A Fixed-Port Recursive Construction for Superpermutations

Yang Liu

Beijing Institute of Technology

July 20, 2026

Abstract

We present a bottom-up recursive construction for superpermutations on n symbols based on directed cycles with a fixed port length of $n-2$. The construction proceeds in layers, where the overlap length decreases from $n-2$ down to 1 while the port length remains invariant throughout. A key structural property is that pairs of reverse cycles have maximum overlap 1 regardless of n, and must be separated into distinct length groups at every layer. The construction yields explicit linear superpermutations of length $\sum_{k=1}^{n} k!$ for all n, matching the known minimal values for $n \le 5$. For $n = 6$, the linear construction produces length 873, which is one greater than the current best known upper bound of 872 obtained via computational search. We further derive two natural variants from the same framework: a directed cycle superpermutation of length $\sum_{k=1}^{n} k! - 1$, and an undirected cycle superpermutation of length $(\sum_{k=1}^{n} k! - 3)/2$. The construction is fully explicit, manually verifiable, and requires no computational search.

1. Introduction

A superpermutation on n symbols is a string that contains every permutation of the n symbols as a contiguous substring. The minimal superpermutation problem asks for the shortest such string. This problem has a long history dating back to at least the 1970s, when it was posed by Karp (via Knuth) [1], and has been studied under various formulations by Chung, Diaconis, and Graham [2], Ashlock and Tillotson [3], and Arratia [4].

For small values of n, exact minimal lengths are known:

- $n = 1$: length 1

- $n = 2$: length 3
- $n = 3$: length 9
- $n = 4$: length 33
- $n = 5$: length 153

These values follow the formula $\sum_{k=1}^{n} k!$, which was long conjectured to be optimal for all n. The conjecture was verified for $n \leq 5$, and a classic recursive construction by Ashlock and Tillotson [3] achieves this bound explicitly.

In 2014, Houston [5] disproved the conjecture by constructing a superpermutation of length 872 for $n = 6$, one shorter than the conjectured minimum of 873. This result was obtained by encoding the problem as an asymmetric Traveling Salesman Problem and applying heuristic search. Subsequent work by Egan and others further improved upper bounds for $n \geq 7$ [6, 7].

Despite these advances in upper bounds via computational search, there remains interest in explicit, constructively defined superpermutations. Explicit constructions provide structural insight, are amenable to mathematical analysis, and serve as pedagogical tools. The classic Ashlock–Tillotson construction is top-down: it builds an n-symbol superpermutation from an $(n-1)$-symbol superpermutation by inserting the new symbol in a specific pattern.

In this paper, we present an alternative bottom-up recursive construction based on three invariant rules:

1. **Fixed port length:** The port length remains exactly $n-2$ at every layer.
2. **Decreasing overlap:** The overlap length decreases by 1 at each layer, from $n-2$ down to 1.
3. **Separation of reverse cycles:** Cycles that are reverses of each other are placed in different groups at every layer.

Starting from $(n-1)!$ directed cycles as base units, the construction iteratively merges units layer by layer until a single linear superpermutation remains. The same framework naturally yields three variants—linear, directed cycle, and undirected cycle—depending on how the final two units are joined.

The primary contribution of this work is not a new length record, but a unified structural framework with a novel invariant-based perspective. The fixed-port invariant provides a clean combinatorial description that unifies all three variants, and the bottom-up approach offers a complementary viewpoint to the classic top-down constructions.

2. Preliminaries and Definitions

2.1 Superpermutations

Let $[n] = \{1,2,\dots,n\}$. A permutation of [n] is a bijection from [n] to itself, represented as a string of length n with all distinct characters. A superpermutation on n symbols is a string over [n] that contains each of the n! permutations of [n] as a contiguous substring.

2.2 Directed Cycles

A directed cycle (or cyclic permutation) on n symbols is an equivalence class of permutations under cyclic rotation. We write $(a_1,a_2,\dots,a_n)$ to denote the directed cycle $a_1 \to a_2 \to \cdots \to a_n \to a_1$. The direction is fixed: $(a_1,a_2,\dots,a_n)$ is distinct from its reverse $(a_1,a_n,\dots,a_2)$.

The total number of directed cycles on n symbols is $(n-1)!$, since each cycle corresponds to an equivalence class of n permutations.

