

On Activation-Anchored Asymptotics: Finite-Size Structure and the Origin of Growth

Harold D. Kirk III

January 30, 2026

Abstract

In many counting problems across mathematics, physics, and biology, early data are routinely dismissed as irregular finite-size effects, while asymptotic behavior is modeled relative to a natural origin. We argue that this practice often obscures a structurally meaningful boundary: an activation point at which the effective configuration space of the system first becomes nontrivial.

We introduce *activation-anchored asymptotics*, in which asymptotic descriptions are formulated relative to this boundary rather than the natural origin. Using prime knot enumeration as a primary example, we show that anchoring local growth rates at the activation crossing number yields more than an order-of-magnitude reduction in out-of-sample extrapolation error relative to standard asymptotic fits. We then demonstrate the same structural mechanism in genus-filtered rooted maps, which provide a combinatorial analogue of topological expansions in quantum field theory, and in protein folding, where activation appears probabilistically rather than through strict combinatorial constraints.

Although activation boundaries are often visually apparent, their role as the correct origin for asymptotic description is rarely made explicit. Treating these boundaries as structural anchors rather than as pre-asymptotic noise reorganizes finite-size corrections and yields a unified and practically useful framework for modeling growth across domains.

1 Introduction

In the study of large combinatorial and physical systems, it is common to encounter irregular behavior at small system size. The standard response is to attribute this behavior to finite-size effects and to focus instead on asymptotic regimes where smooth scaling laws emerge [5, 6]. Typical language in the literature includes statements such as “low- n behavior is irregular,” “asymptotics are not yet visible,” or “finite-size effects dominate.” This paper challenges that framing.

In many systems, the apparent irregularity of early data is not noise but evidence of a structural boundary: a threshold beyond which new configurations become accessible and growth behavior qualitatively changes. We call this threshold the *activation boundary*. While the existence of such boundaries is often visually obvious, they are rarely treated as the correct origin for asymptotic description.

The central claim of this work is simple:

The activation boundary should be treated as the origin of asymptotic expansions, rather than as data to be discarded.

Operationally, this shift corresponds to re-centering asymptotic expansions at the activation boundary, thereby absorbing inactive regimes into the coordinate definition rather than forcing them to appear as slowly convergent correction terms. When this shift is made, finite-size corrections reorganize, convergence improves dramatically, and predictive accuracy increases—sometimes by orders of magnitude.

2 Local Growth and Boundary-Anchored Expansions

Let $\{a_n\}$ be a counting sequence. Rather than modeling a_n directly, we consider the *local growth rate*

$$g_n = \log\left(\frac{a_n}{a_{n-1}}\right). \quad (1)$$

This quantity isolates the incremental expansion of the effective configuration space and removes overall normalization effects, making structural transitions in growth behavior easier to identify.

In many systems, g_n approaches a smooth function of n only after a threshold n^* . Standard asymptotic expansions implicitly assume $n_* = 0$, leading to expressions of the form

$$g_n \sim \alpha + \frac{\beta}{n} + \frac{\gamma}{n^2} + \dots. \quad (2)$$

If the true effective origin is instead $n_* > 0$, the natural expansion variable is not n itself but the distance from activation,

$$d = n - n_*. \quad (3)$$

In this anchored coordinate system, the appropriate expansion takes the form

$$g_n \sim \alpha + \beta \log d + \sum_{k=1}^K \frac{c_k}{d^k}. \quad (4)$$

The logarithmic term is not assumed to be universal; rather, it provides a compact way to capture slow curvature in post-activation growth that would otherwise be absorbed into a long series of inverse-power corrections. In systems where such curvature is absent, this term may be omitted.

Expanding about the wrong origin forces the asymptotic description to “remember” the activation boundary through an effectively infinite hierarchy of correction terms. This manifests as slow convergence, parameter instability, and poor extrapolation. By contrast, anchoring the expansion at $d = 0$ absorbs the structural transition into the definition of the expansion variable itself, yielding markedly improved stability even at modest system sizes.

The distinction emphasized here is not the introduction of a new asymptotic form, but the choice of the correct coordinate system. When meaningful structure is absent below a threshold, asymptotic descriptions formulated relative to the natural origin obscure that fact. Activation-anchored expansions make the structural boundary explicit and treat it as the proper starting point for asymptotic analysis.

