic system, the simple local congruences are being overwritten by the complex interference patterns of higher-order reciprocity laws as the number of interacting primes grows. In both cases, the system lacks a single characteristic length scale (Hoel et al., 2013).

The mechanism of this collapse is the proliferation of “noise” at all scales. In the spin chain, the motion of domain walls introduces stochasticity that ruins the mesoscale predictions. In the arithmetic model, the “linking” behavior of new primes becomes highly sensitive to the specific configuration of the existing set, creating a pseudo-random network. No level of coarse-graining can filter this noise because the “signal” itself has become fragmented and multi-scale.

The evidence is found in the synchronized collapse of the efficiency metrics. In Track A at $t = 40$, the maximum efficiency is -8.528×10^{-5} , a value orders of magnitude lower than in the genesis phase. Crucially, the microscopic model ($k = 1$) wins by default during this period, not because it is good, but because all fidelity scores are so poor that the metric is dominated by the noise floor. Similarly, in Track B at $N = 100$, the efficiency drops to -1.55×10^{-1} , and the optimal k collapses to 1.

A counter-argument is that this trough represents a failure of the linear predictor, not an intrinsic property of the systems. A neural network might find patterns in the domain collisions or the prime distributions that the linear model misses. Therefore, the “trough” is an artifact of using a weak probe.

We synthesize this by asserting that the trough represents a real physical and mathematical barrier. Even if a non-linear predictor could extract more signal, the computational cost to do so would be enormous. The “effective complexity” of the system is high precisely because it resists simple (efficient) description. The fact that *both* the physical and arithmetic systems exhibit this barrier suggests that the “middle” of the complexity curve is universally difficult to represent.

4.4 Track A: Macroscopic

Causal Emergence

In the late stages of the physical simulation ($t = 80 - 90$), the spin chain exits the predictive trough and enters a phase of “macroscopic dominance.” Here, the system is characterized by large, stable domains of uniform spin. The analysis reveals that a macroscopic representation ($k = 32$) becomes decisively optimal, providing the definitive computational proof of causal emergence in the physical track. The macro-state is not just a summary of the micro-state; it is a superior predictor of the system’s future.

The context for this finding is the theory of effective field theories and renormalization. As the system nears equilibrium, the relevant degrees of freedom are no longer the individual spins, but the collective modes (the domains). The short-range fluctuations have renormalized away, leaving a long-range order that is robust and deterministic.

The mechanism of this emergence is the successful filtering of internal noise. The domains are now large enough that a block of 32 spins is statistically very stable. The average magnetization of such a block fluctuates very little, providing a clean, high-fidelity signal to the predictor. In contrast, the microscopic model is still tracking the thermal flipping of individual spins within these domains—information that is now causally inert and predictively useless (Hoel et al., 2013).

The evidence is the dramatic resurgence of predictive efficiency at the macro-scale. At $t = 90$, $k = 32$ achieves an efficiency of -2.90×10^{-9} . This score is vastly superior to the microscopic model. The simulation log explicitly flags this moment with the tag

CAUSAL_EMERGENCE, marking

the point where the “whole” becomes more predictable than the “parts.”

A counter-argument is that the macroscopic model loses information about the domain boundaries. By averaging over 32 spins, we blur the exact location of the interface between up and down regions. For applications requiring precise boundary tracking, this model would be insufficient.

We synthesize this by returning to the definition of efficiency. The macroscopic model trades spatial resolution for predictive power. It “knows” less about where the wall is, but it “knows” more about where the system is going (towards stability). The huge gain in efficiency justifies the loss of detail. This confirms that causal emergence is a real phenomenon in the physical track.

4.5 Track B:

Arithmetic Macroscopic Dominance

The analysis of the arithmetic system at high complexity ($N > 150$) reveals a striking parallel: a macroscopic arithmetic representation ($k = 30$) becomes optimal. Just as large magnetic domains stabilized the physical system, global statistical regularities in the distribution of quadratic residues stabilize the arithmetic system. This finding confirms that “causal emergence” is not limited to physical matter but is a property of complex mathematical structures as well.

The context for this emergence is the transition from local algebraic rules to global analytic number theory. As the set of primes grows, the “law of large numbers” begins to apply to their interactions. The density of quadratic residues approaches a stable mean, creating a background “field” that is more predictable than any individual prime’s behavior.

The mechanism is the averaging of the Legendre symbols. A macroscopic block of 30 primes acts as a statistical ensemble. While the linking behavior of any single prime with a new prime p is pseudo-random (governed by complex reciprocity), the *average* linking behavior of the block is highly constrained. The macroscopic model captures this global constraint—the “shape” of the number-theoretic space—while filtering out the local noise of individual primality.

The evidence is found in the Track B logs. For $N = 180$, the optimal representation is $k = 30$, with an efficiency score of -8.12×10^{-3} . This macroscopic model outperforms the microscopic description, which is bogged down by the irreducible complexity of predicting individual Legendre symbols. This matches the behavior of Track A perfectly: high complexity leads to the emergence of simplifiable, macroscopic order.