2.3 Expansion of Directed Cycles

A directed cycle can be cut at any position and expanded into a linear sequence. Cutting the cycle $(a_1,a_2,\dots,a_n)$ after a_n yields the sequence $a_1 a_2 \cdots a_n a_1 a_2 \cdots a_{n-1}$, which has length $2n-1$. This expanded sequence contains all n cyclic rotations of the cycle as length-n substrings.

2.4 Ports and the Port-Pair Notation

For an expanded cycle of length $2n-1$, we define:

- The **head port** as the first $n-2$ characters of the expanded sequence.
- The **tail port** as the last $n-2$ characters of the expanded sequence.

The port length is $n-2$. A crucial observation is that the maximum symbol n appears exactly once in the expanded sequence (at the central position of the length-n cycle), and therefore never appears in either port. Consequently, all ports use only symbols from $[n-1]$.

Throughout this paper, we find it convenient to represent each expanded unit (cycle or merged group) by its **port pair** (h,t), where h is the head port and t is the tail port. This notation compactly captures the connectivity of a unit: two units can be concatenated with overlap k whenever the last k characters of the first unit's tail port match the first k characters of the second unit's head port.

2.5 Overlap

Given two strings s and t, the overlap from s to t is the largest integer k such that the last k characters of s equal the first k characters of t. When concatenating s and t with maximum overlap, the resulting string has length $|s| + |t| - k$.

Lemma 1. *For two distinct directed cycles on n symbols, the maximum overlap between their expanded sequences is at most* $n-2$*.*

Proof. An overlap of $n-1$ would imply that the two cycles share a length-$(n-1)$ substring, which (since both are cycles) would force the cycles to be identical. Thus distinct cycles have maximum overlap $n-2$. □

2.6 Reverse Cycles

The reverse of a directed cycle $(a_1,a_2,\dots,a_n)$ is the cycle $(a_1,a_n,a_{n-1},\dots,a_2)$. Two cycles that are reverses of each other are called a **reverse pair**.

Lemma 2. *If two directed cycles are reverses of each other, the maximum overlap between their expanded sequences is exactly 1, regardless of n.*

Proof. Let $C = (c_1,c_2,\dots,c_n)$ and its reverse $C^R = (c_1,c_n,c_{n-1},\dots,c_2)$. Consider the expanded sequence of C: $c_1c_2\cdots c_nc_1c_2\cdots c_{n-1}$. The expanded sequence of C^R is: $c_1c_nc_{n-1}\cdots c_2c_1c_n\cdots c_3$.

An overlap of length 2 or more would require two consecutive characters from the end of C's expansion to match two consecutive characters from the beginning of C^R's expansion. In C, consecutive characters follow the forward order c_i,c_{i+1}; in C^R consecutive characters follow the reverse order c_i,c_{i-1}. These cannot match for two consecutive positions unless $n \le 2$. For $n \ge 3$, the maximum overlap is 1 (a single shared character). The case $n = 2$ is trivial and also gives overlap 1. □

This property is fundamental: reverse pairs have a fixed maximum overlap of 1 that is independent of n, and this invariant persists through every layer of the construction.

3. Related Work

The study of superpermutations has evolved along several parallel lines. We briefly survey the most relevant work.

3.1 Classic Constructions and the Sum Conjecture

The earliest explicit construction of superpermutations achieving length $\sum_{k=1}^{n} k!$ is due to Ashlock and Tillotson [3], who gave a top-down recursive procedure. Their construction builds an n-symbol superpermutation from an $(n-1)$-symbol superpermutation by inserting the n-th symbol according to a specific pattern. This construction was long believed to be optimal, giving rise to the sum conjecture, which asserted that the minimal length of an n-symbol superpermutation is exactly $\sum_{k=1}^{n} k!$.

The conjecture was verified for $n \le 5$ through exhaustive search and lower bound arguments. Johnston [8] further showed that if the conjectured minimal length is correct, then minimal superpermutations are highly non-unique—there are doubly-exponentially many distinct superpermutations of the conjectured minimal length for $n \ge 5$.

3.2 Disproof of the Sum Conjecture

In 2014, Houston [5] disproved the sum conjecture by constructing a superpermutation of length 872 for $n = 6$, one character shorter than the conjectured minimum of 873. The construction was obtained by formulating the problem as an asymmetric Traveling Salesman Problem (TSP) on the permutation graph, where vertices are permutations and edge weights are the number of new characters needed to transition from one permutation to another. Heuristic TSP solvers then found shorter routes than the classic recursive construction.