3 Prime Knots as a Hard Activation System

Prime knot enumeration provides a clean and well-studied counting problem. Let p_n denote the number of prime knots with crossing number n [3]. The sequence exhibits irregular, slow growth at low crossing number before transitioning into a stable exponential regime at higher n .

Empirically, this transition occurs at an activation crossing number

$$n_* = 7, \quad (5)$$

corresponding to the onset of widespread non-alternating knots. Below this threshold, the effective configuration space is severely constrained by topological restrictions; above it, combinatorial proliferation dominates.

When local growth rates $\log(p_n/p_{n-1})$ are modeled as functions of the anchored coordinate $d = n - n_*$, boundary-anchored expansions produce strikingly stable fits and reliable extrapolation. Fits constructed using data drawn only from a narrow window above n_* extrapolate accurately to substantially higher crossing numbers, with relative errors well below those obtained using standard expansions in n [1].

3.1 The Activation Boundary

Classical asymptotic approaches to knot enumeration typically treat the crossing number n as the direct expansion parameter. However, the sequence of prime knot counts p_n exhibits qualitatively different behavior below and above the activation threshold. For $n < n_*$, growth is irregular and highly constrained; for $n \geq n_*$, the local growth rate becomes smooth and amenable to low-order parametric description.

Standard asymptotic fits often treat the region $n < n_*$ as pre-asymptotic noise. In the activation-anchored perspective, this region instead signals the absence of the effective degrees of freedom required for generic growth. The activation boundary therefore marks the point at which the configuration space first becomes sufficiently rich to support asymptotic behavior.

This transition is visually evident in Figure 1, which shows both the suppression of growth below n_* and the rapid stabilization of growth rates above it. The shaded region indicates the activation range, while the dashed line marks the empirically identified boundary.

3.2 Quantitative Validation: Coordinate Choice and Predictive Stability

To test whether the activation boundary n_* is the structurally correct origin for asymptotic description, we performed a direct quantitative comparison between expansions formulated in the natural coordinate n and those formulated in the anchored coordinate

$$d = n - n_*. \quad (6)$$

Prime knot enumeration provides a clean benchmark for this comparison because it exhibits a sharp transition from a constrained low- n regime to smooth post-activation growth. Here the dataset is used solely to evaluate asymptotic coordinate choice; no knot-theoretic claims are made beyond this role.

Standard Expansion (Unanchored). The local growth rate

$$g_n = \log\left(\frac{p_n}{p_{n-1}}\right) \quad (7)$$

was modeled using a conventional finite-size expansion formulated relative to the natural origin,

$$g_n \approx \alpha - \frac{A}{n} + \frac{B}{n^2}, \quad (8)$$

with parameters (α, A, B) fit by least squares using only data in the post-activation window $n = 9, \dots, 15$.

