Simple solution of the critical Kauffman model with connectivity one

T. M. A. Fink and F. C. Sheldon
London Institute for Mathematical Sciences, Royal Institution, 21 Albemarle St, London W1S 4BS, UK
(Dated: May 8, 2023)

The Kauffman model is a model of genetic computation that highlights the importance of criticality at the border of order and chaos. But our understanding of its behavior is incomplete, and much of what we do know relies on intricate arguments. We give a simple proof that the number of attractors for the critical Kauffman model with connectivity one grows faster than previously believed. Our approach relies on a link between the critical dynamics and number theory.

The Kauffman model is a simplified model of genetic networks that has been widely studied [1–6]. It highlights the importance of criticality—the border between frozen and chaotic dynamics—at which many biological systems seem poised [7, 8].

The Kauffman model with connectivity one plays a special role because it is exactly solvable [1]. At the same time, new approaches to solving it reveal additional insights [2], suggesting techniques for more realistic models which cannot be solved exactly. If, as is widely believed [2], all critical Kauffman models behave in a similar way, our results for the connectivity one model should apply to other critical versions, too.

In a Kauffman model with connectivity one, N nodes form a random directed network such that each node has one input, but any number of outputs. Thus the network is composed of loops and trees branching off the loops. Because the nodes in the trees are slaves to the loops, they do not contribute to the number or length of attractors, which are set solely by the m nodes in the loops. Each node is randomly assigned one of four Boolean functions: on, off, copy or invert.

In a critical Kauffman model, a perturbation to one node propagates to, on average, one other node. So in the critical model with connectivity one, all of the Boolean functions must be copy or invert [2]. Because of this, all of the 2^m states of the m nodes in loops are part of cycles; there are no transients.

We use the words cycle and attractor interchangeably, and we use the word dynamics to refer to the number and length of cycles in the state space. For a loop of size l, the dynamics depends only on the parity of the number of inverts in the loop. If the number of inverts is even, it is called an even loop, and it has cycles of length k if k divides l. If the number of inverts is odd, it is called an odd loop, and it has cycles of length k if k divides $2l$ but not l. We use the shorthand $\{l\}$ and $\{\bar{l}\}$ for even and odd loops of length l. Let $g \circ x$ denote g copies of a cycle of length x. We can write the dynamics as $d = g_1 \circ x_1 + g_2 \circ x_2 + \ldots$, where "+" means "and". Table I shows some examples.

Multiple loops can give rise to more complicated dynamics, where the cycle lengths of the network are the least common multiples of the cycle lengths of individual loops. We denote a collection of loops by $\{l_1, l_2, \ldots\}$, where $m = l_1 + l_2 + \ldots$ is the number of nodes in loops. The dynamics of multiple loops can be deduced from the dynamics of individual loops by defining a product between them, first introduced in [3]:

$$(g_1 \circ x_1 + g_2 \circ x_2 + \ldots)(h_1 \circ y_1 + h_2 \circ y_2 + \ldots) = \sum_{i,j} g_i h_j \ \gcd(x_i, y_j) \circ \operatorname{lcm}(x_i, y_j). \quad (1)$$

Examples are given in the right of Table I.

A lower bound on the number of attractors can be got from an upper bound on the mean attractor length. Writing $d(l_1, l_2, \ldots) = \nu_1 \circ A_1 + \nu_2 \circ A_2 + \ldots$, where there are ν_i cycles of length A_i, the mean attractor length is

$$\overline{A} = \sum_i \nu_i A_i \Big/ \sum_i \nu_i.$$

Since all 2^m states of the loop nodes belong to attractors, $\sum_i \nu_i A_i = 2^m$. So we have

$$c(m) = \sum_i \nu_i = 2^m / \overline{A}. \quad (2)$$

We can bound the mean attractor length from above by calculating the largest attractor length. In the presence of odd loops, the largest attractor length is double the least common multiple of the individual loop sizes. Consider all ways of partitioning some number m into $l_1, l_2, \ldots$. For $m = 8$, some of the 22 partitions are shown in the right of Table I. What is the maximum value of the lcm of the partitions? This is precisely Landau's function $g(m)$. Its first ten values are 1, 2, 3, 4, 6, 6, 12, 15, 20, 30 (OEIS A000793 [9]). When m is the sum of the first s primes, for $s \in [1, 8]$ $g(m)$ is the product of the primes: $g(2) = 2$, $g(5) = 6$, $g(10) = 30$, and so on. But this is not true in general: for $s = 9$, $\operatorname{lcm}(2, 3, \ldots, 19, 23) < \operatorname{lcm}(9, 3, \ldots, 19, 16)$, where we replaced 2 and 23 with 9 and 16.