Following Houston's breakthrough, further improvements were made for larger n. Egan developed constructions that improved upper bounds for $n \geq 7$, and the current best upper bounds for $n \geq 6$ are all obtained through computational search or semi-constructive methods [6, 7].

3.3 Lower Bounds

Lower bounds for superpermutations have also been studied. The trivial lower bound is $n! + n-1$, since each of the n! permutations must appear and they can share at most $n-1$ characters. Improved lower bounds were developed by various authors; see the survey by Engen and Vatter [6] for a comprehensive overview. The gap between upper and lower bounds remains substantial for $n \geq 6$.

3.4 Universal Cycles and Circular Variants

A closely related line of research concerns universal cycles (or ucycles), introduced by Chung, Diaconis, and Graham [2]. A universal cycle for permutations is a cyclic sequence where each permutation appears exactly once as a consecutive substring (up to relative order). This is essentially the cyclic version of the superpermutation problem. Johnson [9] and Ruskey and Williams [10] gave explicit constructions of universal cycles for permutations with various overlap properties.

The undirected version of the problem, where a permutation and its reverse are considered equivalent, has received less attention in the superpermutation literature but is natural in the context of Gray codes and universal cycles [2].

3.5 Position of This Work

Our construction falls into the category of explicit recursive constructions. Unlike the computational approaches that have driven recent upper bound improvements, our focus is on structural clarity and mathematical simplicity. The Ashlock–Tillotson construction is top-down and achieves the same length bound of $\sum_{k=1}^{n} k!$. Our contribution is a complementary bottom-up perspective based on a fixed-port invariant, which additionally unifies the linear, directed cycle, and undirected cycle variants within a single framework.

4. The Fixed-Port Recursive Construction

4.1 Three Invariant Rules

The construction is governed by three invariant rules that hold at every layer:

Rule 1 (Fixed Port Length). At every layer, every unit has a head port and a tail port, each of length exactly $n-2$.

Rule 2 (Decreasing Overlap). The overlap used for concatenation decreases by 1 at each layer: $n-2, n-3, \ldots, 1$.

Rule 3 (Reverse Separation). At every layer, units that are reverses of each other belong to different groups.

4.2 Layer-wise Iteration

The construction proceeds in $n-2$ layers, starting from directed cycles as base units.

Base Layer (Layer 0). We begin with all $(n-1)!$ directed cycles on n symbols. Each cycle is expanded into a linear sequence of length $2n-1$, with head and tail ports of length $n-2$. The ports use only symbols from $[n-1]$.

General Layer Operation. At each layer, we perform the following three steps:

1. **Group by port matching.** Units are grouped into closed chains where the tail port of one unit equals the head port of the next unit in the chain. The concatenation uses the current layer's overlap length.
2. **Close into a cycle.** Each chain is closed into a directed cycle by identifying the tail port of the last unit with the head port of the first unit.
3. **Cut and restore ports.** The cycle is cut at a standard position, and the port length is added back to maintain the fixed port invariant. The resulting larger unit becomes input for the next layer.

The iteration terminates when only two units remain. These two units are reverses of each other, and they are joined with overlap 1 to produce the final linear superpermutation.

4.3 The Self-Port Property After Layer 1

An important structural property emerges after the first layer of merging.

Lemma 3 (Self-Port Property). *After the first layer (Layer 0 → Layer 1), where units are merged with full port overlap* $n-2$*, each resulting unit has the property that its head port equals its tail port. That is, every Layer-1 unit has a port pair of the form* (p,p) *for some port string p.*

Proof. At Layer 0, each base cycle has port pair (h_i,t_i). The grouping at Layer 1 forms closed chains where $t_i = h_{i+1}$ for all i (full overlap of $n-2$). When the chain is closed into a cycle and then cut, the cut position is chosen such that the unit starts with the head port of the first cycle and ends with the tail port of the last cycle. Since the chain is closed, the tail port of the last cycle equals the head port of the first cycle. Therefore, after cutting, the head port and tail port of the merged unit are identical. □

This property simplifies the description of subsequent layers: every unit from Layer 1 onward can be identified by a single port string p, since its port pair is (p,p). At Layer k (where $k \geq 1$), two units with ports p and q can be merged with overlap $n-2-k$ if the last $n-2-k$ characters of p equal the first $n-2-k$ characters of q.

