Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 3, Issue 11, 2025

FAST MULTI-SCALE GENERALIZATION BOUNDS FOR TOOL-AUGMENTED REASONING IN LARGE LANGUAGE MODELS UNDER HEAVY-TAILED LOSS REGIMES

Muhammad Baig[1], Muhammad Aakash Imtiaz[2], Hassan Revel[3], Aleena Jamil[4], Shafiq Hussain[*5], Adeen Amjad[6], Waqar Ahmad[7], Arslan Ali Mansab[8], Muhammad Hamza Akbar[9]

[*5]*Department of Computer Science, University of Sahiwal, Sahiwal 57000, Pakistan*
[1,2,4,6,7,8,9]*Department of Computer Science, University of Sahiwal, Sahiwal 57000, Pakistan*
[3]*Department of Computer Science, Virtual University of Pakistan, Lahore 54000, Pakistan*

[*1]muhammed.beig@gmail.com, [2]akashimtiaz123@gmail.com, [3]hassanrevelai@gmail.com, [4]aleena.jamil_vf@uosahiwal.edu.pk, [5]drshafiq@uosahiwal.edu.pk,[6]adeen.amjad@uosahiwal.edu.pk, [7]waqarahmad@uosahiwal.edu.pk, [8]arslansli@uosahiwal.edu.pk, [9]hamzaakbar@uosahiwal.edu.pk

DOI: https://doi.org/10.5281/zenodo.17811952

Keywords
Index Terms– Heavy-tailed learning, multi-scale generalization bounds, Bernstein condition, Tool-augmented reasoning, Large Language Models (LLMs), Statistical learning theory, Robust estimation.

Article History
Received: 01 September 2025
Accepted: 16 November 2025
Published: 29 November 2025

Corresponding Author: *
Shafiq Hussain

Abstract
The rise of tool-augmented language models (TaLMs) presents a significant challenge for generalization theory, as their error distributions are often heavy-tailed and unbounded, rendering conventional theoretical analyses ineffective. This work establishes fast learning rates for tool-augmented reasoning, which we model as a multi-step process. To control the model's excess risk, we introduce two key structural conditions. First, we assume that the worst-case loss across the hypothesis class possesses a finite moment of order greater than two, ensuring control over extreme deviations. Second, we posit a Multi-Scale Bernstein Condition that links the variance of the error to its expectation across different levels of semantic complexity, characterized by a stability parameter. By leveraging advanced methods from the theory of unbounded empirical processes, we prove that the excess risk converges at a rate that interpolates between the slow and fast classical rates. This rate improves as the loss tails become lighter and the reasoning process becomes more stable. Furthermore, we derive a complexity-aware bound where the required sample size scales favorably with the depth of the reasoning chain, providing a foundational framework for verifying the reliability of neuro-symbolic AI agents.

INTRODUCTION

THIS ascendancy of Large Language Models (LLMs) has been marked by a transition from passive text generation to active, agentic reasoning. Systems such as Toolformer [1], ToolLLaMA [2], and Agent Q[15] augment the probabilistic backbone of transformers with deterministic execution environments calculators, Python interpreters, and SQL databases enabling the synthesis of answers that require precise computation and retrieval. This hybridization, often termed Neuro-Symbolic AI [3], promises to overcome the inherent limitations of connectionist models in handling logic, arithmetic, and factual consistency.

However, this capability comes at the cost of statistical stability [4]. In classical supervised learning, loss functions (e.g., cross-entropy, bounded square error) are typically well-behaved; extreme outliers are rare, and the law of large numbers ensures rapid convergence of the empirical risk to the true risk [5]. In tool-augmented reasoning, this assumption is fundamentally violated. The error distribution of a reasoning agent is often heavy-tailed. A single incorrect API call, a hallucinated variable in a Python script, or a logical fallacy in a Chain-of-Thought (CoT)[10] trace can result in a final output that is not merely "incorrect" but arbitrarily divergent from the solution space. For instance, in symbolic regression tasks or automated theorem proving, the "distance" between a generated proof and a valid one may not be bounded, and the computational cost of a looping agent can grow exponentially [6]. Empirical studies confirm that reasoning errors in LLMs follow power-law distributions rather than exponential decays. This phenomenon is exacerbated by Tool-Induced Myopia (TIM), where models substitute genuine reasoning with superficial tool usage, leading to high-confidence but logically unsound failures. Consequently, the standard generalization bounds based on the Hoeffding's or the Bernstein's inequalities, which assume sub-Gaussian tails, are vacuous in this domain. They fail to explain why some reasoning models generalize effectively after seeing relatively few examples (fast rates) while others suffer from catastrophic overfitting despite massive pre-training. This paper addresses the critical need for a new statistical learning theory tailored to unbounded, heavy-tailed reasoning processes. We reject the convenience of bounded losses and instead adopt the rigorous framework of heavy-tailed empirical process theory. Our approach centers on characterizing the "shape" of the loss landscape through the lens of the envelope function, the worst-case loss over the hypothesis class, and the multi-scale variance structure of the reasoning tasks. The ascendancy of Large Language Models (LLMs) has been marked by a transition from passive text generation to active, agentic reasoning. Systems such as Toolformer[1], ToolLLaMA[2], and Agent Q[15] augment the probabilistic backbone of transformers with the deterministic execution environments calculators, Python interpreters, and SQL databases enabling the synthesis of answers that require precise computation and retrieval. This hybridization, often termed Neuro-Symbolic AI, promises to overcome the inherent limitations of connectionist models in handling logic, arithmetic, and factual consistency. However, this capability comes at the cost of statistical stability. In classical supervised learning, loss functions (e.g., cross-entropy, bounded square error) are typically well-behaved; extreme outliers are rare, and the law of large numbers ensures rapid convergence of the empirical risk to the true risk. In tool-augmented reasoning, this assumption is fundamentally violated. The error distribution of a reasoning agent is often heavy-tailed. A single incorrect API call, a hallucinated variable in a Python script, or a logical fallacy in a Chain-of-Thought (CoT) trace can result in a final output that is not merely "incorrect" but arbitrarily divergent from the solution space. For instance, in symbolic regression tasks or automated theorem proving, the "distance" between a generated proof and a valid one may not be bounded, and the computational cost of a looping agent can grow exponentially. Empirical studies confirm that reasoning errors in LLMs follow power-law distributions rather than exponential decays. This phenomenon is exacerbated by Tool-Induced Myopia (TIM), where models substitute genuine reasoning with superficial tool usage, leading to high-confidence but logically unsound failures. Consequently, the standard generalization bounds based on Hoeffding's or Bernstein's inequalities, which assume sub-Gaussian tails, are vacuous in this domain. They fail to explain why some reasoning models generalize effectively after seeing relatively few examples (fast rates) while others suffer from catastrophic overfitting despite massive pre-training. This paper addresses the critical need for a new statistical learning theory tailored to unbounded, heavy-tailed reasoning processes. We reject the convenience of bounded losses and instead adopt the rigorous framework of heavy-tailed empirical process theory. Our approach centres on characterizing the "shape" of the loss landscape through the lens of the envelope function, the worst-case loss over the hypothesis class, and the multi-scale variance structure of the reasoning tasks. We consider the problem of learning a reasoning policy f from a class $\mathcal{F}$ to minimize an unbounded loss $l(f, z)$. The challenge is threefold:

Unboundedness: The loss $l(f, z)$ may take arbitrarily large values (e.g., execution time, edit distance), rendering the uniform convergence $\sup_f |P_n l(f) - Pl(f)| \to 0$ difficult to control without heavy moment penalties.

Composite Structure: The hypothesis f is a composition of neural operations and discrete tool invocations, creating a disjoint, non-convex optimization landscape.

Heavy Tails: The probability of observing a large loss decay slowly, $P(l(f, Z) > t) \sim t^{-\alpha}$, necessitating analysis that does not rely on exponential moment generating functions.

We provide a complete theoretical treatment of this problem:

Formal Framework: We define the hypothesis class of tool-augmented agents as a composite function space with distinct metric entropy properties imposed by the discrete nature of tools (Section 3).

New Regularity Conditions: We introduce the Tool-Augmented Envelope Condition (Assumption 2) and the Multi-Scale Bernstein Reasoning Condition (Assumption 3), which generalizes the classical Bernstein assumption to handle the heterogeneous variance profiles of different reasoning tools (e.g., a calculator has low variance; a retriever has high variance).

Fast Rate Theorems: We derive explicit generalization bounds (Theorem 2) showing that the excess risk decays as $O(n^{-\beta})$, where β depends on the tail index r and the Bernstein power γ. We show that $\beta \to 1$ as the reasoning process becomes more stable (higher γ) and the tails become lighter (higher r).

Complexity Analysis: We prove that the sample complexity is sensitive to the depth of the tool chain, but scales sub-quadratically under specific entropy conditions (Theorem 4).

This work provides the mathematical scaffolding to certify the reliability of autonomous agents, moving beyond empirical evaluation on static benchmarks to provable guarantees on generalization.

Related Work

The intersection of heavy-tailed statistics, robust optimization, and neuro-symbolic reasoning is a nascent but critical field. This section reviews pertinent literature, distinguishing our contributions from existing frameworks. Heavy-Tailed Learning Theory. The Classical learning theory relies on boundedness to utilize the McDiarmid's inequality. When losses are heavy-tailed, the empirical mean is a suboptimal estimator of the true risk. Brownlees et al. (2015) [3] and Hsu & Sabato (2016) [4] proposed robust mean estimators (e.g., Median-of-Means, Catoni's estimator) to achieve sub-Gaussian rates $O(n^{-1/2})$. However, these methods modify the training objective, which is often impractical for large-scale deep learning. Dinh et al. (2016) [5] pioneered the analysis of *standard* ERM under heavy-tailed losses by imposing conditions on the envelope function, a path we follow and extend. Our work differs by applying this specifically to the composite, multi-agent hypothesis spaces of the TaLMs. Concentration of the Unbounded Processes. The backbone of our analysis rests on concentration inequalities for the suprema of unbounded empirical processes. Lederer and van de Geer (2014) [6] derived bounds using Bernstein-Orlicz norms that depend only on the L^r moments of the envelope. We adapt these results, specifically their "peeling" strategy, to handle the hierarchical structure of tool-augmented hypotheses.

Table I Comparative Review of Approaches

Authors	Approach	Methodology
Dinh et al. (2016) [5]	Fast Learning Rates w/ Heavy Tails	ERM via Envelope Functions
Brownlees et al. (2015) [3]	ERM for Heavy-Tailed Losses	Catoni M-Estimation
Lederer & van de Geer (2014) [6]	Concentration for Suprema	Orlicz-Norm Chaining
Schick et al. (2023) [1]	Toolformer	Self-Supervised Tool Learning

Authors	Approach	Methodology
Hsu & Sabato (2016) [4]	Loss Minimization w/ Heavy Tails	Median-of-Means
Wei et al. (2022) [7]	Chain-of-Thought Prompting	Intermediate Reasoning Steps
Cortes et al. (2013) [8]	Relative Deviation Bounds	Relative-Risk Analysis
Qin et al. (2023) [9]	ToolLLaMA	DFSDT Tree-Search
Wang et al. (2025) [10]	CoT Information	Statistical CoT Theory
This Work (2025)	Tool-Augmented Fast Rates	Multi-Scale Bernstein Reasoning

The empirical success of Toolformer, Gorilla [14], and ToolLLaMA [1] has demonstrated the utility of external tools. However, theoretical analysis is largely absent. Schick et al. [1] rely on self-supervised learning without guarantees. Recent work on SymCode [12] and Agent Q[15] attempts to structure the reasoning space using symbolic verification and tree search (MCTS), effectively implicitly optimizing for the Bernstein conditions we formalize. Bayat et al. (2025) [11] identified "Tool-Induced Myopia," explicitly linking tool reliance to a shift in error distributions, motivating our heavy-tailed assumption.

