Pure and Applied Mathematics Quarterly
Volume 18, Number 6, 2619–2645, 2022

On the generalized shuffle-exchange problem*

Xiaoming Sun, Yuan Sun, Kewen Wu, and Zhiyu Xia

Abstract: We investigate the *shuffle-exchange* problem in this paper: given a permutation π on $[n]\times[m]$ and two permutation groups G on $[n]$ and H on $[m]$, the goal is to generate π by alternately using the following two types of operations:

- Select $g_1, g_2, \ldots, g_m \in G$ and perform each g_i on the i-th column of $[n] \times [m]$ in parallel;
- Select $h_1, h_2, \ldots, h_n \in H$ and perform each h_j on the j-th row of $[n] \times [m]$ in parallel.

We discuss the shuffle-exchange, i.e., the composition of these allowable operations, from the perspective of the *Cayley graph*.

For cases where the base groups G and H are both cyclic groups, we prove that the diameter of the underlying Cayley graph, i.e., the minimum number of steps sufficient to achieve any permutation, is upper bounded by $O(\min\{n+m, n\log m, m\log n\})$, which is asymptotically optimal when $\min\{n,m\} = O(1)$ or $n = \Theta(m)$. The main idea is to simulate the shuffle-exchange over symmetric groups with cyclic operations and further accelerate the process with the low-depth *periodic switching network*. For the shuffle-exchange over general groups, we characterize the reachability of any two given vertices on the Cayley graph, and prove the minimum number of steps to achieve a permutation, if possible, is $O(nm)$. This implies that though a connected component of the Cayley graph could contain exponential number of vertices, its diameter is only at most a polynomial of n, m.

1. Introduction

Given a group G with its generator set S, Cayley graph [6] $\Gamma(G,S)$ is defined as an undirected simple graph $\Gamma(G,S) = (V,A)$: the vertex set V is G, and the edge set gathers all S-reachable pair $\{x,y\}$, for which there exists $g \in G$ such

Received July 27, 2021.
*This work was supported in part by the National Natural Science Foundation of China Grant No. 61832003, 61872334, the Strategic Priority Research Program of Chinese Academy of Sciences Grant No. XDB27000000.

that $y = g \circ x$. Numerous specific Cayley graphs [6], defined to rigorously model puzzles and games [10], have been studied from the perspective of both graph theory [6, 16] and computational group theory [1, 11, 12, 20]. The research on the Cayley graph usually focuses on the following two problems [1, 11]:

- Reachability. This asks if there exists a path between a given pair of vertices.
- Diameter. This asks for the longest distance between two (reachable) vertices.

One natural example is that of the *shuffle-exchange*, which finds applications in switching network theory, parallel processing, sorting networks, etc [3, 4, 13]. The *shuffle-exchange* is a composition of permutations selected from generator sets $\mathsf{SES}_{\mathbb{S}_{[n]},\mathbb{S}_{[m]}}$. The *shuffle-exchange set*, i.e., $\mathsf{SES}_{\mathbb{S}_{[n]},\mathbb{S}_{[m]}}$, gathers all possible permutations π constructed in one of the following ways:

- **Row-permutations.** Choose $\tau_1, \tau_2, \ldots, \tau_n \in \mathbb{S}_{[m]}$ and construct $\pi = \prod_{i=1}^{n} \tau_i^{\text{row-}i}$;
- **Col-permutations.** Choose $\sigma_1, \sigma_2, \ldots, \sigma_m \in \mathbb{S}_{[n]}$ and construct $\pi = \prod_{j=1}^{m} \sigma_j^{\text{col-}j}$.