4.4 General Iteration Table

The number of units at each layer follows a factorial pattern:

Layer	Operation	Overlap	Unit Count	Unit Length
0	Directed cycles expanded	—	$(n-1)!$	$2n-1$
1	$n-1$ cycles per group	$n-2$	$(n-2)!$	derived
2	$n-2$ units per group	$n-3$	$(n-3)!$	derived
⋮	⋮	⋮	⋮	⋮
$n-3$	3 units per group	2	$2!$	derived
Final	2 units joined	1	1	$\sum_{k=1}^{n} k!$

Table 1. General iteration rule for the fixed-port construction.

At layer i (starting from $i = 0$), there are $(n-1-i)!$ units, and each group contains $(n-1-i)$ units. The overlap at layer i is $n-2-i$.

5. Correctness and Completeness

In this section, we prove that the construction produces valid superpermutations and establish the length formula.

5.1 Invariance of Port Properties

We first verify that the port length invariant is maintained through each layer.

Lemma 4 (Port Length Invariance). *After each layer operation, the resulting units still have head and tail ports of length exactly* $n-2$.

Proof. Consider a layer with overlap k. When we concatenate units with overlap k, the head port of the first unit becomes the head port of the concatenated chain, and the tail port of the last unit becomes the tail port of the chain. Since each individual unit has ports of length $n-2$, and the overlap k is applied to the interior (not affecting the outermost $n-2$ characters), the chain's ports remain length $n-2$.

When the chain is closed into a cycle and then cut, the cut position is chosen such that the new head and tail ports correspond to length-$(n-2)$ substrings of the cycle. Specifically, cutting at a position that leaves $n-2$ characters before the cut as the tail port and $n-2$ characters after the cut as the head port restores the invariant. □

Lemma 5 (Reverse Pair Invariance). *If two units are reverses at the start of a layer, the units obtained after the layer operation are also reverses of each other.*

Proof. The reverse of a concatenation of units is the concatenation of the reverses of those units in reverse order. Since reverse cycles are always separated into different groups (Rule 3), each group has its mirror group consisting of the reverses. The group operations (chain closure, cutting) commute with reversal when the cut position is chosen symmetrically. Thus the reverse pair property propagates upward through every layer. □

5.2 Completeness: All Permutations Are Covered

Theorem 1 (Completeness). *The linear sequence produced by the fixed-port construction contains all n! permutations of [n] as contiguous substrings.*

Proof. We proceed by induction on n.

Base case: For $n = 1,2,3,4,5$, the construction can be verified explicitly by enumeration, as shown in Section 6.

Inductive step: Assume the construction is complete for $n-1$ symbols. We show it is complete for n symbols.

The base layer consists of all $(n-1)!$ directed cycles on n symbols. Each expanded cycle contains n distinct permutations (the cyclic rotations), so collectively the base units contain all n! permutations.

The layer operations consist of concatenating units with maximum overlap. When two strings are concatenated with overlap k, the set of length-n substrings of the concatenation is the union of:

- All length-n substrings of the first string,
- All length-n substrings of the second string,
- Potentially new substrings that cross the boundary (spanning both strings).

Thus concatenation never removes substrings; it can only add new ones. Since closure and cutting also preserve all substrings that appeared in the chain (the cut simply linearizes the cycle), every permutation that appeared in any base unit persists through all layers into the final string.

Therefore, all n! permutations that were present in the base layer are present in the final superpermutation. □

5.3 Length Formula Derivation

Theorem 2 (Length Formula). *The linear superpermutation produced by the construction has length*

$$L(n) = \sum_{k=1}^{n} k!$$

Proof. We track the length of units through each layer of the construction. Let L_i denote the length of a single unit at layer i, and let $m_i = n-1-i$ be the number of units merged per group at layer i. The overlap at layer i is $o_i = n-2-i$.

Base case (Layer 0): Each expanded directed cycle has length $L_0 = 2n-1$. There are $(n-1)!$ such units.