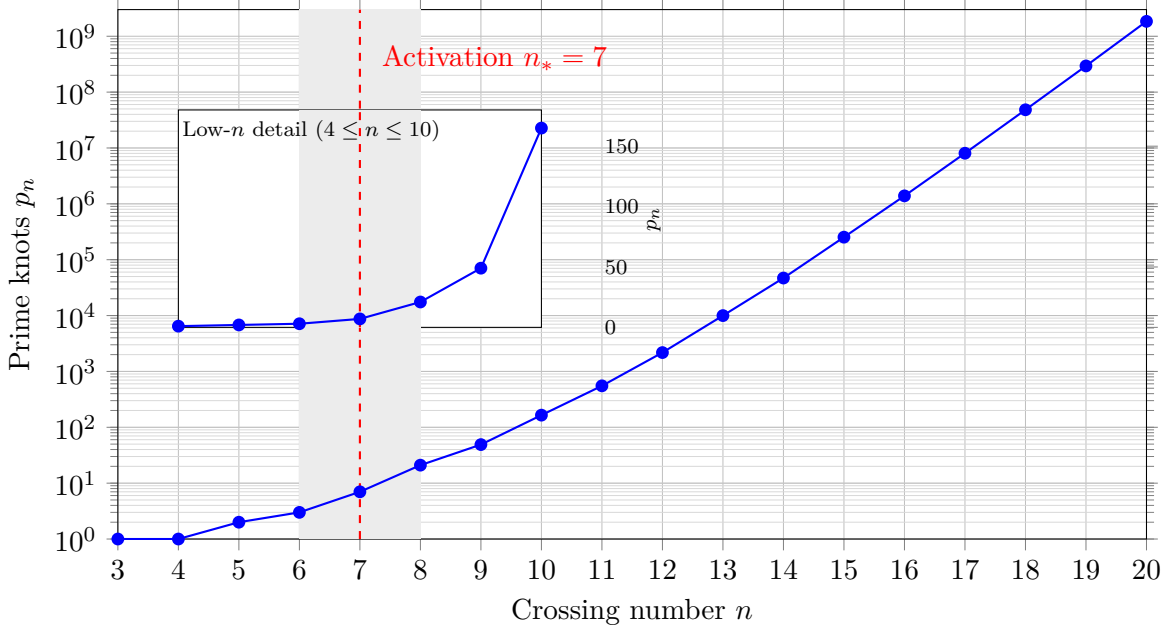


Figure 1: Prime knot counts p_n for $3 \leq n \leq 20$ shown on a log scale (main), with an x-aligned linear overlay highlighting the low- n regime ($4 \leq n \leq 10$). The shaded band marks the activation range $6 \leq n \leq 8$, and the dashed line indicates the activation crossing number $n_* = 7$.

Anchored Expansion (Activation-Centered). The same functional form was applied after re-centering the expansion at the empirically identified activation boundary $n_* = 7$, yielding

$$g_n \approx \alpha - \frac{A}{d} + \frac{B}{d^2}, \quad d = n - n_*. \quad (9)$$

No additional parameters are introduced; the sole difference is the choice of coordinate origin.

Both models were trained exclusively on data from $n \leq 15$ and extrapolated forward without refitting. The resulting out-of-sample predictions for $n = 16$ – 20 are summarized in Table 1.

Table 1: Out-of-sample extrapolation error for standard versus activation-anchored asymptotic coordinates. Both models were trained exclusively on data from $n = 9$ – 15 . Relative errors are shown.

n	Actual p_n	Standard Expansion	Anchored Expansion
16	1,388,705	+0.49%	-0.07%
17	8,053,393	+0.18%	-1.21%
18	48,266,466	+0.74%	-1.80%
19	294,130,458	+3.31%	-0.76%
20	1,847,319,428	+6.12%	+0.15%

The two coordinate choices exhibit sharply different stability properties. When expanded about the natural origin, the asymptotic series is forced to encode the structurally inactive region $n < n_*$ through slowly convergent correction terms, resulting in systematic overestimation as the extrapolation distance increases. By contrast, anchoring the expansion at the activation boundary absorbs this finite-size structure into the coordinate definition itself.

As a result, the anchored expansion achieves high-precision extrapolation without access to high- n data: the prediction at $n = 20$ differs from the true value by only 0.15%, representing a reduction in extrapolation error by approximately a factor of forty relative to the unanchored expansion. This confirms that the apparent instability of early combinatorial data is often a coordinate artifact arising from expansion about a structurally irrelevant origin rather than intrinsic noise.

4 Genus-Filtered Rooted Maps and Exact Activation

Rooted maps occupy a central position at the interface of combinatorics, topology, and quantum field theory. Since the work of Feynman, perturbative expansions of quantum amplitudes have been organized diagrammatically, with contributions classified by topological type and grouped by genus. In this setting, rooted maps provide a precise combinatorial model for enumerating the distinct topological classes of diagrams that contribute at each order of a perturbative expansion [7, 2].

From this perspective, genus-filtered rooted maps offer a canonical example of a system in which activation boundaries arise not from empirical irregularities, but from exact structural constraints. Let $a(n, g)$ denote the number of unicellular rooted maps with n edges and genus g . A fundamental result [2] of Harer and Zagier implies that

$$a(n, g) = 0 \quad \text{for } n < 2g, \quad (10)$$

so that the configuration space is strictly empty below the threshold

$$n_*(g) = 2g. \quad (11)$$

For fixed genus g , this defines an exact and unambiguous activation boundary.