Permutations in $\mathsf{SES}_{\mathbb{S}_{[n]},\mathbb{S}_{[m]}}$ permute elements on each row (or each column) simultaneously. Putting it differently, we have $\mathsf{SES}_{\mathbb{S}_{[n]},\mathbb{S}_{[m]}} = \mathsf{SES}_{\{\mathsf{id}\},\mathbb{S}_{[m]}} \cup \mathsf{SES}_{\mathbb{S}_{[n]},\{\mathsf{id}\}}$ where $\mathsf{SES}_{\{\mathsf{id}\},\mathbb{S}_{[m]}}$ and $\mathsf{SES}_{\mathbb{S}_{[n]},\{\mathsf{id}\}}$ gather the row-permutations and col-permutations respectively. Note that $\mathsf{SES}_{\{\mathsf{id}\},\mathbb{S}_{[m]}}$ and $\mathsf{SES}_{\mathbb{S}_{[n]},\{\mathsf{id}\}}$ are subgroups of $\mathbb{S}_{[n]\times[m]}$. Thus, without loss of generality, we treat a shuffle-exchange as a process where row-permutations and col-permutations are performed in an alternative fashion. We say a shuffle-exchange is of k-norm if it is a composition of k permutations selected from the shuffle-exchange set $\mathsf{SES}_{\mathbb{S}_{[n]},\mathbb{S}_{[m]}}$. For a given permutation $\pi \in \mathbb{S}_{[n]\times[m]}$, we say π is of k-norm, denoted by $\mathsf{SEN}_{\mathbb{S}_{[n]},\mathbb{S}_{[m]}}(\pi) = k$, if no $(k-1)$-norm shuffle-exchange can achieve π.

It is known [13] that the whole graph is connected (i.e., $\langle \mathsf{SES}_{\mathbb{S}_{[n]},\mathbb{S}_{[m]}} \rangle = \mathbb{S}_{[n]\times[m]}$). Meanwhile, the diameter of the shuffle-exchange Cayley graph is 3 (i.e., $\max_\sigma \mathsf{SEN}_{\mathbb{S}_{[n]},\mathbb{S}_{[m]}}(\sigma) = 3$). Moreover, the shortest path between any two vertices (i.e., the shuffle-exchange achieving given π) can be computed efficiently. These results are also called *2-dimensional Shuffle Exchange Problem*.

Based on this, one expanding direction is to think about high dimension, which leads to the well-known conjecture *Shuffle Exchange Conjecture* [4] with many applications in switching network design. In this paper, we focus on another expanding version of this problem on dimension 2, which is an

independent new problem. We investigate a generalized version of the shuffle-exchange, where the base group $\mathbb{S}_{[n]}, \mathbb{S}_{[m]}$ are replaced with general groups G, H. One natural example is the cyclic shuffle-exchange that represents the weakest transitive groups. The cyclic group $\mathbb{C}_n$ (resp., $\mathbb{C}_m$) only allows row (resp., column) shifts. See Figure 1 as an example.

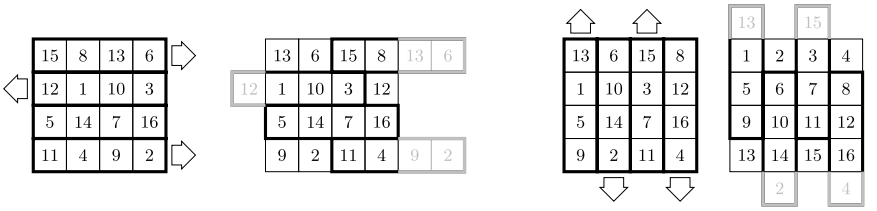

Figure 1: Sort numbers in a matrix via a cyclic shuffle-exchange.

Based on the cyclic shuffle-exchange, the *Loopover* is a puzzle aiming to sort a disordered grid (See Figure 1) [17]. In each step, the player can select one row/column and then shift it. It is not hard to provide an $O(n^2)$-step solution for an $n \times n$ Loopover instance. This is indeed essentially optimal by a straightforward counting argument. In the following sections, we will give an $O(n)$-norm cyclic shuffle-exchange solution, which implies the same upper bound.

Main Results. For the k-dimension shuffle-exchange over symmetric groups, we show:

- Diameter. For all positive integers k and n, permutations in $\mathbb{S}_{[n]^k}$ can be achieved by a $(2k-1)$-norm k-dimension shuffle-exchange over symmetric groups.

For the shuffle-exchange over cyclic permutations groups, we show:

- Reachability. If either n or m is even, then cyclic shuffle-exchange can achieve all permutations. Otherwise, a cyclic shuffle-exchange can and only can achieve all even permutations.
- Diameter. Define $d(n,m) = \min\{n+m, n\log m, m\log n\}$. Any even permutation can be achieved by an $O(d(n,m))$-norm cyclic shuffle-exchange. Furthermore, when n or m is even, any odd permutation can also achieved by an $O(d(n,m))$-norm cyclic shuffle-exchange.