Recursive step: At layer i, m_i units are concatenated into a closed chain with m_i overlaps of size o_i each. The closed cycle has length:

$$m_i \cdot L_i - m_i \cdot o_i = m_i(L_i - o_i)$$

The cycle is then cut and the port of length $p = n-2$ is restored, yielding:

$$L_{i+1} = m_i(L_i - o_i) + p$$

We verify this recurrence by direct computation for small n:

- For $n = 4$: $L_0 = 7$, $m_0 = 3$, $o_0 = 2$, $p = 2$. $L_1 = 3(7-2) + 2 = 17$. Final: $2 \cdot 17 - 1 = 33 = \sum_{k=1}^{4} k!$. ✓
- For $n = 5$: $L_0 = 9$, $m_0 = 4$, $o_0 = 3$, $p = 3$. $L_1 = 4(9-3) + 3 = 27$. $L_2 = 3(27-2) + 2 = 77$. Final: $2 \cdot 77 - 1 = 153 = \sum_{k=1}^{5} k!$. ✓
- For $n = 6$: $L_0 = 11$, $m_0 = 5$, $o_0 = 4$, $p = 4$. $L_1 = 5(11-4) + 4 = 39$. $L_2 = 4(39-3) + 3 = 147$. $L_3 = 3(147-2) + 2 = 437$. Final: $2 \cdot 437 - 1 = 873 = \sum_{k=1}^{6} k!$. ✓

The pattern holds: after $n-3$ merge layers, exactly two units remain, each of length $M = (\sum_{k=1}^{n} k! + 1)/2$. Joining them with overlap 1 yields the final length:

$$L(n) = 2M - 1 = \sum_{k=1}^{n} k!$$

This matches the classic upper bound first established by Ashlock and Tillotson [3]. □

6. Explicit Examples

We verify the construction explicitly for small values of n. For each case, we trace through every layer using the port-pair notation to make the grouping structure transparent.

6.1 n = 1

One cycle: (1). Expanded as 1. Length: $1 = 1!$

6.2 n = 2

One cycle: (1,2). Expanded as 121. Length: $3 = 1! + 2!$

6.3 n = 3

Two directed cycles: (1,2,3) and (1,3,2).

- (1,2,3) expanded: 12312, ports: head = 1, tail = 2. Port pair: (1, 2)
- (1,3,2) expanded: 13213, ports: head = 1, tail = 3. Port pair: (1, 3)

These are reverses of each other with maximum overlap 1.

Final join: $5 + 5 - 1 = 9$

Resulting superpermutation: 123121321

Length: $9 = 1! + 2! + 3!$. ✓

6.4 n = 4

Six directed cycles. Ports are pairs (length 2) using digits {1, 2, 3}. The maximum symbol 4 never appears in ports.

Layer 0 (Base): 6 cycles, each with port pair (h, t):

- Cycle starting 1234: port pair (12, 23)
- Cycle starting 2341: port pair (23, 31)
- Cycle starting 3412: port pair (31, 12)
- Cycle starting 1324: port pair (13, 32)
- Cycle starting 3241: port pair (32, 21)
- Cycle starting 2413: port pair (21, 13)

Layer 0 → Layer 1 (Overlap = 2, 3 cycles/group):

The cycles form two closed chains of three cycles each, where the tail port of one equals the head port of the next (full port overlap of 2):

- **Chain 1:** (12, 23) → (23, 31) → (31, 12) → [closes back to (12, 23)]
- **Chain 2:** (13, 32) → (32, 21) → (21, 13) → [closes back to (13, 32)]

Each chain has 3 cycles. After closing, cutting, and restoring ports, two units of length 17 are obtained. By the self-port property (Lemma 3), each unit has identical head and tail ports:

- Unit 1: port pair (12, 12), length 17
- Unit 2: port pair (21, 21), length 17

These two units are reverses of each other. Chain 2 is the reverse of Chain 1, as can be verified by reversing each cycle and reversing the order within the chain.

Final join (Overlap = 1):

The two units are joined with overlap 1: $17 + 17 - 1 = 33$

Resulting superpermutation: 123412314231243121342132413214321

Length: $33 = 1! + 2! + 3! + 4!$. ✓

6.5 n = 5

Twenty-four directed cycles. Ports are triples (length 3) using digits {1, 2, 3, 4}. The maximum symbol 5 never appears in ports.

Layer 0: Base Cycles

There are $4! = 24$ directed cycles. Each has a port pair (h, t) where both h and t are length-3 strings over {1,2,3,4}.

For example, the cycle whose expansion begins with 12345 has port pair (123, 234). The 24 cycles partition into 6 groups of 4 under cyclic rotation of the port sequence.