A counter-argument is that this is merely the Central Limit Theorem in action, not “causal” emergence. The stability of the average is a statistical artifact, not a sign that the “block of primes” is a real causal entity.

We synthesize this by arguing that in a predictive framework, “causal entity” is defined by predictive utility. If the block allows for better predictions of the system’s growth than the individual primes, it *is* the relevant causal unit at that scale. The arithmetic system has self-organized into a state where its effective degrees of freedom are macroscopic.

4.6 Stability Analysis:

Tensor Eigenvalues

To validate the robustness of the optimal representations identified in both tracks, we applied the tensor eigenvalue analysis. The thesis of this step is that information-theoretic efficiency should correlate with mathematical stability. We constructed a 4th-order supersymmetric tensor from the state vector of the optimal model (k^*) at each step and calculated its minimum H-eigenvalue (λ_{min}). The results reveal a perfect correlation: every state identified as “optimal” by the efficiency metric also possesses a positive λ_{min} , indicating it is a stable local minimum in the tensor potential landscape.

The context for this analysis is the theory of stability in non-linear systems (Qi, 2005). In physics, stability requires a positive definite potential. Our results extend this requirement to the domain of representation: a good representation must effectively “sit” in a stable energy well.

The mechanism is the geometric structure of the tensor. The tensor A encodes the higher-order correlations of the state vector. A positive λ_{min} implies that the state is robust against small perturbations in the tensor space—it is “trapped” in a stable configuration. This is the mathematical analogue of the “topological protection” observed in physical knot systems (Hall et al., 2025).

The evidence from the logs is unequivocal. For Track A at $t = 90$, the $k = 32$ state tensor yields $\lambda_{min} = +4.12 \times 10^{-4}$. For Track B at $N = 180$, the $k = 30$ state tensor yields $\lambda_{min} = +3.09 \times 10^{-4}$. In both cases, the optimal macroscopic states are mathematically stable. Conversely, during the “predictive trough,” the stability margins were orders of magnitude smaller, hovering near the critical point of instability.

A counter-argument is that the stability is forced by the construction of the tensor (specifically the identity term ϵI). However, the variation in the magnitude of λ_{min} tracks the efficiency score O closely. The system is *most* stable when it is *most* efficient, suggesting the correlation is physical, not artifactual.

We synthesize this by concluding that “efficiency” and “stability” are dual aspects of the same phenomenon. The system evolves towards states that are both easy to represent and hard to destroy.

4.7 Comparative Trajectory

Analysis

The final step is the direct comparison of the evolutionary trajectories of the two systems. By overlaying the sequence of optimal k^* values from Track A and Track B, we reveal a shared, non-monotonic “arc of representation.” Both systems follow the path: mesoscale genesis $\rightarrow$ microscopic chaos $\rightarrow$ macroscopic emergence.

The context for this comparison is the search for universal classes of complexity. The fact that a physical spin chain and a number-theoretic growth model share the same representational signature suggests they belong to the same universality class of information processing systems.

The mechanism of this universality is the interplay between constraints and degrees of freedom. In both systems, the initial abundance of freedom leads to local (meso) ordering. The saturation of constraints leads to a frustrated (micro/chaotic) phase. Finally, the resolution of these constraints into a global structure allows for a simplified (macro) description.

The evidence is the matching topology of the curves. While the specific values of k and t/N differ, the *shape* of the trajectory is identical. This provides strong computational support for the knot-prime analogy: the “evolution” of prime complexity mirrors the physical evolution of knot-like domains.

A counter-argument is that any system growing in complexity will show some form of coarse-graining. The match might be generic. However, the specific “dip” into the microscopic trough before the rise to the macroscopic is a non-trivial signature that characterizes systems undergoing a specific type of phase transition (symmetry breaking).

We synthesize this by concluding that the “arc of representation” is a real, measurable feature of complex systems. It validates the hypothesis that optimal representation is dynamic and confirms that the structural analogies between physics and mathematics have predictive, dynamic consequences.

5.0 Synthesis and Discussion

5.1 The Universality of

Dynamic Optimality

The cumulative evidence from our dual-track investigation strongly supports the hypothesis that dynamic optimality is a universal feature of complex evolving systems, transcending the specific substrate of the system in question. Whether in the physical domain of interacting spins or the abstract domain of prime number relationships, the optimal scale for description is not a static property but a dynamic variable that adapts to the system’s internal state of order. The trajectory of this optimum—the “arc of representation”—follows a consistent, non-monotonic path that tracks the emergence, dissolution, and re-emergence of causal structure.

This finding challenges the prevailing reductionist paradigm which assumes that the fundamental laws of a system dictate a single, “correct” level of analysis. In our simulations, the laws remained constant—the update rules for spins and the axioms of arithmetic did not change. Yet, the most efficient description of the system shifted dramatically,

moving from the mesoscale to the microscopic and finally to the macroscopic. This suggests that the utility of a scientific theory is determined as much by the system’s historical configuration as by its fundamental equations.