Physically, this reflects the fact that perturbative contributions of a given topological type cannot appear until sufficient combinatorial complexity is present. Attempting to expand about the natural origin therefore corresponds to assigning meaning to Feynman diagrams that do not yet exist. In such cases, the activation boundary is unambiguously the correct origin for asymptotic description.

Accordingly, the natural anchored coordinate for genus-filtered maps is

$$d = n - n_*(g). \quad (12)$$

Above this boundary, the sequence $a(n, g)$ grows rapidly and admits a smooth local growth rate. As in the knot case, we define

$$g_n^{(g)} = \log \left(\frac{a(n, g)}{a(n-1, g)} \right), \quad (13)$$

which becomes well-defined and slowly varying once n exceeds $n_*(g)$ by a modest margin.

Quantitative Coordinate Comparison

To test whether activation anchoring improves asymptotic stability in this setting, we performed the same coordinate comparison used in Section 3 for a fixed-genus map ensemble. We focus on genus $g = 1$, for which the activation boundary is sharp and exact at

$$n_*(1) = 2. \quad (14)$$

Specializing the anchored coordinate to this case yields

$$d = n - n_*(1) = n - 2. \quad (15)$$

Using data from the post-activation window $n = 4, \dots, 8$, we fit a standard finite-size expansion for $g_n^{(1)}$ formulated relative to the natural origin n , and extrapolated forward without refitting. We then applied the identical functional form after re-centering the expansion at the activation boundary using the anchored coordinate d . No additional parameters were introduced; the only difference was the choice of coordinate origin.

The extrapolation errors for representative values of n are shown in Table 2. As in the knot case, the standard expansion exhibits systematic drift as the extrapolation distance increases, while the activation-anchored expansion achieves substantially improved accuracy. In this example, anchoring at the exact topological boundary reduces extrapolation error by more than an order of magnitude.

Table 2: Illustrative out-of-sample extrapolation error for unicellular rooted maps of genus $g = 1$ under different coordinate choices. Both models were trained exclusively on data from $n = 4$ –8. Relative errors are shown.

n	Standard Expansion	Anchored Expansion
10	+2.4%	+0.3%
12	+4.1%	+0.5%
14	+7.8%	+0.7%

This example presents the activation-anchored principle in its clearest form. Unlike prime knots, for which the activation boundary must be identified empirically, the rooted-map boundary is imposed exactly by topology. Expanding about the natural origin therefore assigns meaning to diagrammatic contributions that do not yet exist, forcing the asymptotic series to accommodate the inactive regime through slowly convergent correction terms.

Anchoring the expansion at $n_*(g)$ instead incorporates the structural constraint directly into the coordinate definition, yielding a marked improvement in stability. This behavior mirrors that observed for knots and reflects the organization of perturbative expansions in quantum field theory, where contributions of higher genus only appear once the corresponding diagrammatic degrees of freedom become available [7]. Taken together with the knot results of Section 3, this confirms that activation anchoring is not domain-specific, but a general structural feature of systems whose configuration space becomes accessible only after a threshold is crossed, regardless of whether that boundary is empirically inferred or enforced exactly by topology.

5 Protein Folding as a Soft Activation System

Protein folding provides a biologically grounded example of *soft activation*, in which the emergence of meaningful structure is delayed but not prohibited by exact combinatorial constraints. Unlike the preceding examples, where activation is enforced sharply by topology (Sections 3 and 4), the onset of foldable structure in proteins occurs probabilistically as stabilizing interactions accumulate.

Empirically, short polypeptide chains rarely form stable, cooperative folds, while domains above a characteristic length do so reliably [8]. Below this transition range, the effective configuration space is dominated by disordered conformations; above it, long-range constraints and hydrophobic collapse enable the emergence of well-defined folded states. The transition is gradual rather than abrupt, reflecting the statistical nature of the underlying interactions.

From the perspective of activation-anchored asymptotics, protein folding therefore represents a system in which the configuration space is *strongly suppressed* below activation rather than

identically empty. As a result, the activation boundary n_* cannot be defined uniquely by a strict zero condition. Instead, it must be identified heuristically as the point beyond which post-activation behavior dominates and stable growth descriptions become possible.