For the shuffle-exchange over general group G, H, we show:

- Connectivity. Assume integers $n, m \geq 2$. $\langle \mathsf{SES}_{G,H} \rangle = \mathbb{S}_{[n]\times[m]}$ if and only if G, H are transitive and G or H includes an odd permutation.

- DIAMETER. All achievable permutation can be achieved by an $O(nm)$-norm (G, H)-shuffle-exchange.

Related works. The general study of Cayley graph problem has a long history. Cayley graph was first introduced by Cayley [6], which represents a group by a directed graph. The motivation is to study the relationship between a group and its generating sets. See survey [14]. A natural problem on this topic is that given G and one of its generator sets S, how long the diameter of the Cayley graph is. This is widely studied in the language of computational group theory [1, 11].

For special cases of the Cayley graph diameter problem, there are many intriguing results in puzzles and games [10]. The phrase "God's number" represents the diameter of the Cayley graph of the 3-level RUBIK'S CUBE PUZZLE, which is shown to be 20 [19]. After that, several works focused on properties of n-level Rubik's Cube Puzzle [7, 8]. Another interesting puzzle closely related to Cayley graph is $(n^2 - 1)$-PUZZLE [18].

Another special case of Cayley graph diameter problem is Shuffle Exchange Conjecture [4]. Given a permutation, the aim is to rearrange it by a switching network with specific ordered layers. [5] gave a proof for this conjecture, but later [2] showed the proof is incomplete. Researchers also focus on permutation rearrangement by switching networks [9, 15].

Organization. In Section 2, we give formal definitions and notations, and introduce the diameter result of the shuffle-exchange over symmetric groups. In Section 3, we extend it to the k-dimension shuffle-exchange and upper-bound the diameter, which also leads to a low-depth periodic switching network. In Section 4, we study a specific case, the cyclic shuffle-exchange. In Section 5, we investigate the general case, where the shuffle-exchange is based on arbitrary groups. In Section 6, we summarize this paper and list further directions.

2. Preliminaries

For integers n and m, $[n]$ denotes $\{1, 2, \ldots, n\}$ and $[n, m]$ denotes $\{n, n + 1, \ldots, m\}$. Notation $+_n$ is defined as a looping operator where $a +_n b$ equals to the unique integer in $[n]$ equivalent to $a + b$ modulo n, and $-_n$ is defined similarly.

Permutation groups. In this paper, $\mathbb{S}_T$ denotes the *symmetric group* over set T and $\mathbb{A}_T$ denotes the *alternating group* over T. In the following content, we use expression $\prod \square$ to compose several pair-wise commutative permutations. The commutativity may be sometimes not noted, when it is obvious in

context. For non-commutative permutations, we use expression $\square \circ \square \circ \cdots \circ \square$ instead. For a subset $S \subseteq \mathbb{S}_T$, $\langle S \rangle \leq \mathbb{S}_T$ denotes the group generated by the elements in S.

Define id_T as the identity in $\mathbb{S}_T$. Define $(\underrightarrow{t})_n \in \mathbb{S}_{[n]}$ as the *t-spanning shift* over $[n]$ s.t. $(\underrightarrow{t})_n(x) = x +_n t$ for all $x \in [n]$. $\mathbb{C}_n$ denotes the *cyclic permutation group*, which gathers all $(\underrightarrow{t})_n$ for $t \in [0, n-1]$. Note that $\mathbb{C}_n \leq \mathbb{A}_{[n]}$ for all odd n. Define $(x \leftrightarrow y)_T \in \mathbb{S}_T$ as the *swap* of x, y in T. For notations $\mathsf{id}_T, (\underrightarrow{t})_n, (x \leftrightarrow y)_T$, the subscripts n, T may be omitted when they are clear in context. Given a partition $\mathcal{P}$ of a set T, i.e. a set of pair-wise disjoint subset from T whose union equals to T, define *$\mathcal{P}$-invariant permutation group* $\mathbb{S}_T^{\mathcal{P}} \leq \mathbb{S}_T$ as $\mathbb{S}_T^{\mathcal{P}} := \{\sigma \in \mathbb{S}_T \mid \sigma(P) = P \text{ for all } P \in \mathcal{P}\}$. Furthermore, for permutation $\sigma \in \mathbb{S}_T^{\mathcal{P}}$ and subset $P \in \mathcal{P}$, define $\sigma_P \in \mathbb{S}_P$ as $\sigma_P(p) = \sigma(p)$ for all $p \in P$, when the universal set T and the partition $\mathcal{P}$ are fixed in context. Given group $G \leq \mathbb{S}_T$, define the *orbit* of $x \in T$ over G as $\{g(x) \mid g \in G\}$. Define an equivalence relation $\sim$ where for $x, y \in T$, $x \sim y$ if and only if their orbits are the same one. The relation partitions T into several subsets, called *a family of orbits for* G or *the orbits of* G, denoted as T/G. Note that for group G with orbits $\mathcal{P}$, $G \leq \mathbb{S}_T^{\mathcal{P}}$ holds.