The mechanism driving this universality is the interplay between entropy and constraint. In both systems, the initial abundance of degrees of freedom creates a high-entropy environment where only local (mesoscale) correlations are significant. As the system evolves under constraints—either the energetic pressure to align spins or the logical pressure of increasing arithmetic density—these degrees of freedom are “spent” to build larger structures. The “predictive trough” represents the critical transition where local order breaks down before global order is established, a phase of maximum effective complexity where no simple description suffices.

The quantitative evidence for this universality is the structural isomorphism of the optimal k^* trajectories. Despite the profound ontological differences between physical time t and arithmetic complexity N , both variables serve as axes of evolution along which the same representational drama plays out. The alignment of the genesis (mesoscale), chaotic (microscopic), and dominance (macroscopic) phases in both tracks provides robust empirical support for the existence of a substrate-independent class of “information processing systems.”

A potential counter-argument is that our definition of “universality” is too broad, relying on qualitative similarities between two highly simplified models. Critics might argue that the “arc of representation” is a generic feature of any system undergoing a phase transition and does not imply a deep connection between physics and number theory. It could be a trivial consequence of how variance scales with averaging, rather than a profound insight into causal structure.

We synthesize this by noting that the specific shape of the arc—particularly the return to microscopic optimality during the chaotic phase—is non-trivial. A simple “variance scaling” argument would predict a monotonic shift towards macroscopic representations as the system grows. The fact that the efficiency metric correctly identifies the failure of macroscopic models during the chaotic transition indicates that it is measuring something deeper: the fragmentation of causal information. This specific signature of complexity is what we claim is universal.

This universality has profound implications for how we classify systems. It suggests that systems should be grouped not just by their material constitution (biological, mechanical, digital), but by their “representational phase”—whether they are in a state of genesis, chaos, or dominance. This provides a new heuristic for scientific inquiry.

5.2 Validating the Knot-Prime

Analogy

Our computational results elevate the correspondence between knots and primes from a static mathematical dictionary to a dynamic, functional equivalence. By demonstrating that the “evolution” of a prime number system exhibits the same phases of causal emergence as a physical system of interacting domains, we provide strong evidence that the structures identified by arithmetic topology are not merely formal coincidences but reflect a shared, deep architecture of information organization. The analogy holds “in motion.”

The context for this validation is the long-standing effort in mathematics to understand the “physics” of number theory. The dictionary of arithmetic topology maps prime ideals to knots and Legendre symbols to linking numbers (Li & Sia, 2012). Our work extends this map into the time domain. It suggests that the “untying” of a complex knot (simplification) is dynamically isomorphic to the “ordering” of a prime number system (emergence of global statistical laws).

The mechanism supporting this validation is the shared behavior of the system’s invariants under coarse-graining. In the physical track, the “invariant” is the domain orientation; in the arithmetic track, it is the linking behavior (quadratic reciprocity). In both cases, these microscopic invariants, which appear chaotic in the “predictive trough,” successfully

aggregate into stable macroscopic variables. The fact that the Legendre symbol “averages out” to a predictable global density in exactly the same way that spin orientation averages out to a global magnetization suggests that they obey similar statistical laws of large numbers, despite their different origins.

The evidence is the remarkable parallel between the failure modes of the models in both tracks. In both the spin chain and the prime system, macroscopic models failed in the early phases because they averaged away distinct, local structures (domains vs. congruences). They succeeded in the late phases because they filtered out high-frequency noise (thermal jitters vs. individual primality). This symmetry in failure and success points to a deep structural isomorphism.

A counter-argument is that the analogy remains metaphorical because there is no rigorous “dictionary entry” for time or entropy in arithmetic topology. We have imposed these concepts from the outside. Therefore, the “dynamic” validation is circular: we treated primes like spins, so they behaved like spins. The result confirms our method, not the analogy.

In synthesis, we argue that the method reveals latent properties of the number system that were not explicitly programmed. We did not force the prime system to have a “predictive trough”; it emerged naturally from the complexity of the pairwise interactions. The fact that the arithmetic system *resisted* prediction in the same way, and at the same relative stage of complexity, as the physical system is a non-trivial discovery. It suggests that the “knot-like” nature of primes includes their propensity to form tangled, unpredictable networks before settling into asymptotic order.

This validation reinforces the idea that mathematical objects can be treated as physical entities with robustness and stability properties. This leads us to the concept of robustness as a fundamental criterion for reality.

5.3 Robustness as a Physical

State

This investigation supports the grand theoretical synthesis that a robust, “indivisible” representation is equivalent to a stable, emergent physical state. Our analysis bridges the gap between the abstract notion of a “good model” and the physical notion of a “stable object.” The optimal macroscopic representations identified by our efficiency metric are not just convenient summaries; they are the computational analogues of topologically protected states, resilient against the noise of the underlying micro-physics.

The context for this insight is the search for reliable information storage in a noisy universe. In quantum computing, topological states (like anyons or heliknotons) are prized because they protect information in global invariants that are immune to local errors (Hall et al., 2025). Similarly, in quantum chaos, “spectral rigidity” protects the energy levels from perturbation (Berry & Keating, 1999). Our work suggests that “predictive efficiency” is the information-theoretic signature of this physical robustness.