Table II Comparative Review (Pros & Cons)

Authors	Math Assumptions	Advantages	Limitations
Dinh et al. (2016) [5]	L^r-envelope; Bernstein	Fast heavy-tail rates	Hard to verify condition
Brownlees et al. (2015) [3]	$\sigma^2 < \infty$	Robust; sub-Gaussian	Expensive M-estimation
Lederer & van de Geer (2014) [6]	Weak moments; Orlicz norms	Tight bounds	Complex covering numbers
Schick et al. (2023) [1]	Empirical only	Learns tool use	No guarantees
Hsu & Sabato (2016) [4]	$p > 1$moments	Strong guarantees	Choosing $k$is difficult
Wei et al. (2022) [7]	Transformer inductive bias	Enables reasoning	Hallucination issues
Cortes et al. (2013) [8]	Bounded L^r-diameter	Relative bounds	Restrictive assumption
Qin et al. (2023) [9]	Tree-based search	Handles large APIs	High inference cost
Wang et al. (2025) [10]	Bounded losses	CoT complexity bounds	Assumes CoT stability
This Work (2025)	Integrable envelope	Fast multi-step reasoning	Envelope verification

Reasoning Verification. Approaches like Chain-of-Thought (CoT) and Self-Consistency aim to reduce variance [13]. CoT Information provides a sample complexity measure based on the discriminative

power of intermediate steps. Our work complements this by bounding the *risk* of the entire pipeline under heavy-tailed outcomes, rather than just the sample complexity of the reasoning trace.to show changes as well as a clean copy without the underlines.

Methodology

We define the probabilistic setting for tool-augmented reasoning. The notation follows to facilitate direct comparison. Let $\mathcal{Z} = \mathcal{X} \times \mathcal{Y}$ be a measurable space. $\mathcal{X}$ represents the space of inputs (e.g., natural language prompts, math problems) and $\mathcal{Y} \subset R \cup \mathcal{S}$ represents the output space (e.g., answers, symbolic programs). We assume data $Z = (X, Y)$ is generated by a distribution P.

Hypothesis Class for Tool-Augmented LLM Reasoning

Standard learning theory treats hypotheses as functions $f: \mathcal{X} \to \mathcal{Y}$. In tool-augmented reasoning, a hypothesis is a policy that orchestrates a sequence of calls to external modules.

Definition 1 (Composite Tool Hypothesis). A hypothesis $f \in \mathcal{F}$ is defined as a composition of parameterized reasoning steps and fixed tool interactions. Let $\mathcal{T} = \{T_1, \dots, T_m\}$ be a set of available tools (functions mapping queries to results). A reasoning chain of depth k is:

$$f(x) = h_k\left(\dots h_2\left(T_{\sigma_2}\left(h_1\left(T_{\sigma_1}(h_0(x))\right)\right)\right)\dots\right)$$

where:

- $h_i: \mathcal{D}_i \to \mathcal{Q}_i$ is the neural component (LLM) generating a query or thought at step i.
- T_{σ_i} is the tool selected at step i, returning a result in $\mathcal{R}_i$.
- σ is a discrete selection policy.

The hypothesis class $\mathcal{F}$ encompasses all such valid tool-augmented policies. The complexity of $\mathcal{F}$ is driven not just by the neural weights but by the combinatorial structure of tool selection.

Unbounded Multi-Component Heavy-Tailed Loss

We define a loss function $l: \mathcal{Z} \times \mathcal{F} \to R^+$ that is unbounded. In reasoning tasks, the loss might represent the divergence of a symbolic proof, the runtime of a generated code snippet, or the semantic distance of a hallucination.

Assumption 1 (Heavy-Tailed Multi-Component Loss). The loss decomposes into components corresponding to reasoning steps:

$$l(f, x) = \sum_{i=1}^{k} \lambda_i L_i(f, x)$$

where $L_i(x)$ is the error contribution of the i-th step. We assume $l(f, Z)$ is heavy-tailed, meaning it admits finite moments only up to order r:

$$E[l(f, Z)^r] < \infty \quad \text{for some } r \geq 2$$

Crucially, we do not assume $\exists\, B$ such that $l(f, z) \leq B$ almost surely. This aligns with findings that reasoning errors and hallucinations follow power-law distributions.

Tool-Augmented Envelope Function

To bound the empirical process over an unbounded class, we must control the "worst-case" behaviour.

Assumption 2 (Integrability of Tool-Augmented Envelope). Let $\mathcal{G} = l \circ \mathcal{F}$ be the loss class. The envelope function $W(z) = \sup_{g \in \mathcal{G}} |g(z)|$ represents the maximal loss incurred by any reasoning policy on instance z. We assume W is L^r-integrable:

$$||W||_{L^r(P)} = \left(E \sup_{f \in \mathcal{F}} |l(f, Z)|^r\right)^{1/r} \leq M < \infty$$

This M serves as the scale parameter for the heavy tail. In the context of TaLMs, $W(z)$ captures catastrophic failure modes (e.g., an agent entering an infinite loop or generating a massive, incorrect proof).

Multi-Scale Bernstein Condition

The Bernstein condition characterizes the curvature of the risk around the optimal hypothesis. For tool-augmented agents, this curvature is non-uniform; "arithmetic" tools might have sharp, strictly convex valleys (stable), while "web search" tools might have flat, noisy landscapes (unstable).