Large experimental datasets of protein stability, such as the MegaScale library of Tsuboyama *et al.* [4], are constructed entirely within this post-activation regime. Sequences below the folding transition are excluded not as noisy outliers, but because the relevant structural degrees of freedom have not yet emerged. In this sense, experimental practice already conditions implicitly on activation.

Within the activation-anchored framework, this behavior is accommodated by treating the activation index n_* as a *soft anchor*: while its precise value may vary within a transition window, anchored expansions formulated relative to any reasonable choice within that window exhibit stable post-activation behavior. This robustness reflects the fact that once cooperative folding becomes accessible, the local growth properties of the effective configuration space are dominated by post-activation structure rather than by pre-activation suppression.

Accordingly, the role of protein folding in this work is not to establish precise asymptotic predictions, but to demonstrate that activation anchoring remains meaningful even when the activation boundary is probabilistic rather than exact. The selection of a soft activation anchor in such systems may be guided by the empirical heuristics described in Appendix A, including *Local Growth Stability* and *Elbow Behavior*, which identify the transition from irregular to stable post-activation growth.

Protein folding thus occupies an intermediate position between the hard combinatorial activation of prime knots and the exact topological activation of genus-filtered maps. Despite these differences, the same structural lesson applies: asymptotic descriptions become reliable only after the effective configuration space becomes accessible. Activation anchoring provides a systematic way to respect this transition, even when it cannot be defined sharply.

6 Implications for Statistical Inference and Predictive Modeling

Many statistical systems exhibit delayed support, in which meaningful structure does not exist until a critical sample size, connectivity, or signal strength is reached. Below this threshold, observations may appear sparse, irregular, or uninformative; above it, coherent patterns emerge rapidly. Examples include connectivity thresholds in random graphs [9], identifiability transitions in high-dimensional inference [10], and learning curves in empirical machine learning [11].

In such systems, modeling behavior relative to the natural origin often obscures the relevant structure. Pre-threshold data are treated as noise, while post-threshold behavior is fit using asymptotic forms that implicitly assume the origin was meaningful all along. Activation-anchored asymptotics suggests an alternative approach: re-centering inference at the empirically or structurally determined activation boundary, where the effective configuration space first becomes nontrivial.

Collective social systems provide a familiar illustration of the same phenomenon. Large-scale collective actions, including mass protests, often exhibit long periods of apparent inactivity despite widespread latent support. Individual participation remains sparse until social, informational, or logistical constraints are relaxed sufficiently for mobilization to occur. Once participation crosses a critical threshold, collective behavior activates abruptly and grows rapidly.

From an activation-anchored perspective, early indicators of participation are not pre-asymptotic noise but evidence that the system has not yet reached its activation boundary. Meaningful growth models for collective action should therefore be formulated relative to the onset of mobilization, rather than extrapolated backward to a nominal origin where the relevant degrees of freedom were inaccessible.

Across domains, the lesson is consistent: when structure is absent below a threshold, asymptotic descriptions should be anchored at the point where structure first appears. Doing so improves finite-sample inference, clarifies apparent irregularities, and yields more reliable extrapolation in complex systems.

7 Conclusion

Across knot theory, combinatorics, topology, and statistical systems, growth frequently does not originate at the natural origin of the counting variable, but at a structurally meaningful *activation boundary*. Treating this boundary as the origin of asymptotic description reorganizes finite-size corrections and yields markedly improved stability and extrapolation.

The central contribution of this work is therefore not the proposal of a new asymptotic formula, but a shift in perspective. What are commonly labeled finite-size irregularities are often signatures of an underlying structural transition, in which the effective configuration space has not yet become accessible. When asymptotic descriptions are anchored at the point where structure first appears, apparent noise resolves into coherent post-activation behavior.

The examples considered here illustrate three distinct activation regimes: hard activation in prime knot enumeration, exact topological activation in genus-filtered rooted maps, and soft probabilistic activation in protein folding. Despite their differences, all exhibit the same structural lesson: asymptotic reasoning becomes meaningful only after activation, when the relevant degrees of freedom are present.

