A Heuristic Approach to P ≠ NP Based on the Hyperuniformity of Langford Sequences

Wino Research*

Published on 2025-09-02; revised on 2025-09-19

*sci@cosmoduli.com

Introduction to Langford's Problem

Arrange m sets of numbers 1 to n in a sequence, so that any two consecutive occurrences of k are separated by exactly k numbers [1–5].

Let $L(m, n)$ denote the number of distinct Langford sequences up to a reversal symmetry. We have $L(2, 3) = L(2, 4) = 1$ and $L(3, 9) = 3$:

```
3 1 2 1 3 2
4 1 3 1 2 4 3 2
1 9 1 6 1 8 2 5 7 2 6 9 2 5 8 4 7 6 3 5 4 9 3 8 7 4 3
1 9 1 2 1 8 2 4 6 2 7 9 4 5 8 6 3 4 7 5 3 9 6 8 3 5 7
1 8 1 9 1 5 2 6 7 2 8 5 2 9 6 4 7 5 3 8 4 6 3 9 7 4 3
```

Asymptotic Formulas for Counting Langford Sequences

Conjecture 1. The number of Langford sequences $L(m, n)$ has the following asymptotic formula [6]

$$L(m, n) \sim n! e^{-\ell(m)n},$$

where $\ell(m)$ is an exponential coefficient depending only on m, *i.e.*

$$\ell(m, n) = \frac{1}{n} \log \frac{n!}{L(m, n)}$$

converges to a constant when $n \to \infty$.

Conjecture 2. The exponential coefficient $\ell(2,n)$ for the number of Langford sequences $L(2,n)$ converges to Taniguchi's constant

$$\lim_{n\to\infty} \ell(2,n) = \prod_{p\in\mathbb{P}} \left(1 - \frac{3}{p^3} + \frac{2}{p^4} + \frac{1}{p^5} - \frac{1}{p^6}\right) = 0.678234491\cdots,$$

where the product runs over the primes $\mathbb{P}$. More precisely, the specific value can be approximated as

$$\ell(2,n) \simeq \prod_{i=1}^{N} \left(1 - \frac{3}{p_i^3} + \frac{2}{p_i^4} + \frac{1}{p_i^5} - \frac{1}{p_i^6}\right).$$

Table 1. Number of Langford sequences $L(2, n)$, OEIS A014552. In Ref. [7], the approximate values $L(2, 31) \simeq 5.381 \cdot 10^{24}$ and $L(2, 32) \simeq 8.812 \cdot 10^{25}$ are obtained using a parallel tempering algorithm.

n	exact		approximate		error
	$L(2,n)$	$\ell(2,n)$	$\ell(2,n)$	$L(2,n)$	
3	1	0.597253	0.765625	1	$\sim 0\%$
4	1	0.794513		1	$\sim 0\%$
7	26	0.752438		24	-7.7%
8	150	0.699246	0.701560	147	-2.0%
11	17792	0.701437		$1.777 \cdot 10^4$	-0.1%
12	108144	0.699666		$1.057 \cdot 10^5$	-2.2%

15	39809640	0.693310	0.687148	$4.367 \cdot 10^{7}$	+9.7%
16	326721800	0.691703		$3.514 \cdot 10^{8}$	+7.6%
19	256814891280	0.687803		$2.600 \cdot 10^{11}$	+1.3%
20	2636337861200	0.686760		$2.616 \cdot 10^{12}$	−0.8%
23	3799455942515488	0.684045	0.681745	$4.006 \cdot 10^{15}$	+5.4%
24	46845158056515936	0.683296		$4.862 \cdot 10^{16}$	+3.8%
27	111683611098764903232	0.681309		$1.104 \cdot 10^{20}$	−1.2%
28	1607383260609382393152	0.680745	0.680305	$1.627 \cdot 10^{21}$	+1.2%
31				$5.701 \cdot 10^{24}$	
32				$9.240 \cdot 10^{25}$	
35			0.679426	$4.861 \cdot 10^{29}$	
36				$8.871 \cdot 10^{30}$	

Relations between Permutations and Langford Sequences

For any permutation $\sigma(n)$ of the set $\{1, 2, ..., n\}$, we know that its Lehmer code forms a factoradic number x, which can be used to index a permutation in the lexicographic ordering.