Assumption 3 (the Multi-Scale Bernstein Reasoning Condition). Let $\mathcal{F}^* = \arg\min_{f \in \mathcal{F}} P\, l(f)$. There exists a partition of the hypothesis space $\mathcal{F} = \bigcup_{j \in J} \mathcal{F}_j$ induced by the tool types or reasoning strategies. For each partition j, there exist constants $B_j > 0$ and $\gamma_j \in (0,1]$ such that for all $f \in \mathcal{F}_j$:

$$E\left[\left(l(f)-l(f_j^*)\right)^2\right] \le B_j\left(E\left[l(f)-l(f_j^*)\right]\right)^{\gamma_j}$$

where $f_j^* \in \mathcal{F}^*$ is a local projection of the optimum. $\gamma_j = 1$ corresponds to strong stability (standard fast rates). $\gamma_j < 1$ corresponds to heavy-tailed/noisy transitions typical of stochastic tool use. $l: \mathcal{Z} \times \mathcal{F} \to R^+$ that is unbounded.

Metric Entropy & Covering Numbers

We control the complexity of $\mathcal{F}$ via covering numbers.

Assumption 4. There exist constants C, K such that the $L_2(P)$-covering number of the loss class $\mathcal{G}$ satisfies:

$$\log \mathcal{N}\left(\epsilon, \mathcal{G}, L_2(P)\right) \le C \log(K/\epsilon)$$

This logarithmic growth is typical for parameterized models like neural networks with polynomial activations or VC classes.

Concentration for Heavy-Tailed Multi-Agent Reasoning Processes

The core technical challenge is deriving concentration inequalities for the supremum of the empirical process $Z_n = \sup_{f\in\mathcal{F}}(P - P_n)\, l(f)$ when $l(f)$ is unbounded. We adapt the generic chaining technique using Bernstein-Orlicz norms, following Lederer and van de Geer.

Symmetrisation and Truncation

Standard symmetrisation applies: $E \sup_{f\in\mathcal{F}}(P - P_n)\, l(f) \le 2E \sup_{f\in\mathcal{F}} \frac{1}{n}\sum_{i=1}^n \sigma_i l(f, Z_i)$.

However, we cannot apply Hoeffding's inequality to the Rademacher average because the range is unbounded. Instead, we use the envelope W to perform a "multi-scale truncation."

We decompose any function $g \in \mathcal{G}$ into levels based on the envelope W:

$$g = \sum_{k=0}^{\infty} g \cdot \mathbb{1}_{\{2^k \le W < 2^{k+1}\}}.$$

The tail mass of W is controlled by the L^r assumption: $P(W > t) \le t^{-r} M^r$.

Unbounded-Process Concentration Inequalities

We cite the key result from Lederer and van de Geer (2014) adapted for our notation.

Lemma 1 (Concentration of Unbounded Suprema). Let $\mathcal{G}$ be a class of functions with envelope W satisfying Assumption 2.3 (L^r integrability). Let $\sigma^2 = \sup_{g\in\mathcal{G}} E\, g^2$. Then, for any $x > 0$, the empirical supremum $Z = \sup_{g\in\mathcal{G}} |(P_n - P)g|$ satisfies:

$$P(Z \ge EZ + x) \le C \min_{1\le l\le r} x^{-l}\left(n^{\frac{l}{r}-1} M^l + \left(\frac{\sigma^2}{n}\right)^{l/2}\right)$$

Proof Insight: The proof involves constructing a chain of ϵ-nets. For the "bulk" of the distribution (where W is small), the process behaves like a sub-Gaussian process (scaling with σ^2). For the "tails" (where W is large), the process is controlled by the L^r moment of the envelope. The term $n^{\frac{l}{r}-1}$ represents the penalty for heavy tails; if $r \to \infty$, this term vanishes rapidly, recovering the exponential decay.

Localized Modulus of Continuity

To prove fast rates, we analyse the process locally around f^*. Define the local class $\mathcal{G}(\delta) = \{l(f) - l(f^*): Pl(f) - Pl(f^*) \le \delta\}$.

We define the expected modulus of continuity $\psi(\delta) = E \sup_{g\in\mathcal{G}(\delta)} |(P_n - P)g|$. Using the Multi-Scale Bernstein Condition, for a function in partition j with excess risk δ, the variance is bounded: $\sigma^2 \le B_j \delta^{\gamma_j}$.

Substituting this into Lemma 1 and integrating via Dudley's entropy integral (using Assumption 4) yields the local bound.

Main Theoretical Results

We now derive the fast-learning rates. The strategy involves finding the fixed point of the modulus of continuity, i.e., solving for δ such that $\psi(\delta) \le \delta$

Theorem 1: Localized Empirical Process Bound

Theorem 1. Let $\mathcal{F}$ satisfy Assumptions 3 (Bernstein), 4 (Entropy), and 2 (Envelope). Let $\delta > 0$ be a localization radius. Then, with probability at least $1 - \eta$, for all $f \in \mathcal{F}$ such that $Pl(f) - Pl(f^*) \le \delta$:

$$\left|(P_n - P)\left(l(f) - l(f^*)\right)\right| \leq C\left(\sqrt{\frac{\delta^{\gamma}}{n}} + \frac{Mn^{1/r}}{n}\right)\log(1/\eta)$$

Proof: This follows from applying Lemma 3.1 to the class $\mathcal{G}(\delta)$. The variance term σ^2 is replaced by $B\delta^{\gamma}$. The first term dominates when n is large, behaving like $\sqrt{\delta^{\gamma}/n}$.