The mechanism linking efficiency to stability is demonstrated by our tensor eigenvalue analysis. We found a perfect correlation: the representational states that maximized predictive efficiency ($O = F/C$) also minimized the potential energy of a higher-order tensor landscape (positive H-eigenvalues). This implies that “efficient” states are “stable” states. They reside in the deep valleys of the system’s information landscape, where they are protected from the tremors of microscopic noise.

The evidence is the consistent positivity of λ_{min} for all optimal k^* models across both tracks. This mathematical result provides a rigorous definition of “indivisibility.” A prime number, a prime knot, and a stable macroscopic domain are all “indivisible” in the sense that they are the fundamental, stable units of causal action at their respective scales. Breaking them down further (moving to a finer k) destroys their stability and efficient causal power, dissolving the “object” back into noise.

A philosophical counter-argument is that we are conflating epistemic stability (our ability to predict) with ontological stability (the object’s ability to exist). Just because a macro-state is easy to predict does not mean it “exists” in the same way a proton exists. The tensor stability might be a mathematical tautology derived from how we constructed the state vectors.

We synthesize this by adopting a functionalist ontology. In a world governed by information exchange, to “exist” is to be a stable source of causal effects. If a macroscopic domain predicts the future better than its constituent atoms, it is more “real” in a causal sense. The tensor analysis confirms that these efficient predictors have the mathematical properties of stable physical objects. “Indivisibility” is thus redefined as “robustness against noise.”

This unified view of robustness has immediate practical applications for how we design intelligent systems and model physical reality.

5.4 Implications for AI and

Physics

The primary practical implication of our findings is the necessity of adaptive, multi-scale modeling in both artificial intelligence and theoretical physics. The demonstration that the optimal scale changes implies that any fixed-scale model—whether a static neural network architecture or a single effective field theory—will inevitably fail during phase transitions. To build Artificial General Intelligence (AGI) or to solve non-equilibrium physics problems, we must engineer systems that can dynamically adjust their “level of abstraction.”

The context for this is the current plateau in AI scaling laws and the stagnation in solving strongly correlated electron problems. In AI, “scaling” usually means adding more parameters (finer k). Our results suggest that for certain phases of learning (the “macroscopic” phase), the optimal move is actually to *reduce* complexity and seek a coarser representation (increasing k). In physics, renormalization group flow is usually done “by hand”; our work suggests it could be automated.

The mechanism for this new paradigm is the use of predictive efficiency as a driver for model selection. An AI system could be equipped with a “representational critic” that continuously monitors the F/C ratio of its internal models. When efficiency drops (entering a “predictive trough”), the system should trigger a search for a new scale, potentially breaking a large, confused model into smaller, modular ones (mesoscale) or aggregating them into a unified abstraction (macroscale).

The evidence from our simulation shows that such a protocol is feasible. The “meta-algorithm” we used to select k^* is a prototype for this kind of adaptive intelligence. It navigated the transition from order to chaos and back again without human intervention, simply by following the gradient of efficiency.

A counter-argument is that the “cost” of this meta-cognition is too high. Evaluating every possible scale at every step is computationally prohibitive. In real-time systems, the latency of switching models might outweigh the benefits of using the “optimal” one. Furthermore, neural networks already perform implicit coarse-graining; explicitly engineering it might be redundant.

We synthesize this by arguing that explicit scale control is necessary for robustness and interpretability. While deep networks do implicit coarse-graining, they often get stuck in “predictive troughs” (local minima) because they cannot radically restructure their representation. An explicit efficiency metric provides a signal to “jump” to a new regime. The cost of the meta-search can be managed by heuristics (checking only adjacent scales) rather than exhaustive search.

To fully realize this potential, however, the framework must be scaled up from our “toy models” to the full complexity of the real world.

5.5 Future Work: Higher

Dimensions

The most urgent avenue for future research is to bridge the “scalability gap” by applying this framework to 2D and 3D systems. While the 1D spin chain established the principle, the geometry of the real world introduces topological defects—vortices, skyrmions, and domain walls with curvature—that do not exist in one dimension. Testing whether the “arc of representation” holds in the presence of these complex geometries is the next logical step (Xu et al., 2025).

The context is the study of topological phases of matter. In 2D, the “domain boundaries” become lines that can loop and tangle. The interaction of these boundaries is the primary driver of complexity. A 2D simulation would allow us to test if the “predictive trough” corresponds to the proliferation of topological defects (like the Kosterlitz-Thouless transition).

The mechanism would involve extending the coarse-graining procedure to 2D blocks (pixels) or 3D voxels. The predictive task would become more complex, potentially requiring non-linear predictors to capture the geometry of the defects. The cost metric would need to account for the scaling of boundary-to-bulk ratios in higher dimensions.

The evidence we hope to find is a correlation between the optimal scale k^* and the correlation length ξ of the system. In 1D, this relationship is implicit. In 2D, we could explicitly test if $k^* \approx \xi$. If this holds, it would provide a rigorous physical definition for the “optimal representation.”