Since every Langford sequence can be represented as a permutation, such as `1 4 1 5 6 7 4 2 3 5 2 6 3 7` can be represented as $(1, 4, 5, 6, 7, 2, 3)$, an intriguing question arises: **how many permutations on the index interval $[x, x + d)$ correspond to Langford sequences?**

The Hyperuniformity of Langford Sequences

For the Langford pairing problem $\mathbb{L}(2, n)$, the pairing ratio $r(n, d)$ is defined as

$$r(n, d) = \frac{n!\mu(n, d)}{2L(2, n)d},$$

where $\mu(n, d)$ is the sampling mean of the number of Langford sequences on the index interval $[x, x + d)$. From Table 2, we can see that

$$r(n, d) \simeq 1 - \frac{1}{d} \quad \Rightarrow \quad \mu(n, d) \simeq 2(d - 1)e^{-\ell(2,n)n}.$$

Table 2. Pairing ratios $r(n,d)$ for different interval lengths.

$r(n,d)$	10^7 samples		10^8 samples		10^9 samples	
	$n=11$	$n=12$	$n=15$	$n=16$	$n=19$	$n=20$
$d=2$	0.497950	0.504275	0.504712	0.505264	0.504812	0.494178
$d=5$	0.794028	0.802189	0.795189	0.798817	0.799836	0.811909
$d=10$	0.898364	0.901406	0.894949	0.895355	0.897554	0.908530
$d=20$	0.948910	0.950527	0.952901	0.951805	0.953246	0.943667
$d=50$	0.979904	0.982732	0.983297	0.979322	0.976178	0.980197
$d=100$	0.990496	0.991568	0.987725	0.986786	0.988295	0.989223
$d=200$	0.995367	0.994959	0.993111	0.995348	0.995793	0.992925
$d=500$	0.998519	0.998965	0.998034	1.001088	0.998281	0.997641
$d=1000$	0.999376	0.999480	0.998831	1.000253	0.999208	0.999534

Referring to concepts in physics, we introduce global and local density as

$$\rho_g(n) = \frac{2L(2,n)}{n!}, \quad \rho_l(n,d) = \frac{\mu(n,d)}{d}.$$

Then, we have $r(n,d) = \rho_l(n,d)/\rho_g(n)$. Similarly, density fluctuations are defined as

$$s^2(n,d) = \frac{1}{N-1} \sum_{\text{samples}} \left[\rho_l(n,d) - \rho_g(n)\right]^2.$$

From Table 3, we can see that ρ_l is nearly constant and s^2 decays as d increases, which are main characteristics of hyperuniform systems.

Table 3. Local densities $\rho_l(n,d)$ and density fluctuations $s^2(n,d)$ for different interval lengths. The number of samples $N = 10^8$ is used for all calculations.

$\rho_l(n,d)$ / $s^2(n,d)$	$n=11$	$n=12$	$n=15$	$n=16$	$n=19$	$n=20$
$d=2$	$4.442\cdot10^{-4}$	$2.254\cdot10^{-4}$	$2.965\cdot10^{-5}$	$1.586\cdot10^{-5}$	$2.250\cdot10^{-6}$	$1.110\cdot10^{-6}$
	$2.221\cdot10^{-4}$	$1.127\cdot10^{-4}$	$1.482\cdot10^{-5}$	$7.927\cdot10^{-6}$	$1.125\cdot10^{-6}$	$5.550\cdot10^{-7}$
$d=5$	$7.150\cdot10^{-4}$	$3.623\cdot10^{-4}$	$4.867\cdot10^{-5}$	$2.497\cdot10^{-5}$	$3.286\cdot10^{-6}$	$1.690\cdot10^{-6}$
	$1.483\cdot10^{-4}$	$7.533\cdot10^{-5}$	$1.008\cdot10^{-5}$	$5.172\cdot10^{-6}$	$6.788\cdot10^{-7}$	$3.476\cdot10^{-7}$
$d=10$	$8.026\cdot10^{-4}$	$4.065\cdot10^{-4}$	$5.480\cdot10^{-5}$	$2.803\cdot10^{-5}$	$3.791\cdot10^{-6}$	$1.926\cdot10^{-6}$
	$8.740\cdot10^{-5}$	$4.418\cdot10^{-5}$	$5.932\cdot10^{-6}$	$3.021\cdot10^{-6}$	$4.089\cdot10^{-7}$	$2.046\cdot10^{-7}$