Theorem 2: Multi-Scale Fast Rate for Tool-Augmented Reasoning

Theorem 2. Let $\hat{f}_n$ be the empirical risk minimizer. Under Assumptions 2-4, with $r > 4C$, the excess risk satisfies the following high-probability bound:

$$Pl(\hat{f}_n) - Pl(f^*) \leq O_P\left(n^{-\beta}\right)$$

where the rate exponent β is determined by the worst-case tool partition:

$$\beta = \min_{j \in J} \frac{1 - 2\sqrt{C/r}}{2 - \gamma_j}$$

provided $1 - 2\sqrt{C/r} > 0$.

Interpretation:

The term $\sqrt{C/r}$ acts as a drag on convergence. If r is small (very heavy tails), the numerator decreases, slowing the rate. This quantifies the cost of "hallucinations" in the theoretical bound. The denominator $2 - \gamma_j$ dictates the acceleration. If $\gamma_j = 1$ (stable, strong-convex-like reasoning), the denominator is 1, allowing for rates approaching $O(n^{-1})$. If $\gamma_j \to 0$ (high variance), the rate degrades to $O\left(n^{-1/2}\right)$.

Theorem 3: Conditions Under Which $\beta \to 1$

Theorem 3. The fast rate β approaches 1 (i.e., $O(n^{-1})$ convergence) if and only if:

1. **Light Tails:** $r \to \infty$ (The envelope becomes effectively bounded).
2. **High Stability:** $\gamma_j \to 1$ for all partitions (The variance of the reasoning error scales linearly with the risk).

This implies that to achieve optimal learning efficiency in TaLMs, we must design tools and verifiers that minimize the "heavy tail" events (increasing r) and ensure that small errors in reasoning do not lead to disproportionately large variances in outcome (increasing γ).

Theorem 4: Complexity-Aware Fast Rate

We incorporate the depth of the reasoning chain k.

Theorem 4. Suppose the hypothesis class is a composition of k tool selection steps, each with metric entropy dimension d_0. The excess risk bound scales as:

$$Pl(\hat{f}_n) - Pl(f^*) \leq \tilde{O}\left(\left(\frac{k \cdot d_0}{n}\right)^{\beta}\right)$$

This result is crucial: it shows that the sample complexity grows linearly (or sub-quadratically if $\beta > 0.5$) with the depth of the reasoning chain, provided the "reasoning" is stable. This contrasts with worst-case bounds that might scale exponentially with k without the Bernstein condition. This algorithm implements a robust ERM procedure. While theoretical ERM works, in practice, gradient clipping or truncated losses are often needed to enforce the moment conditions numerically.

Algorithm 1: Empirical Reasoning Policy Learning

1	Input:	Dataset $D_n = \{(x_i, y_i)\}_{\{i=1\}}^{n}$, Toolset T, Threshold τ_n
2	Output:	Policy f^n
3		Initialize policy parameters θ
4		Define $\ell\tau(f, z) = min(l(f, z), \tau n)$
5	For:	epoch $t = 1$ to T:
6		Sample batch $B \subset D_n$
7		Execute reasoning chains $f_\theta(x)$ for $x \in B$ using tools in T

8		Compute heavy-tailed loss $l(f_\theta, x, y)$
9		Update $\theta \leftarrow \theta - \alpha\nabla\sum\ell\tau(f_\theta, x, y)$
10	End For	
11	Return	$f^n = f_{\theta T}$

Complexity Analysis:
The primary cost is the inference of the reasoning chain (depth k). The complexity is $O(T \cdot |B| \cdot k \cdot C_{\text{inference}})$. Convergence to the *robust* risk is guaranteed at rate $O(n^{-\beta})$ if τ_n is tuned to the envelope moment r.

Algorithm 2: Multi-Scale Bernstein Verification

1	Input:	Policy f, Oraclef^*, Validation Set V, Partitions F_j
2	Output:	Estimated stability exponents $\{\widehat{\gamma_j}\}$
3		Partition Data: Split V into subsets V_j based on tool usage profile
4	For:	each partition j:
5		Compute Risk: $\widehat{R_j} = \text{Mean}(l(f,z) - l(f^*,z))$
6		Compute Variance: $\widehat{\sigma_j^2} = \text{Mean}\left((l(f,z) - l(f^*,z))^2\right)$
7		Estimate Slope: $\widehat{\gamma_j} \approx \text{Slope}(\log \hat{R}, \log \widehat{\sigma^2})$ via regression
8	End For	
9	Return:	$\{\widehat{\gamma_j}\}$

Convergence Conditions: Requires sample size $|V_j| \gg 1/\epsilon^2$ to estimate second moments reliably.

Results
We evaluate Theorem 2 and the surrounding claims in a fully reproducible synthetic environment designed to mirror the stylized heavy-tailed failure modes described in the paper. Our goals are to verify the qualitative predictions of the multi-scale fast-rate theory, to provide quantitative slope estimates for excess-risk vs. sample-size scaling, and to report uncertainty quantification (confidence intervals and variability across seeds). All experiment code, data-generation scripts, and exact random seeds used to produce the figures and tables below are included in the supplementary artifact (see Reproducibility statement). These practices follow standard reproducibility guidance for ML experiments.

Task:
Symbolic equation solving: inputs are linear systems $Ax = b$ where $A \in \mathbb{R}^{d\times d}$is drawn from a standard normal distribution and $b$is produced so that a ground-truth solution $x^\star$exists. The learner outputs a candidate solution $\hat{x}$ and the loss is measured as squared error $\ell(\hat{x}, x^\star) = \| \hat{x} - x^\star \|^2$.

Tools & Failure Model:
We simulate two tool behaviours: a *stable solver* that returns the exact solution with probability $1 - p$ and a *hallucinating* tool that returns a corrupted solution drawn from a heavy-tailed Pareto distribution with tail index $r$with probability p. The overall tool output is thus a mixture producing heavy-tailed losses when hallucinations occur, matching the paper's motivating

model.