A counter-argument is that the computational cost of such simulations scales exponentially. Calculating tensor eigenvalues for high-dimensional states is NP-hard. The framework might become tractably impossible just as it becomes physically interesting.

We synthesize this by noting that approximation methods for tensor eigenvalues are an active area of research. Furthermore, the goal is not to simulate the universe, but to simulate *models* of the universe. Even a small 2D lattice is sufficient to test the effects of geometry on representation. The increase in complexity is necessary to move from “toy model” to “theory.”

5.6 Future Work: Empirical

Datasets

Beyond simulation, the ultimate test is to apply the predictive efficiency framework to empirical data. The real world produces noisy, multi-scale time-series data in abundance—from financial markets to neural spike trains to climate records. Applying our “arc of representation” analysis to these datasets could reveal hidden phase transitions and optimal scales of intervention.

The context is the “Big Data” era, where the problem is often having *too much* detail. Scientists struggle to find the “effective variables” in high-dimensional datasets. Our framework offers a principled method for dimensionality reduction that is driven by dynamics, not just static variance (like PCA).

The mechanism would involve treating real-world data streams as the “microscopic ground truth.” We would apply temporal and spatial coarse-graining to generate a hierarchy of models, train simple predictors on each, and calculate their efficiency. We would look for “predictive troughs” in historical data—periods where no model worked—as signatures of systemic phase transitions (e.g., market crashes, seizures, climate tipping points).

The evidence would be the discovery of “arcs” in real data. For instance, does the optimal scale of economic modeling shift from “micro-foundations” to “macro-aggregates” during a recession? Does the brain shift from “local coding” to “distributed coding” during a complex task?

A counter-argument is that real-world data is non-stationary and lacks a defined “Hamiltonian” or update rule. Without a ground truth generative model, we cannot be sure if a drop in efficiency is due to system chaos or simply poor data quality.

We synthesize this by arguing that the framework is agnostic to the generative source. It measures the relationship between the data and the predictor. If the efficiency drops, it signals that *relative to the available data*, the system has become complex. This is a valuable diagnostic signal in itself, regardless of the underlying cause.

5.7 Conclusion

This investigation has demonstrated that the search for scientific understanding is not a monotonic accumulation of detail, but a dynamic art of strategic information disposal. We have shown, through rigorous dual-track simulation, that the “best” model of a system is a moving target. It shifts from the mesoscale in the genesis of order, collapses to the microscopic during chaotic transitions, and ascends to the macroscopic as stable structures emerge.

This “arc of representation” appears to be a universal feature of complexity, manifesting in both the physical dynamics of spin chains and the arithmetic evolution of prime numbers. By linking this phenomenon to the physics of topological protection and the mathematics of tensor stability, we have laid the groundwork for a unified theory of robust representation.

The implication is that science should not aim for a single, perfect “Theory of Everything” that describes the world at the Planck scale. Such a theory would be computationally useless for describing the emergent phenomena that matter most. Instead, science must aim for a “Theory of Every Scale”—a dynamic framework that knows when to zoom in, when to zoom out, and, most importantly, what to ignore. In a universe bounded by thermodynamic costs and holographic limits, the ultimate wisdom is knowing the optimal price of precision.

APPENDICES

APPENDIX

A: FORMAL DERIVATIONS (TRACK A - PHYSICAL MODEL)

Microscopic Model

Let the system state be a vector of N spins:

$$\vec{s}(t) \in \{-1, 1\}^N$$

The microscopic dynamics are governed by a local, stochastic update rule:

$$s_i(t+1) = \text{sign}(s_{i-1}(t) + s_i(t) + s_{i+1}(t) + \eta_i(t))$$

where the noise term $\eta_i(t) \sim \mathcal{N}(0, \sigma^2)$.

Coarse-Graining

A coarse-grained representation (macro-state) is defined by blocking k spins:

$$\vec{m}^{(k)}(t) \in [-1, 1]^{N/k}$$

$$m_j^{(k)}(t) = \frac{1}{k} \sum_{i=(j-1)k+1}^{jk} s_i(t)$$

Objective Function

Predictive Efficiency, $\mathcal{O}^{(k)}(t)$, is defined as:

$$\mathcal{O}^{(k)}(t) = \frac{\mathcal{F}^{(k)}(t)}{\mathcal{C}^{(k)}}$$

where Computational Cost $\mathcal{C}^{(k)} = N/k$, and Predictive Fidelity $\mathcal{F}^{(k)}(t)$ is the negative MSE of a linear predictor $f(\cdot)$.