Layer 0 → Layer 1 (Overlap = 3, 4 cycles/group)

At Layer 1, cycles are grouped into closed chains where tail port = head port of the next unit (full overlap of 3). Each group contains exactly 4 cycles, corresponding to the 4 cyclic shifts of a length-4 base string over {1,2,3,4}.

Example Group 1 (base string 1234):

- Cycle 1: port pair (123, 234)
- Cycle 2: port pair (234, 341)
- Cycle 3: port pair (341, 412)
- Cycle 4: port pair (412, 123)

These form a closed chain: (123,234) → (234,341) → (341,412) → (412,123) → closes back to cycle 1.

Other groups are formed similarly from other length-4 strings: 1243, 1324, 1342, 1423, 1432. In total there are 6 groups (since $24/4 = 6 = 3!$), each corresponding to a permutation of {1,2,3,4} up to cyclic rotation.

After closing each chain into a cycle, cutting at the standard position, and restoring ports, we obtain 6 units of length 27. By the self-port property (Lemma 3), each Layer -1 unit has identical head and tail ports:

- Unit G1: port pair (123, 123), length 27
- Unit G2: port pair (124, 124), length 27
- Unit G3: port pair (132, 132), length 27
- Unit G4: port pair (134, 134), length 27
- Unit G5: port pair (142, 142), length 27
- Unit G6: port pair (143, 143), length 27

Closing calculation: $4 \cdot 9 - 4 \cdot 3 = 24$ (cycle length). Cut and add port: $24 + 3 = 27$. ✓

Layer 1 → Layer 2 (Overlap = 2, 3 units/group)

At Layer 2, the 6 units are grouped into closed chains where the last 2 characters of one unit's tail port match the first 2 characters of the next unit's head port (overlap = 2). Since each unit has self-port form (p, p), we simply check if the suffix of p matches the prefix of q with length 2.

Group A:

- Unit 1: port (123, 123)
- Unit 2: port (231, 231)
- Unit 3: port (312, 312)

Chain: (123,123) → (231,231) → (312,312) → closes back. The suffix-prefix matching: "23" (last 2 of 123) = "23" (first 2 of 231); "31" (last 2 of 231) = "31" (first 2 of 312); "12" (last 2 of 312) = "12" (first 2 of 123).

Group B (reverse of Group A):

- Unit 1: port (132, 132)
- Unit 2: port (321, 321)
- Unit 3: port (213, 213)

After closing, cutting, and restoring ports, we obtain 2 units of length 77. These two units are reverses of each other.

Closing calculation: $3 \cdot 27 - 3 \cdot 2 = 75$ (cycle length). Cut and add port: $75 + 2 = 77$. ✓

Final Join (Overlap = 1)

The two final reverse units are joined with overlap 1: $77 + 77 - 1 = 153$

Length: $153 = 1! + 2! + 3! + 4! + 5!$. ✓

6.6 n = 6

One hundred twenty directed cycles. Ports are quadruples (length 4) using digits {1, 2, 3, 4, 5}. The maximum symbol 6 never appears in ports.

Layer 0: Base Cycles

There are $5! = 120$ directed cycles. Each has a port pair (h, t) where both h and t are length-4 strings over {1,2,3,4,5}.

Layer 0 → Layer 1 (Overlap = 4, 5 cycles/group)

At Layer 1, cycles are grouped into closed chains with full port overlap (overlap = 4). Each group contains 5 cycles, corresponding to cyclic shifts of a length-5 base string.

Example Group (base string 12345):

- Cycle 1: port pair (1234, 2345)
- Cycle 2: port pair (2345, 3451)
- Cycle 3: port pair (3451, 4512)
- Cycle 4: port pair (4512, 5123)
- Cycle 5: port pair (5123, 1234)

Closed chain: (1234,2345) → (2345,3451) → (3451,4512) → (4512,5123) → (5123,1234) → closes back.

In total there are $120/5 = 24 = 4!$ groups. After closing, cutting, and restoring ports, we obtain 24 units of length 39, each with the self-port property (p, p):

Closing: $5 \cdot 11 – 5 \cdot 4 = 35$. Cut and add port: $35 + 4 = 39$. ✓

Layer 1 → Layer 2 (Overlap = 3, 4 units/group)

At Layer 2, the 24 units are grouped with overlap = 3 (last 3 of tail port = first 3 of head port). Since units have self-port form (p, p), we match suffix(3) = prefix(3).