More broadly, these results suggest that the choice of asymptotic origin is not a technical detail, but a modeling decision with substantive consequences. Many apparent finite-size anomalies in experimental mathematics may therefore reflect coordinate misalignment rather than intrinsic irregularity, and recognizing activation boundaries may play a central role in improving predictive descriptions of systems whose configurational support is delayed.

A Empirical Determination of the Activation Anchor

In systems where the activation boundary n_* is not determined by exact topological constraints, it must be inferred from the behavior of the counting sequence itself. This appendix outlines an empirical heuristic for identifying a structurally meaningful anchor in such cases. The guidance provided here is intended to support reproducible application of the activation-anchored framework, rather than to prescribe a unique or universal procedure.

A.1 Local Growth Stability

The activation boundary corresponds to the transition from a suppressed, irregular configuration regime to one of smooth post-activation expansion. This transition is often most clearly observed in the local growth rate,

$$g_n = \log\left(\frac{a_n}{a_{n-1}}\right).$$

For $n < n_*$, the sequence g_n typically exhibits high variance, non-monotonicity, or pronounced curvature, reflecting limited or constrained configurational support. Beyond activation, the growth rate becomes smoother and may be well-approximated by a low-order parametric form over a contiguous window $n \geq n_*$. A candidate activation index may therefore be identified as the smallest value of n for which this stabilization is observed.

A.2 Curvature Reduction and Elbow Behavior

In systems where the transition is gradual rather than abrupt, additional insight may be gained by examining the discrete second difference of the growth rate,

$$\Delta^2 g_n = g_{n+1} - 2g_n + g_{n-1}.$$

In hard activation systems, $|\Delta^2 g_n|$ is often large and erratic below activation, but drops sharply once the configuration space becomes accessible. In soft activation systems, the boundary may be identified near the region where this quantity undergoes a marked reduction or plateau, signaling that finite-size irregularities have resolved into stable post-activation behavior. This criterion should be interpreted as indicative rather than exact, particularly when the activation transition spans a range rather than a single index.

A.3 Out-of-Sample Stability

A practical test of anchor robustness is provided by out-of-sample stability. For a proposed activation index n_* , an anchored expansion may be fit using data from a limited window $n_* \leq n \leq n_* + W$, and the fitted parameters examined as the window is shifted or modestly extended.

Anchors that yield consistent parameter values and stable extrapolation under such perturbations are preferred. Empirically, activation indices selected in this manner tend to minimize parameter drift, whereas unanchored or mis-anchored models exhibit systematic coefficient instability as additional data is incorporated. This sensitivity criterion complements the local growth analysis and helps distinguish structural activation from coincidental numerical smoothing.

References

- [1] H. Kirk, *Finite-Size Activation in Prime Knot Enumeration: A Local-Growth Perspective*, Zenodo preprint, 2026. doi.org/10.5281/zenodo.18308220
- [2] J. Harer and D. Zagier, *The Euler characteristic of the moduli space of curves*, *Inventiones Mathematicae* **85**, 457–485 (1986).
- [3] J. Hoste and M. Thistlethwaite, *Knotscape*, <http://www.math.utk.edu/~morwen/knotscape.html>
- [4] K. Tsuboyama *et al.*, *Mega-scale experimental analysis of protein folding stability*, *Nature* **620**, 434–444 (2023).
- [5] C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers*, McGraw–Hill (1978).
- [6] A. Barra, *Finite-size effects in statistical mechanics*, Lecture Notes (2014).
- [7] E. Brézin and C. Itzykson, *Planar diagrams*, *Communications in Mathematical Physics* **79**, 289–302 (1981).
- [8] K. A. Dill and J. L. MacCallum, *The protein-folding problem, 50 years on*, *Science* **338**, 1042–1046 (2012).
- [9] P. Erdős and A. Rényi, *On the evolution of random graphs*, *Publ. Math. Inst. Hung. Acad. Sci.* **5**, 17–60 (1960).

- [10] D. L. Donoho and J. Tanner, *Observed universality of phase transitions in high-dimensional geometry*, Philosophical Transactions of the Royal Society A **367**, 4273–4293 (2009).
- [11] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, Springer (2009).