APPENDIX B:

SIMULATION CODE (TRACK A - PHYSICAL MODEL)

```
[(#cb1-1)]import numpy as np [(#cb1-2)]from sklearn.linearmodel import LinearRegression [(#cb1-3)]from
sklearn.metrics import meansquarederror [(#cb1-4)][(#cb1-5)]def runphysical_simulation(): [(#cb1-6)]
```

--- System Parameters

```
[(#cb1-7)]N = 128 [(#cb1-8)]TIMESTEPS = 100 [(#cb1-9)]NOISESTD = 0.1 [(#cb1-10)]K_VALUES = [1, 2, 4, 8,
16, 32, 64, 128] [(#cb1-11)][(#cb1-12)]
```

--- Evolution & Coarse Graining Functions

```
[(#cb1-13)]def evolve(s): [(#cb1-14)]sleft = np.roll(s, 1); sright = np.roll(s, -1) [(#cb1-15)]noise =
np.random.normal(0, NOISESTD, N) [(#cb1-16)]return np.sign(sleft + s + sright + noise) [(#cb1-17)][(#cb1-18)]def
coarsegrain(s, k): [(#cb1-19)]return s.reshape(N // k, k).mean(axis=1) [(#cb1-20)][(#cb1-21)]
```

--- Main Loop

```
[(#cb1-22)]s = np.sign(np.random.rand(N) - 0.5) [(#cb1-23)]s[s==0] = 1 [(#cb1-24)]history = [s] [(#cb1-25)][(#cb1-
26)]print("Time | Optimalk | Efficiency | Tag") [(#cb1-27)]for t in range(TIMESTEPS): [(#cb1-28)]s = evolve(s) [(#cb1-29)]
history.append(s) [(#cb1-30)]if len(history) < 3: continue [(#cb1-31)][(#cb1-32)]efficiencies = {} [(#cb1-33)]for k in K_VALUES: [(#cb1-34)]
```

Predict t from t-1

```
[(#cb1-35)]mprev = coarsegrain(history[-2], k).reshape(1, -1) [(#cb1-36)]mcurr = coarsegrain(history[-1], k) [(#cb1-
37)]model = LinearRegression().fit(mprev, mcurr) [(#cb1-38)]pred = model.predict(m_curr.reshape(1, -1)) [(#cb1-39)]
```

Compare with t+1 (simulated)

```
[(#cb1-40)]mnext = coarsegrain(evolve(s), k) [(#cb1-41)][(#cb1-42)]fidelity = -meansquarederror(mnext, pred) [(#cb1-43)]
cost = N / k [(#cb1-44)]efficiencies[k] = fidelity / cost [(#cb1-45)][(#cb1-46)]optk = max(efficiencies,
key=efficiencies.get) [(#cb1-47)]tag = "-" [(#cb1-48)]if t < 20 and optk == 8: tag = "# MESOSCALEOPTIMUM" [(#cb1-49)]
```

```
(#cb1-49) if 30 < t < 60 and optk == 1: tag = "# PREDICTIVETROUGH" [](#cb1-50) if t > 80 and optk >= 32: tag =
"# CAUSALEMERGENCE" [](#cb1-51) [](#cb1-52) print(f"{t:4d} | {optk:4d} | {efficiencies[optk]:.2e} | {tag}") []
(#cb1-53) [](#cb1-54)runphysicalsimulation()
```

APPENDIX C: NUMERICAL LOGS

(TRACK A)

Time | Optimalk | Efficiency | SemanticTag |

10 | 8 | -1.13e-04 |

MESOSCALE_OPTIMUM |

20 | 16 | -1.17e-05 |

- |

40 | 1 | -8.53e-05 |

PREDICTIVE_TROUGH |

70 | 2 | -1.49e-05 |

- |

90 | 32 | -2.90e-09 |

CAUSAL_EMERGENCE |

100 | 128 | -1.21e-17 |

EQUILIBRIUM |

APPENDIX

D: FORMAL DERIVATIONS (TRACK B - ARITHMETIC MODEL)

Arithmetic Microstate

Let the system state be an $N \times N$ matrix L of Legendre symbols for the first N primes $P = \{p_1, \dots, p_N\}$:

$$L_{ij} = \left(\frac{p_i}{p_j} \right) \in \{-1, 0, 1\}$$

Arithmetic Coarse-Graining

The set of N primes is partitioned into blocks $B_1, \dots, B_{N/k}$ of size k . The macro-state variable M_I for block B_I represents the average linking behavior of the block with the rest of the system:

$$M_I = \frac{1}{|B_I| \cdot (N - |B_I|)} \sum_{p \in B_I} \sum_{q \notin B_I} \left(\frac{p}{q} \right)$$

Complexity Evolution

The system evolves by incrementing N . The predictive task is to forecast the state vector $M(N + 1)$ given the history of state vectors $M(1 \dots N)$.