Example Group:

- Unit 1: port (1234, 1234)
- Unit 2: port (2341, 2341)
- Unit 3: port (3412, 3412)
- Unit 4: port (4123, 4123)

Chain: (1234,1234) → (2341,2341) → (3412,3412) → (4123,4123) → closes back.
Verification: "234" (last 3 of 1234) = "234" (first 3 of 2341); "341" = "341"; "412" = "412"; "123" = "123".

In total there are $24/4 = 6 = 3!$ groups. After closing, cutting, and restoring ports, we obtain 6 units of length 147.

Closing: $4 \cdot 39 – 4 \cdot 3 = 144$. Cut and add port: $144 + 3 = 147$. ✓

Note: At this layer, the units no longer have the simple self-port property (p, p). The port pair of each merged unit is determined by the cut position, and the head and tail ports may differ.

Layer 2 → Layer 3 (Overlap = 2, 3 units/group)

At Layer 3, the 6 units are grouped with overlap = 2 (last 2 of tail port = first 2 of head port). There are 2 groups of 3 units each.

Example Group (forward):

- Unit 1: port pair (1234, 4123)
- Unit 2: port pair (2314, 4231)
- Unit 3: port pair (3124, 4312)

Closed chain verification:

- Unit 1 tail "4123" → last 2 = "23"; Unit 2 head "2314" → first 2 = "23". Match ✓
- Unit 2 tail "4231" → last 2 = "31"; Unit 3 head "3124" → first 2 = "31". Match ✓
- Unit 3 tail "4312" → last 2 = "12"; Unit 1 head "1234" → first 2 = "12". Match ✓

The second group consists of the reverse units, forming the mirror chain.

After closing, cutting, and restoring ports, we obtain 2 units of length 437. These two units are reverses of each other.

Closing: $3 \cdot 147 - 3 \cdot 2 = 435$. Cut and add port: $435 + 2 = 437$. ✓

Final Join (Overlap = 1)

The two final reverse units are joined with overlap 1: $437 + 437 - 1 = 873$

Length: $873 = \sum_{k=1}^{6} k!$. ✓

Note: For $n = 6$, the linear length of 873 is one greater than the current best known upper bound of 872 due to Houston [5], which was obtained via computational search.

7. Three Variants

The construction naturally yields three variants depending on how the final two reverse units are treated. Let M denote the length of each of the two final units. From the linear construction:

$$2M - 1 = \sum_{k=1}^{n} k! \Rightarrow M = \frac{\sum_{k=1}^{n} k! + 1}{2}$$

7.1 Linear Superpermutation

Join the two final units with a single overlap of 1:

$$L_{linear}(n) = \sum_{k=1}^{n} k!$$

This is the standard superpermutation as defined in the literature.

7.2 Directed Cycle Superpermutation

Join the two final units into a directed cycle, with two overlaps of 1 (one at each junction):

$$L_{directed}(n) = \sum_{k=1}^{n} k! - 1$$

For $n = 6$, this gives length 872. We emphasize that this is the length of a directed cyclic superpermutation, not a linear one. The relationship $L_{directed} = L_{linear} - 1$ is a general fact: any linear superpermutation can be made cyclic by overlapping the first and last $n-1$ characters, saving $n-1$ characters—but the maximum overlap between the beginning and end of our linear construction is only 1 (since the two halves are reverses), hence the saving is only 1.

7.3 Undirected Cycle Superpermutation

In the undirected setting, a permutation and its reverse are considered the same object. The two reverse units are identified as one. Taking a single unit and closing it into a cycle:

$$L_{undirected}(n) = M - 2 = \frac{\sum_{k=1}^{n} k! - 3}{2}$$

The subtraction of 2 accounts for the two port overlaps when closing the single unit into a cycle.

7.4 Summary Table

n	Linear	Directed Cycle	Undirected Cycle
1	1	—	—
2	3	—	—
3	9	8	3
4	33	32	15
5	153	152	75

6	873	872	435

Table 2. Lengths of the three superpermutation variants for small n.