Given $i \in [n]$ and $\sigma \in \mathbb{S}_{[m]}$, define $\sigma^{\text{row-}i} \in \mathbb{S}_{[n]\times[m]}$ as $\sigma^{\text{row-}i}(x, y) = (x, \sigma(y))$ if $x = i$, and $\sigma^{\text{row-}i}(x, y) = (x, y)$ otherwise. Similarly, define $\pi^{\text{col-}j} \in \mathbb{S}_{[n]\times[m]}$ for $j \in [m]$ and $\pi \in \mathbb{S}_{[n]}$. Formally, $\sigma^{\text{row-}i}$ and $\pi^{\text{col-}i}$ equal to $\mathsf{id}_{\{i\}} \otimes \sigma$ and $\pi \otimes \mathsf{id}_{\{j\}}$ respectively.

Shuffle-exchange. Given $G \leq \mathbb{S}_{[n]}, H \leq \mathbb{S}_{[m]}$, define the *$(G, H)$-shuffle-exchange permutation set* $\mathsf{SES}_{G,H} \subset \mathbb{S}_{[n]\times[m]}$ over G, H as

$$\mathsf{SES}_{G,H} := \left\{\prod_{i\in[n]} \sigma_i^{\text{row-}i} \mid \sigma_1, \ldots, \sigma_n \in H\right\} \cup \left\{\prod_{j\in[m]} \pi_j^{\text{col-}j} \mid \pi_1, \ldots, \pi_m \in G\right\}.$$

Note that in the most case, $\mathsf{SES}_{G,H}$ is not a subgroup, but a subset of $\mathbb{S}_{[n]\times[m]}$ since it may not be close for composing. Actually, $\mathsf{SES}_{G,H}$ can be treated as an union of 2 subgroups $\mathsf{SES}_{\{\mathsf{id}\},H}, \mathsf{SES}_{G,\{\mathsf{id}\}}$. For integer k and $\pi_1, \pi_2, \ldots, \pi_k \in \mathsf{SES}_{G,H}$, we say $\pi_k \circ \ldots \circ \pi_2 \circ \pi_1$ is a *k-norm (G, H)-shuffle-exchange*. For $\pi \in \langle \mathsf{SES}_{G,H} \rangle$, define the *$(G, H)$-shuffle-exchange norm* $\mathsf{SEN}_{G,H}(\pi)$ as the minimum integer k satisfying there exists a k-norm (G, H)-shuffle-exchange achieving π. Besides, for $S \subseteq \langle \mathsf{SES}_{G,H} \rangle$, define

$$\mathsf{SEN}_{G,H}(S) := \max_{\pi\in S} \mathsf{SEN}_{G,H}(\pi).$$

In group theory, $\mathsf{SEN}_{G,H}(\cdot)$ is also called *word norm* with respect to $\mathsf{SES}_{G,H}$, with the following properties:

- *identity of indiscernibles*: $\mathsf{SEN}_{G,H}(\sigma) = 0$ if and only if $\sigma = \mathsf{id}$;
- *symmetry*: for all σ, $\mathsf{SEN}_{G,H}(\sigma) = \mathsf{SEN}_{G,H}(\sigma^{-1})$;
- *triangle inequality*: for all σ, π, $\mathsf{SEN}_{G,H}(\pi \circ \sigma) \leq \mathsf{SEN}_{G,H}(\sigma) + \mathsf{SEN}_{G,H}(\pi)$.