Landau proved $\ln g(m)$ asymptotically approaches $\sqrt{m \ln m}$. More recently, Massias [10] proved

$$g(m) \leq 2^{1.52\sqrt{m \ln m}},$$

with equality at $m = 1{,}319{,}766$. Therefore the mean attractor length $\overline{A}$ satisfies

$$\overline{A}(l_1, l_2, \ldots) < 2 \cdot 2^{1.52\sqrt{m \ln m}}. \quad (3)$$

Even loops	*Odd loops*	*Multiple even loops*
$d(1) = 2 \circ 1$	$d(\overline{1}) = 1 \circ 2$	$d(1,7) = 4 \circ 1 + 36 \circ 7$
$d(2) = 2 \circ 1 + 1 \circ 2$	$d(\overline{2}) = 1 \circ 4$	$d(2,6) = 4 \circ 1 + 6 \circ 2 + 4 \circ 3 + 38 \circ 6$
$d(3) = 2 \circ 1 + 2 \circ 3$	$d(\overline{3}) = 1 \circ 2 + 1 \circ 6$	$d(3,5) = 4 \circ 1 + 4 \circ 3 + 12 \circ 5 + 12 \circ 15$
$d(4) = 2 \circ 1 + 1 \circ 2 + 3 \circ 4$	$d(\overline{4}) = 2 \circ 8$	$d(4,4) = 4 \circ 1 + 6 \circ 2 + 60 \circ 4$
$d(5) = 2 \circ 1 + 6 \circ 5$	$d(\overline{5}) = 1 \circ 2 + 3 \circ 10$	$d(2,3,3) = 8 \circ 1 + 4 \circ 2 + 40 \circ 3 + 20 \circ 6$
$d(6) = 2 \circ 1 + 1 \circ 2 + 2 \circ 3 + 9 \circ 6$	$d(\overline{6}) = 1 \circ 4 + 5 \circ 12$	$d(2,2,4) = 8 \circ 1 + 28 \circ 2 + 48 \circ 4$
$d(7) = 2 \circ 1 + 18 \circ 7$	$d(\overline{7}) = 1 \circ 2 + 9 \circ 14$	$d(2,2,2,2) = 16 \circ 1 + 120 \circ 2$
$d(8) = 2 \circ 1 + 1 \circ 2 + 3 \circ 4 + 30 \circ 8$	$d(\overline{8}) = 16 \circ 16$	$d(1,1,1,1,1,1,1,1) = 256 \circ 1$

TABLE I: **Number and length of cycles for single and multiple loops.** For example, $d(3)$, the dynamics of a loop of size 3, reads as 2 cycles of length 1 and 2 cycles of length 3. The cycle lengths of multiple loops are the least common multiples of those of the individual loops.

We can now write down our lower bound on the number of attractors. Inserting (3) into eq. (2) gives

$$c(m) > 2^{m-1.52\sqrt{m\ln m}}/2. \tag{4}$$

We can re-express this in terms of the total number of nodes N, where m is now a random variable. The mean number of nodes in loops $\overline{m}$ is asymptotically $\sqrt{\frac{\pi}{2}N}$ [1, 2]. Since eq. (4) is convex, by Jensen's inequality we can replace m with its mean, giving

$$c(N) > 2^{1.25\sqrt{N}-\sqrt[4]{N}\sqrt{1.45\ln N+0.65}-1}.$$

Despite the simplicity of our calculation, our bounds are a marked improvement on the best known bounds for the critical Kauffman model with connectivity one. Our upper bound for the mean attractor length $\overline{A}$ is less than $2^{0.5m}$ in [1], $2^{0.53m+1}$ in [2], and $2^{\sqrt{2m}\log_2\sqrt{2m}}$ in [3]. These are plotted in the top of Fig. 1. Similarly, our lower bound for the number of attractors c is greater than $2^{0.63\sqrt{N}}$ in [1], $2^{0.59\sqrt{N}}$ in [2], and $2^{1.25\sqrt{N}-\sqrt[4]{N}(0.57\ln N+1.05)}$ in [3]—the last of which was the result of a much lengthier calculation. These are plotted in the bottom of Fig. 1.