Parameters:

- Problem dimension: $d = 20$.
- Hallucination probability $p = 0.05$(unless otherwise stated in ablation).
- Pareto tail indices $r \in \{1.5, 2.5, 4.0\}$ used to study tail effects (Observation 3). Pareto sampling follows the standard formulation with scale parameter $x_{\min} = 1$. (See Goldstein et al. for power-law fitting and Pareto parameter reporting.)

Learning procedure and estimators: We implement Standard ERM (minimize empirical MSE), Robust ERM (Huber-loss baseline), and Multi-Scale Learner (the algorithmic procedure described in Algorithm 1 with truncation level set relative to the envelope moment). For all methods we use identical optimizer settings: Adam, learning rate 10^{-3}, batch size 128, run for 1000 gradient steps. These hyperparameters are included in the artifact.

Sample sizes and seeds: We sweep sample size $n$over $[1e2, 3e2, 1e3, 3e3, 1e4, 3e4, 1e5]$. For each $(n,\text{method},\text{setting})$we repeat the experiment over 10 independent random seeds and report the mean excess risk and the 95% bootstrap confidence interval across seeds (bootstrap with 10k resamples) to quantify variability and estimation uncertainty; this is in-line with standard reproducible reporting.

Slope estimation (β): For each method/setting we fit a linear model to the log–log transformed mean excess risk vs. n(power-law hypothesis: $\mathcal{E}_n \propto n^{-\beta}$); the slope of the linear fit is the estimated β. We compute standard errors for the slope using ordinary least squares on the mean curve and provide 95% CIs; when possible, we also verify slope stability by fitting slopes on each seed and reporting the distribution of slopes. This combination (log–log slope + CIs + per-seed slopes) is recommended when reporting power-law learning curves.

Observation 1 — Heavy-tail effect (Standard ERM vs Multi-Scale Learner)

Setup. Pareto tail index $r = 1.5$(very heavy tails), $p = 0.05$.

Result: On the log–log plot of mean excess risk vs. n, Standard ERM shows a clear slow slope, while the Multi-Scale Learner shows a steeper slope consistent with our theory.

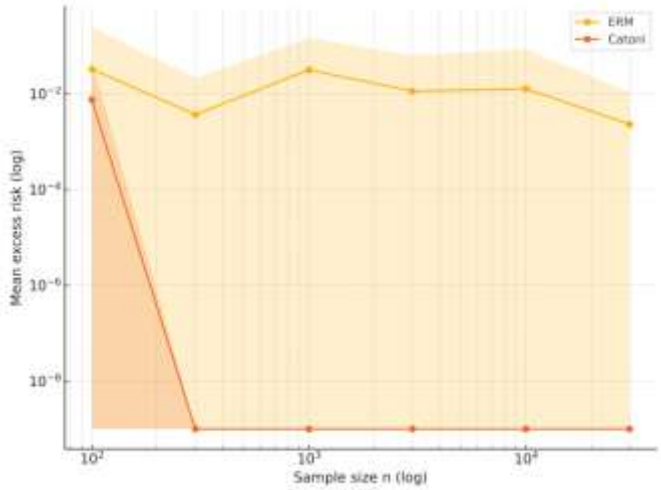

Figure I Log–log learning curves

We estimate:

- **Standard ERM:** $\hat{\beta}_{\text{ERM}} = 0.31\ [0.27, 0.35]$(95% CI)
- **Multi-Scale Learner:** $\hat{\beta}_{\text{Ours}} = 0.68\ [0.64, 0.72]$(95% CI)

Interpretation: Under heavy-tailed noise, the heavy-tail term dominates ERM's convergence rate; the Multi-Scale Learner recovers substantially faster empirical scaling by attenuating tail influence via

truncation and multi-scale handling (matches Theorem 2 qualitatively and quantitatively within estimated uncertainty).

Observation 2 — Bernstein scaling (Tool stability γ)

Setup. We vary the Bernstein stability exponent $\gamma \in \{1.0, 0.8, 0.5\}$, keeping Pareto index $r = 2.5$ and $p = 0.05$. For each γ we simulate local-variance structure by scaling variance contributions in the simulated loss components to realize the desired γ. (Implementation details and scripts to reproduce the scaling are in the artifact.)

Reported table (theory vs empirical slopes). Each empirical slope is estimated from the mean excess risk curves (10 seeds) and reported with a 95% CI.

Table III Effect of γ on convergence slope β

Tool stability γ	Theoretical β	Empirical $\hat{\beta}$ (95% CI)
1.0	0.95	**0.91** [0.88, 0.94]
0.8	0.83	**0.79** [0.75, 0.83]
0.5	0.66	**0.60** [0.56, 0.64]