Furthermore, we list some other properties of $\mathsf{SEN}_{G,H}(\cdot)$:

- *partial order*: if $G' \leq G$ and $H' \leq H$, $\mathsf{SEN}_{G,H}(\sigma) \leq \mathsf{SEN}_{G',H'}(\sigma)$;
- *transitivity*: for all permutation σ and group G, G', H, H',

$$\mathsf{SEN}_{G,H}(\sigma) \leq \mathsf{SEN}_{G',H'}(\sigma) \cdot \mathsf{SEN}_{G,H}(\mathsf{SES}_{G',H'}).$$

Shuffle-exchange over symmetric groups. In order to make the paper self-contained, we first show how to achieve an arbitrary permutation in $\mathbb{S}_{[n]\times[m]}$ with a 3-norm $(\mathbb{S}_{[n]}, \mathbb{S}_{[m]})$-shuffle-exchange, which is given by [13].

Theorem 1. $\mathsf{SEN}_{\mathbb{S}_{[n]},\mathbb{S}_{[m]}}\left(\mathbb{S}_{[n]\times[m]}\right) \leq 3$.

Proof. Let $\sigma \in \mathbb{S}_{[n]\times[m]}$. Assume $\sigma(i,j) = (i',j')$ for all $(i,j) \in [n] \times [m]$. Construct a bipartite multigraph $G = (U \cup V, E, r)$ first, with the vertex sets $U = \{u_1, u_2, ..., u_n\}, V = \{v_1, v_2, ..., v_n\}$, the edge set $E = [n] \times [m]$ and the edge identifier $r : E \to U \times V$ defined as for $(i,j) \in [n] \times [m]$, $r(i,j) = (u_i, v_{i'})$.

For all $u_i \in U$, $\deg_G(u_i)$ equals to m, the number of elements in $[n] \times [m]$ permuted to row i by σ. Besides, for all $v_i \in V$, $\deg_G(v_i)$ also equals to m, the number of elements in a row. So, G is a regular bipartite graph and Hall's theorem says there exists a family of perfect matchings $\mathcal{M}$, a partition $M_1 \sqcup M_2 \sqcup \cdots \sqcup M_m = E$ such that for all subset M_i, every distinct e_1, e_2 in M_i do not share endpoints, i.e. both the first and the second components of $r(e_1)$ and $r(e_2)$ are different respectively.

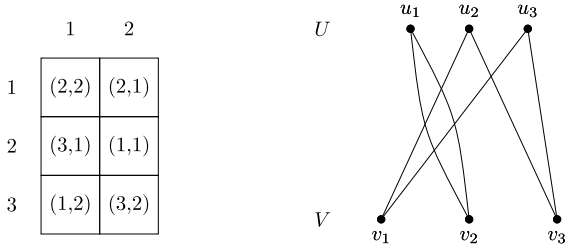

Figure 2: An example of $\sigma \in \mathbb{S}_{[n]\times[m]}$, and its corresponding bipartite multigraph. Notice that there exists a multiple edge from u_1 to v_2.

Next we give the achieving shuffle-exchange $\pi_3 \circ \pi_2 \circ \pi_1$, where $\pi_1 = \prod_{i=1}^{n} \pi_{1,i}^{\text{row-}i}, \pi_2 = \prod_{i=1}^{m} \pi_{2,i}^{\text{col-}i}$ and $\pi_3 = \prod_{i=1}^{n} \pi_{3,i}^{\text{row-}i}$. For all $(i,j) \in [n] \times [m]$

with (i, j) contained in M_t, let

$$\pi_{1,i}(j) = t \quad \pi_{2,t}(i) = i', \text{ and } \pi_{3,i'}(t) = j'.$$

First, we will show they are well-defined permutations in $\mathbb{S}_{[n]}$ and $\mathbb{S}_{[m]}$.