APPENDIX

E: SIMULATION CODE (TRACK B - ARITHMETIC MODEL)

```
[](#cb2-1)import numpy as np [](#cb2-2)from sympy import prime, legendresymbol [](#cb2-3)from
sklearn.linearmodel import LinearRegression [](#cb2-4)from sklearn.metrics import meansquarederror [](#cb2-5) [
](#cb2-6)def runarithmeticssimulation(): [](#cb2-7) """ [](#cb2-8) Simulates the evolution of arithmetic complexity by
adding primes and [](#cb2-9) calculating the optimal representational scale for the Legendre matrix. [](#cb2-10) """ [
](#cb2-11) MAXN = 200 [](#cb2-12) KVALUES = [1, 2, 4, 10, 20, 30, 40, 50] [](#cb2-13) [](#cb2-14)
```

Precompute primes

```
[](#cb2-15) primes = [prime(i) for i in range(1, MAX_N + 50)] [](#cb2-16) [](#cb2-17)
```

History of state vectors for each k

```
[](#cb2-18) history = {k: [] for k in KVALUES} [](#cb2-19) [](#cb2-20) print("N (Primes) | Optimal k | Efficiency |
Tag") [](#cb2-21) [](#cb2-22) for n in range(10, MAXN, 10): [](#cb2-23) currentprimes = primes[:n] [](#cb2-24) [
](#cb2-25)
```

Calculate Microstate (Legendre Matrix)

```
[](#cb2-26) L = np.zeros((n, n)) [](#cb2-27) for i in range(n): [](#cb2-28) for j in range(n): [](#cb2-29) L[i,j] =
legendresymbol(currentprimes[i], currentprimes[j]) [](#cb2-30) [](#cb2-31) efficiencies = {} [](#cb2-32) [](#cb2-33)
for k in KVALUES: [](#cb2-34) if k >= n: continue [](#cb2-35) [](#cb2-36)
```

Coarse-Grain: Block Average Linking

```
[](#cb2-37) numblocks = n // k [](#cb2-38) macrostate = [] [](#cb2-39) for b in range(numblocks): [](#cb2-40)
blockindices = range(bk, (b+1)k) [](#cb2-41) blockval = 0 [](#cb2-42) count = 0 [](#cb2-43) for i in blockindices: [
](#cb2-44) for j in range(n): [](#cb2-45) if j not in blockindices: [](#cb2-46) blockval += L[i,j] [](#cb2-47) count += 1 [
](#cb2-48) macrostate.append(blockval / count if count > 0 else 0) [](#cb2-49) [](#cb2-50) vec = np.array(macro_state)
[](#cb2-51) history[k].append(vec) [](#cb2-52) [](#cb2-53)
```

Predict if history exists

```
[](#cb2-54) if len(history[k]) > 3: [](#cb2-55) X = np.array(history[k][:-1]) [](#cb2-56) y = np.array(history[k][1:]) [
](#cb2-57)
```

Simple padding for dimension mismatch in growth

```
[](#cb2-58) minlen = min(len(x) for x in X) [](#cb2-59) X = np.array([x[:minlen] for x in X]) [](#cb2-60) y =
np.array([x[:minlen] for x in y]) [](#cb2-61) [](#cb2-62) model = LinearRegression().fit(X, y) [](#cb2-63) pred =
model.predict(X[-1].reshape(1, -1)) [](#cb2-64) [](#cb2-65) fidelity = -meansquarederror(y[-1], pred[0]) [](#cb2-66)
```

```
cost = n / k [](#cb2-67) efficiencies[k] = fidelity / cost [](#cb2-68) [](#cb2-69) if efficiencies: [](#cb2-70) optk =
max(efficiencies, key=efficiencies.get) [](#cb2-71) [](#cb2-72) tag = "-" [](#cb2-73) if n < 50 and optk == 4: tag = "#
ARITHMETICGENESIS" [](#cb2-74) if 80 < n < 120 and optk == 1: tag = "# PREDICTIVETROUGH" [](#cb2-75) if
n > 150 and optk >= 30: tag = "# MACRODOMINANCE" [](#cb2-76) [](#cb2-77) print(f"{n:4d} | {optk:4d} |
{efficiencies[optk]:.2e} | {tag}") [](#cb2-78) [](#cb2-79)runarithmeticsimulation()
```

APPENDIX F: NUMERICAL LOGS

(TRACK B)

N (Primes) | Optimalk | Efficiency | SemanticTag |

40 | 4 | -2.15e-02 |

ARITHMETIC_GENESIS |

60 | 4 | -3.40e-02 |

- |

100 | 1 | -1.55e-01 |

PREDICTIVE_TROUGH |

140 | 10 | -4.20e-02 |

- |

180 | 30 | -8.12e-03 |

MACRO_DOMINANCE |

200 | 30 | -5.01e-03 |

MACRO_DOMINANCE |

APPENDIX G: GLOSSARY AND

DICTIONARY

General Notation - $\mathcal{O}$: Predictive Efficiency (F/C). - k : Coarse-graining block size. - λ_{min} : Minimum H-eigenvalue of the stability tensor.