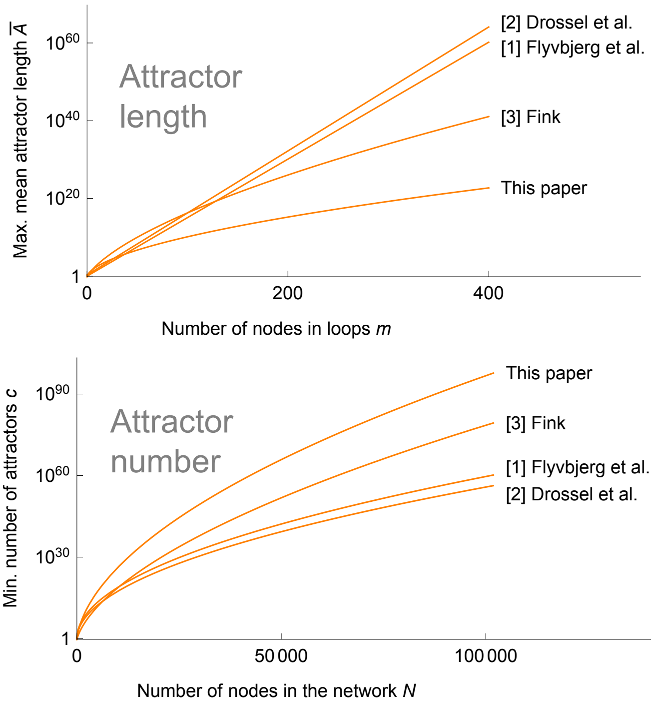

FIG. 1: **Top.** We plot the best known upper bounds for the mean attractor length $\overline{A}$. **Bottom.** We plot the best known lower bounds for the number of attractors c. In both cases our results are a marked improvement.

Simplified proofs are useful because they boost our confidence in the conclusion and improve our intuition for why it is so. Often, they motivate new lines of attack, leading to further insight. Our results suggest that the scaling of the number of attractors for the critical Kauffman model remains unsettled, despite the series of advances made in [1, 2, 5] and elsewhere. Our further analysis, not included here, leads us to conjecture that the scaling is in fact exponential in N, and we hope others will extend our results.

[1] H. Flyvbjerg, N. Kjaer, Exact solution of Kauffman's model with connectivity one, J Phys A **21**, 1695 (1988).

[2] B. Drossel, T. Mihaljev, F. Greil, Number and length of attractors in a critical Kauffman model with connectivity one, Phys Rev Lett **94**, 088701 (2005).

[3] T. Fink, Exact dynamics of the critical Kauffman model with connectivity one, arXiv:2302.05314 (2023).

[4] J. Socolar, S. Kauffman, Scaling in ordered and critical random Boolean networks, Phys Rev Lett **90**, 068702 (2003).

[5] B. Samuelsson, C. Troein, Superpolynomial growth in the number of attractors in Kauffman networks, Phys Rev Lett **90**, 098701 (2003).

[6] T. Peixoto, Redundancy and error resilience in Boolean networks, Phys Rev Lett **104**, 048701 (2010).

[7] M. Muñoz, Criticality and dynamical scaling in living systems, Rev Mod Phys **90**, 031001 (2018).

[8] B. Daniels et al., Criticality distinguishes the ensemble of biological regulatory networks, Phys Rev Lett **121**, 138102 (2018).

[9] N. Sloane, editor, On-Line Encyclopedia of Integer Sequences, https://oeis.org, 2023.

[10] J.-P. Massias, Majoration explicite de l'ordre maximum d'un élément du groupe symétrique, Ann Fac Sci Toulouse Math **6**, 269 (1984).