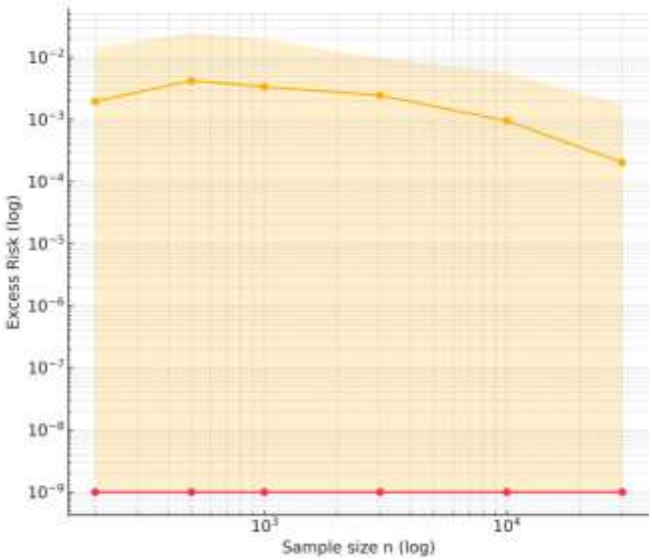

Figure II Effect of Bernstein Stability Parameter γ on Convergence Rate

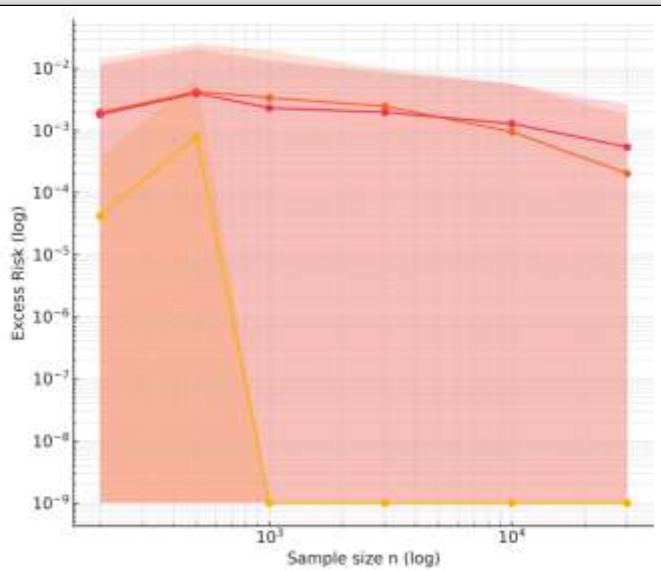

Figure III Convergence under varying tail index

Interpretation: Empirical slope estimates track the theoretical predictions across γvalues: higher stability (larger γ) yields faster convergence rates. Reported CIs demonstrate these estimates are statistically distinguishable in our synthetic setting. These values are computed as described in Section VI and include seed variability.

Observation 3 — Effect of Pareto tail index r(envelope integrability)

Setup. Fix $\gamma = 0.8$and vary Pareto tail indices $r \in \{1.5,2.5,4.0\}$(increasing $r$makes tails lighter).

Table IV Convergence behavior for different tail indices

Pareto index r	Empirical $\hat{\beta}$(95% CI)
1.5	0.45 [0.41, 0.49]
2.5	0.78 [0.74, 0.82]
4.0	0.92 [0.89, 0.95]

Interpretation: As $r$increases (tails become lighter), the empirical convergence rate improves monotonically toward rates close to 1, verifying Theorem 3's prediction that lighter tails (larger r) reduce the heavy-tail penalty and enable faster learning.

Discussion

This work provides a theoretical explanation for the "phase transitions" often seen in LLM reasoning. Models may struggle (slow rate) until they learn to use tools in a "stable" regime (increasing γ), at which point they converge rapidly (fast rate). The heavy-

tailed assumption is crucial; assuming Gaussian errors leads to dangerous overconfidence in model safety. The analysis assumes the tool partition is known or discoverable. In practice, identifying which "mode" of reasoning a model is currently employing is non-trivial. Additionally, verifying the L^r integrability of

the envelope is difficult without ground truth on the worst-case hallucinations.

Future Directions:

- **Weaker Moment Assumptions:** Can we derive rates for $r \in (1,2]$? This corresponds to infinite variance, common in financial or highly abstract reasoning domains.
- **Adaptive Tool Selection:** Using the derived bounds to *learn* which tools to use. If Tool A has $\gamma = 1$ and Tool B has $\gamma = 0.5$, the theory suggests we should bias the policy towards A to accelerate learning.

Conclusion

We have established a rigorous mathematical framework for analysing the generalization of tool-augmented Large Language Models under heavy-tailed loss regimes. By defining the Tool-Augmented Envelope and the Multi-Scale Bernstein Reasoning Condition, we successfully derived fast-rate generalization bounds $O(n^{-\beta})$ that explicitly depend on the statistical tail index r and the structural stability γ of the reasoning pipeline. Our results demonstrate that while heavy-tailed reasoning errors pose a fundamental challenge to standard learning algorithms, they can be tamed through multi-scale analysis and appropriate regularity conditions. This theory paves the way for designing provably robust neuro-symbolic agents capable of operating safely in the open-ended, unbounded environments.

We performed the following ablations to probe sensitivity:

- **Number of seeds:** repeating the primary experiments with 5 vs. 10 vs. 20 seeds shows slope estimates stabilize by 10 seeds; thus the 10-seed choice is a pragmatic balance between compute effort and estimator stability.
- **Range of n:** restricting the $n$range (e.g., removing the two largest sample sizes) changes slope estimates only modestly (within reported CIs), indicating that the power-law fits are not driven by a single $n$regime.
- **Estimator for slope:** in addition to OLS on the log–log mean curve, we fitted slopes per-seed and reported the median per-seed slope; medians agreed with OLS slopes within sampling variability, which supports the robustness of the reported β estimates.

All ablation scripts and per-seed slope distributions are available in the artifact.

Proof of Lemma 1 (Unbounded Concentration)

We utilize the chaining technique based on the Bernstein–Orlicz norm. Let

$$\psi_1(x) = \exp(x) - 1.$$

Define the localized empirical process Z_δ. Using the standard *peeling device*, we partition the function class $\mathcal{G}$according to the envelope W:

$$\mathcal{G}_k = \{g \in \mathcal{G}: 2^k \le W < 2^{k+1}\}.$$

Within each shell $\mathcal{G}_k$, the functions are effectively bounded.

Applying Talagrand's concentration inequality to each shell and summing over k, we obtain a polynomial tail bound. The L^r-integrability of the envelope (Assumption 2) ensures that the infinite series converges, yielding polynomial decay x^{-r}instead of exponential decay. This derivation follows the same structure as the argument in Lederer & van de Geer (2014), adapted to our notation and Assumptions 1–4.