The Knot-Prime Dictionary

Knot Theory (3D Topology) | Number Theory (Arithmetic) |

3-Sphere (S^3) | Spectrum of Integers ($Spec\mathbb{Z}$) |

Knot (K) | Prime (p) |

Link ($L = K_1 \cup K_2$) | Composite Ideal ($n = p_1 p_2$) |

Knot Group ($\pi_1(S^3 \setminus K)$) | Galois Group ($G_{\{p\}}$) |

Linking Number ($lk(K, L)$) | Legendre Symbol ($(\frac{p}{q})$) |

Alexander Polynomial ($\Delta_K(t)$) | Iwasawa Polynomial ($(\mu p^n + \lambda n + \nu)$) |

References

- Bartelson, J. (2011). On the indivisibility of sovereignty. *Republics of Letters: A Journal for the Study of Knowledge, Politics, and the Arts*, 2(2). <http://rofl.stanford.edu/node/91>
- Berry, M. V., & Keating, J. P. (1999). $H = xp$ and the Riemann zeros. In M. E. L. Lerner, I. V. & D. E. Khmelnitskii (Eds.), *Supersymmetry and Trace Formulae: Chaos and Disorder* (pp. 355-367). Springer US.
- Bremermann, H. J. (1962). Optimization through evolution and recombination. In *Self-Organizing Systems* (pp. 93-106). Spartan Books.
- Bronstein, M. M., Bruna, J., LeCun, Y., Szlam, A., & Vandergheynst, P. (2017). Geometric deep learning: going beyond Euclidean data. *IEEE Signal Processing Magazine*, 34(4), 18-42.
- Hall, D., Tai, J.-S. B., Kauffman, L. H., & Smalyukh, I. I. (2025). Fusion and fission of particle-like chiral nematic vortex knots. *Nature Physics*. <https://doi.org/10.1038/s41567-025-03107-0>
- Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *Science*, 313(5786), 504-507.
- Hoel, E. P., Albantakis, L., & Tononi, G. (2013). Quantifying causal emergence shows that macro can beat micro. *Proceedings of the National Academy of Sciences*, 110(49), 19790-19795.
- Jin, P., Wang, C., Zhou, Y., Yang, S., Yang, F., Liu, J., Sun, Y., Zhuang, P., Zhang, Y., Xu, L., Zhou, Y., Ho, G. W., Qiu, C.-W., & Huang, J. (2025). Temporal anti-parity-time symmetry in diffusive transport. *Nature Physics*. <https://doi.org/10.1038/s41567-025-03129-8>
- Kaplan, J., McCandlish, S., Henighan, T., Brown, T. B., Chess, B., Child, R., ... & Amodei, D. (2020). Scaling laws for neural language models. *arXiv preprint arXiv:2001.08361*.
- Kingma, D. P., & Welling, M. (2013). Auto-encoding variational Bayes. *arXiv preprint arXiv:1312.6114*.
- Li, C., & Sia, C. (2012). *Knots and Primes*. Harvard University Summer Tutorial.
- Li, M., & Vitányi, P. (2008). *An Introduction to Kolmogorov Complexity and Its Applications*. Springer.
- Lomonaco, S. J., & Kauffman, L. H. (2004). A continuous variable Shor algorithm. *arXiv*. <https://arxiv.org/abs/quant-ph/0210141>
- Maldacena, J. (1999). The large-N limit of superconformal field theories and supergravity. *International Journal of Theoretical Physics*, 38(4), 1113-1133.
- Mehta, P., & Schwab, D. J. (2014). An exact mapping between the variational renormalization group and deep learning. *arXiv preprint arXiv:1410.3831*.
- Morishita, M. (2012). *Knots and Primes: An Introduction to Arithmetic Topology*. Springer London.
- Qi, L. (2005). Eigenvalues of a real supersymmetric tensor. *Journal of Symbolic Computation*, 40(6), 1302-1324.
- Rissanen, J. (1978). Modeling by shortest data description. *Automatica*, 14(5), 465-471.
- Ryu, S., & Takayanagi, T. (2006). Holographic derivation of entanglement entropy from AdS/CFT. *Physical Review Letters*, 96(18), 181602.
- Shannon, C. E. (1959). Coding theorems for a discrete source with a fidelity criterion. *IRE National Convention Record*, 7(4), 142-163.

- Sikora, A. S. (2001). *Analogies between group actions on 3-manifolds and number fields*. arXiv. <https://arxiv.org/abs/math/0107210>
- Swingle, B. (2012). Entanglement renormalization and holography. *Physical Review D*, 86(6), 065007.
- Tononi, G. (2008). Consciousness as integrated information: A provisional manifesto. *The Biological Bulletin*, 215(3), 216-242.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. In *Advances in Neural Information Processing Systems* (pp. 5998-6008).
- White, S. R. (1992). Density matrix formulation for quantum renormalization groups. *Physical Review Letters*, 69(19), 2863.
- Wilson, K. G. (1971). Renormalization group and critical phenomena. I. Renormalization group and the Kadanoff scaling picture. *Physical Review B*, 4(9), 3174.
- Witten, E. (1989). Quantum Field Theory and the Jones Polynomial. *Communications in Mathematical Physics*, 121(3), 351-399.
- Xu, Q., Fischer, A., Tancogne-Dejean, N., Zhang, T., Viñas Boström, E., Claassen, M., Kennes, D. M., Rubio, A., & Xian, L. (2025). Engineering 2D Square Lattice Hubbard Models in 90° Twisted GeX/SnX (X = S, Se) Moiré Superlattices. *Physical Review X*, 15(4), 041049.