$d = 20$	$8.478 \cdot 10^{-4}$	$4.281 \cdot 10^{-4}$	$5.735 \cdot 10^{-5}$	$2.958 \cdot 10^{-5}$	$4.004 \cdot 10^{-6}$	$2.052 \cdot 10^{-6}$
	$4.829 \cdot 10^{-5}$	$2.435 \cdot 10^{-5}$	$3.238 \cdot 10^{-6}$	$1.666 \cdot 10^{-6}$	$2.236 \cdot 10^{-7}$	$1.128 \cdot 10^{-7}$
$d = 50$	$8.732 \cdot 10^{-4}$	$4.431 \cdot 10^{-4}$	$5.961 \cdot 10^{-5}$	$3.064 \cdot 10^{-5}$	$4.097 \cdot 10^{-6}$	$2.123 \cdot 10^{-6}$
	$2.258 \cdot 10^{-5}$	$1.130 \cdot 10^{-5}$	$1.486 \cdot 10^{-6}$	$7.560 \cdot 10^{-7}$	$9.892 \cdot 10^{-8}$	$5.089 \cdot 10^{-8}$
$d = 100$	$8.823 \cdot 10^{-4}$	$4.470 \cdot 10^{-4}$	$6.028 \cdot 10^{-5}$	$3.102 \cdot 10^{-5}$	$4.160 \cdot 10^{-6}$	$2.138 \cdot 10^{-6}$
	$1.257 \cdot 10^{-5}$	$6.221 \cdot 10^{-6}$	$8.075 \cdot 10^{-7}$	$4.122 \cdot 10^{-7}$	$5.359 \cdot 10^{-8}$	$2.739 \cdot 10^{-8}$
$d = 200$	$8.875 \cdot 10^{-4}$	$4.494 \cdot 10^{-4}$	$6.041 \cdot 10^{-5}$	$3.117 \cdot 10^{-5}$	$4.162 \cdot 10^{-6}$	$2.161 \cdot 10^{-6}$
	$7.578 \cdot 10^{-6}$	$3.694 \cdot 10^{-6}$	$4.576 \cdot 10^{-7}$	$2.330 \cdot 10^{-7}$	$2.982 \cdot 10^{-8}$	$1.539 \cdot 10^{-8}$
$d = 500$	$8.901 \cdot 10^{-4}$	$4.507 \cdot 10^{-4}$	$6.080 \cdot 10^{-5}$	$3.116 \cdot 10^{-5}$	$4.199 \cdot 10^{-6}$	$2.160 \cdot 10^{-6}$
	$3.857 \cdot 10^{-6}$	$1.852 \cdot 10^{-6}$	$2.196 \cdot 10^{-7}$	$1.091 \cdot 10^{-7}$	$1.392 \cdot 10^{-8}$	$7.056 \cdot 10^{-9}$
$d = 1000$	$8.906 \cdot 10^{-4}$	$4.512 \cdot 10^{-4}$	$6.078 \cdot 10^{-5}$	$3.118 \cdot 10^{-5}$	$4.216 \cdot 10^{-6}$	$2.171 \cdot 10^{-6}$
	$2.489 \cdot 10^{-6}$	$1.190 \cdot 10^{-6}$	$1.348 \cdot 10^{-7}$	$6.595 \cdot 10^{-8}$	$8.166 \cdot 10^{-9}$	$4.100 \cdot 10^{-9}$

We then consider the number of Langford sequences whose indices x belonging to distinct residue classes modulo k:

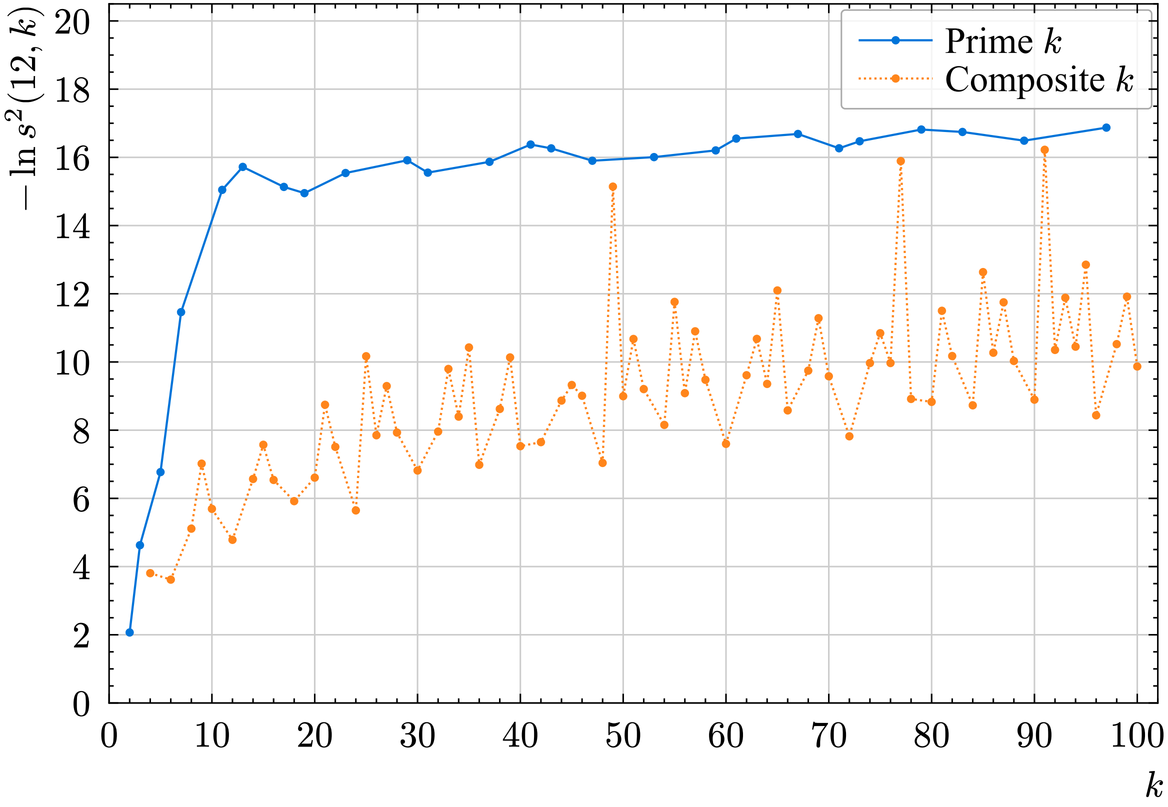
$$c(n, k, a) = \#\{x \in \mathbb{L}(2, n) : x \equiv a \bmod k\}.$$

The variance across different moduli is defined as

$$s^2(n, k) = \frac{1}{k-1} \sum_{a=0}^{k-1} \left[\frac{c(n, k, a)}{2L(2, n)} - \frac{1}{k} \right]^2.$$

Interestingly, we find that the uniformity of distribution across different residue classes of indices is closely related to the number of prime factors of modulo k. See Figure 1 for an example of $n = 12$.

Figure 1. The negative log variance across different moduli k.



A Heuristic Approach to P ≠ NP

For any given index x and a fixed length d, find a permutation $\sigma(n)$ on the interval $[x, x+d)$ which provides a solution to the Langford pairing problem $\mathbb{L}(2, n)$.

When $d > 1 + \frac{1}{2}e^{\ell(2,n)n}$ and $n \to \infty$, we can expect that there is at least one solution on the interval $[x, x+d)$ regardless of x. It is easy to verify the solution, but there is no efficient algorithm to find it due to the hyperuniformity of Langford sequences, unless we exhaustively traverse all $O\left(e^{\ell n}\right)$ permutations on the interval. Therefore, we can conclude that **P ≠ NP**.

Bibliography

1. Langford CD. Problem 228. The Mathematical Gazette. 1958;42.
2. Miller JE. Langford's Problem. Available: https://dialectrix.com/langford.html
3. Walsh T. CSPLib Problem 024: Langford's Number Problem. Jefferson C, Miguel I, Hnich B, Walsh T, Gent IP, editors. Available: https://www.csplib.org/Problems/prob024

4. Krajecki M, Loiseau J, Alin F, Jaillet C. Many-Core Approaches to Combinatorial Problems: case of the Langford Problem. Supercomputing Frontiers and Innovations. 2016;3: 21–37. doi:10.14529/jsfi160202

5. Akgün Ö, Miguel I. Modelling Langford's Problem: A Viewpoint for Search. 2018. doi:10.48550/arXiv.1808.09847

6. Pan Z. Conjectures on the Number of Langford Sequences. 2021. doi:10.5281/zenodo.13824675

7. Assarpour A, Barnoy A, Liu O. Counting Skolem Sequences. 2015. doi:10.48550/arXiv.1507.00315

Appendix: Commands in Wino Studio

- `oeis.langford.count_pairings(n, start, end)`
- `oeis.langford.find_pairings(n, start, end, count)`
- `oeis.langford.pairing_moduli(n, start, end, m)`
- `oeis.langford.estimate_pairings(n, d, samples)`
- `oeis.langford.pairing_density(n, d, samples)`
- `oeis.langford.pairing_ratio(n, d, samples)`

Examples

```
 Input: oeis.langford.count_pairings(12, 1000000, 2000000)
Output: 622
```