8. Discussion

8.1 Comparison with the Classic Construction

The Ashlock–Tillotson construction [3] achieves the same linear length bound of $\sum_{k=1}^{n} k!$ via a top-down recursive approach. Our construction is bottom-up and based on a different structural principle—the fixed-port invariant. While the resulting lengths coincide, the two approaches offer complementary insights:

- The top-down approach emphasizes how an n-symbol superpermutation extends an $(n-1)$-symbol one, highlighting the inductive structure.
- Our bottom-up approach emphasizes the hierarchical grouping of cycles and the role of reverse pairs as a structural constraint.

The unification of all three variants (linear, directed cycle, undirected cycle) within a single framework is a feature that emerges naturally from the bottom-up perspective.

8.2 Structural Properties

The fixed-port invariant is the most distinctive feature of this construction. Because ports never contain the maximum symbol n, and the port length stays constant at $n-2$ throughout, the matching of units at each layer reduces to a problem on $n-1$ symbols. This hierarchical self-similarity is what drives the factorial counting of units at each layer.

The separation of reverse pairs is equally important. Because reverse units have maximum overlap 1 (independent of n), they can never be merged at high-overlap layers and must always remain in separate groups. This ensures that exactly two units survive to the final layer, which is essential for the clean termination of the construction.

The self-port property after Layer 1 (Lemma 3) further simplifies the structure: from Layer 1 onward until the penultimate layer, each unit can be identified by a single port string, and inter-unit matching reduces to suffix-prefix matching on these port strings.

8.3 Pedagogical Value

One notable advantage of this construction is its transparency. The three rules are simple to state, and the entire procedure can be carried out by hand for $n \leq 5$ without any computation. The port-pair notation makes the grouping structure at each layer explicit and verifiable. This makes it well-suited for teaching introductory combinatorics, where superpermutations serve as an engaging example of combinatorial optimization and recursive construction.

8.4 Limitations and Future Work

The construction does not improve upon the best known upper bounds for $n \geq 6$. The linear length of 873 for $n = 6$ exceeds the current record of 872 by one character. Whether the fixed-port framework can be modified or extended to incorporate the kind of "shortcuts" that enable sub-sum-conjecture lengths is an open question.

Another direction for future work is to investigate whether the port-based perspective yields new lower bounds. The invariant structure might provide a route to combinatorial lower bound arguments that complement existing approaches.

9. Conclusion

We have presented a bottom-up recursive construction for superpermutations based on three invariant rules: fixed port length $n-2$, decreasing overlap from $n-2$ to 1, and separation of reverse cycles at every layer. The construction yields linear superpermutations of length $\sum_{k=1}^{n} k!$ for all n, matching the classic upper bound and achieving the known minimum for $n \leq 5$. The same framework naturally produces directed cycle and undirected cycle variants with closed-form lengths. While the construction does not set new length records, it offers a novel structural perspective based on the fixed-port invariant, unifies three variants in a single framework, and provides an explicit, manually verifiable procedure.

References

[1] D. E. Knuth, *The Art of Computer Programming, Volume 4, Fascicle 2: Generating All Tuples and Permutations*. Addison-Wesley, 2005.

[2] F. Chung, P. Diaconis, and R. Graham, "Universal cycles for combinatorial structures," *Discrete Mathematics*, vol. 110, no. 1–3, pp. 43–59, 1992.

[3] D. Ashlock and J. Tillotson, "Superpermutations," *Congressus Numerantium*, vol. 93, pp. 91–98, 1993.

[4] R. Arratia, "On the Stanley–Wilf conjecture for the number of permutations avoiding a given pattern," *Electronic Journal of Combinatorics*, vol. 6, no. 1, p. N1, 1999.

[5] R. Houston, "Tackling the Minimal Superpermutation Problem," arXiv:1408.5108, 2014.

[6] M. Engen and V. Vatter, "Containing All Permutations," arXiv:1810.08252, 2018.

[7] J. Egan, "Improved upper bounds for superpermutations," unpublished manuscript, 2019.

[8] N. Johnston, "Non-Uniqueness of Minimal Superpermutations," arXiv:1303.4150, 2013.

[9] J. R. Johnson, "Universal Cycles for Permutations," arXiv:0710.5611, 2007.

[10] F. Ruskey and A. Williams, "An Explicit Universal Cycle for the (n-1)-Permutations of an n-Set," *ACM Transactions on Algorithms*, vol. 6, no. 3, 2010.

[11] OEIS Foundation Inc., "The On-Line Encyclopedia of Integer Sequences, A007489," 2026.

[12] G. Dumas, "Superpermutation Matrices," arXiv:1908.04708, 2019.