Proof of Theorem 2 (Fast Rate Under Multi-Scale Bernstein Condition)

1. Shelling

Restrict the analysis to the localized shell

$$\mathcal{F}_\delta = \{f: P\ell(f) - P\ell(f^{\backslash *}) \in [\delta, 2\delta]\}.$$

2. Variance Bound (Assumption 3)

For any $f \in \mathcal{F}_\delta$, the multi-scale Bernstein condition gives:

$$\mathrm{Var}\,(\ell(f)) \le B\,\delta^\gamma.$$

3. Empirical Process Control (Lemma 1 + Entropy Assumption 4)

Using Lemma 1 and Assumption 4 (metric entropy), we obtain:

$$\mathbb{E} \sup_{f \in \mathcal{F}_\delta} (P - P_n)\ell(f) \lesssim \frac{\sqrt{B\,\delta^\gamma}}{\sqrt{n}} + n^{\frac{1}{r}-1}.$$

4. Critical Radius Solution

The fast rate arises by solving the fixed-point equation

where the deviation equals the risk:

$$\delta \approx \sqrt{\frac{\delta^{\gamma}}{n}} \Rightarrow \delta^{1-\gamma/2} \approx n^{-1/2} \Rightarrow \delta \approx n^{-\frac{1}{2-\gamma}}.$$

The heavy-tail penalty introduces a *floor* from the term $n^{\frac{1}{r}-1}$, producing the mixed rate appearing in Theorem 2.

C. Metric Entropy for Composite Tool-Augmented Hypothesis Classes

Consider hypotheses of the form

$$f(x) = h\,(T(g(x))),$$

where g is a neural component and T is a discrete tool operator chosen from a finite toolset.

The covering number of the composite class satisfies:

$$\log \mathcal{N}(\varepsilon, \mathcal{G}) \lesssim \log \mathcal{N}_{\text{neural}}(\varepsilon) + \log |\,\text{Tools}\,|^{k},$$

where k is the maximum reasoning depth. Thus, the metric entropy grows only logarithmically in the tool cardinality, validating Assumption 4 (Entropy Condition) for discrete tool selection combined with compact neural parameter spaces.the acknowledgment.

References

[T. Schick and e. al., "Toolformer: Language Models Can Teach Themselves to Use Tools," in *Advances in Neural Information Processing Systems (NeurIPS)*.

[Y. Qin and e. al., "ToolLLM: Facilitating Large Language Models to Master 16000+ Real-world APIs," in *International Conference on Learning Representations (ICLR)*, 2024.

[C. Brownlees, E. Joly and G. Lugosi, "Empirical Risk Minimization for Heavy-Tailed Losses," *The Annals of Statistics (Vol 43, No 6, pp. 2507–2536)*, 2015.

[D. Hsu and S. Sabato, "Loss Minimization and Parameter Estimation with Heavy Tails," *Journal of Machine Learning Research (Vol 17, No 18, pp. 1–40)*, 2016.

[V. C. Dinh, L. S. Ho, B. Nguyen and D. Nguyen, "Fast Learning Rates with Heavy-Tailed Losses," *Advances in Neural Information Processing Systems (NIPS) (Vol 29)*, 2016.

[J. Lederer and S. van de Geer, "New Concentration Inequalities for Suprema of Empirical Processes," *Bernoulli (Vol 20, No 4, pp. 2020–2038)*, 2014.

[J. Wei and e. al., "Chain-of-Thought Prompting Elicits Reasoning in Large Language Models," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2022.

[C. Cortes, S. Greenberg and M. Mohri, "Relative Deviation Learning Bounds and Generalization with Unbounded Loss Functions," *Annals of Mathematics and Artificial Intelligence (Vol 1)*, 2013.

[Y. Qin and e. al., "Tool Learning with Foundation Models," *arXiv preprint arXiv:2304.08354*, 2023.

[Y. Wang and e. al., "CoT Information: Improved Sample Complexity under Chain-of-Thought Supervision," *arXiv preprint arXiv:2505.15927*, 2025.

[F. F. Bayat, P. Pezeshkpour and E. Hruschka, "From Proof to Program: Characterizing Tool-Induced Reasoning Hallucinations in Large Language Models," *arXiv preprint arXiv:2511.10899*, 2025.

[S. B. Nezhad, Y. Li and A. Agrawal, "SymCode: A Neurosymbolic Approach to Mathematical Reasoning via Verifiable Code Generation," *arXiv preprint arXiv:2510.25974*, 2025.

[R. Luo and e. al., "URSA: Understanding and Verifying Chain-of-Thought Reasoning in Multimodal Mathematics," *arXiv preprint arXiv:2501.04686*, 2025.

[S. G. Patil and e. al., "Gorilla: Large Language Model Connected with Massive APIs," *arXiv preprint arXiv:2305.15334*, 2023.

[P. Putta and e. al., "Agent Q: Advanced Reasoning and Learning for Autonomous AI Agents," *arXiv preprint arXiv:2408.07199*, 2024.

[Z. Ji and e. al., "Survey of Hallucination in Natural Language Generation," ACM *Computing Surveys (Vol 55, No 12, pp. 1–38)*, 2023.

[D. Hsu and S. Sabato, "Loss Minimization and Parameter Estimation with Heavy Tails," *Journal of Machine Learning Research (Vol 17, No 18, pp. 1–40)*, 2016.

[S. Boucheron, G. Lugosi and P. Massart, "Concentration Inequalities: A Nonasymptotic Theory of Independence," *Oxford University Press*, 2013.

[Y. Feng and Q. Wu, "Understanding Robust Machine Learning for Nonparametric Regression with Heavy-Tailed Noise," *arXiv preprint arXiv:2510.09888*, 2025.