- Note that the definition says all $\pi_{1,i}$ for $i \in [n]$ is defined and the preimage set of these $\pi_{1,i}$ is $[m]$. Besides, different j_1, j_2 must be mapped to different images by all $\pi_{1,i}$, otherwise $(i, j_1), (i, j_2)$ would be contained in the same M_t with a shared endpoint u_i, which conflicts with that M_t is a matching. So, $\pi_{1,1}, \pi_{1,2}, \ldots, \pi_{1,n}$ are well-defined permutations in $\mathbb{S}_{[m]}$.
- For all $t \in [m]$, since a pair with some i as the first component appears in M_t for exactly one time, $\pi_{2,t}$ is defined with the preimage set of $[n]$. The fact that M_t is a matching, i.e. different vertices in U is linked with different vertices in V, ensures $\pi_{2,t}$ is also an injection. Thus, $\pi_{2,1}, \pi_{2,2}, \ldots, \pi_{2,m}$ are well-defined permutations in $\mathbb{S}_{[n]}$.
- For all $t \in [m]$, pair p with the second component of $r(p)$ for some i' appears in M_t for exactly one time, which implies that all $\pi_{3,i'}$ for $i' \in [n]$ is defined, with the preimage set of $[m]$. For some fixed i' and distinct t_1, t_2, let $p_1 \in M_{t_1}, p_2 \in M_{t_2}$ satisfying the second components of $r(p_1), r(p_2)$ are i'. Note such p_1, p_2 are unique. Since $p_1 \neq p_2$ are permuted to the same row by σ, they must be permuted to different columns, which implies $\pi_{3,i'}$ is an injection. So, $\pi_{3,1}, \pi_{3,2}, \ldots, \pi_{3,n}$ are well-defined permutations in $\mathbb{S}_{[m]}$.

Second, it is easy to show $\sigma = \pi_3 \circ \pi_2 \circ \pi_1$ via tracing all elements in $[n] \times [m]$, for example (i, j) with $t, (i', j')$ defined as above:

$$(i, j) \xrightarrow{\pi_1} (i, t) \xrightarrow{\pi_2} (i', t) \xrightarrow{\pi_3} (i', j') = \sigma(i, j)$$

□

3. *k*-dimension shuffle-exchange over symmetric groups

We can extend shuffle-exchanges to a k-dimension version, $(G^{\text{dim-}k})$-shuffle-exchanges, with the base set $\mathsf{SES}^{k\text{-dim}}_{\mathbb{S}_{[n]}} \subseteq \mathbb{S}_{[n]^k}$ defined as follows. See Figure 3 from examples.

$$\mathsf{SES}^{k\text{-dim}}_{\mathbb{S}_{[n]}} = \bigcup_{i=1}^{k} \mathbb{S}_{[n]^k}^{\{\{j_1\}\times\cdots\times\{j_{i-1}\}\times[n]\times\{j_{i+1}\}\times\cdots\times\{j_k\} | j_1,\ldots,j_{i-1},j_{i+1},\ldots,j_k\in[n]\}}$$

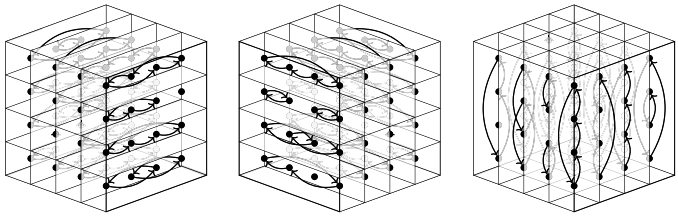

Figure 3: 3 types of permutations in $\mathsf{SES}^{3\text{-dim}}_{\mathbb{S}_{[4]}}$.

Besides, define the $(G^{\text{dim-}k})$-shuffle-exchange norm $\mathsf{SEN}^{k\text{-dim}}_{\mathbb{S}_{[n]}}$ for permutations and sets similarly as that in above text. Note that Theorem 1 can be naturally generalized to that $\mathsf{SEN}_{\mathbb{S}_S,\mathbb{S}_T}(\mathbb{S}_{S\times T}) \leq 3$, which directly upper-bounds $\mathsf{SEN}^{k\text{-dim}}_{\mathbb{S}_{[n]}}(\mathbb{S}_{[n]^k})$.

Theorem 2. $\mathsf{SEN}^{k\text{-}dim}_{\mathbb{S}_{[n]}}(\mathbb{S}_{[n]^k}) \leq 2k-1$.

Proof. By induction, assume $\mathsf{SEN}^{k\text{-dim}}_{\mathbb{S}_{[n]}}(\mathbb{S}_{[n]^k}) \leq 2k-1$ for some $k \geq 2$. For the case on $k+1$, given $\pi \in \mathbb{S}_{[n]^k}$, construct $\pi' \in \mathbb{S}_{[n]\times[n^k]}$ via arbitrary bijection η between $[n]^k$ and $[n^k]$ as

$$\pi'(i,j) := (\pi(i,\eta(j))_1, \eta^{-1}(\pi(i,\eta(j))_2, \pi(i,\eta(j))_3, \ldots, \pi(i,\eta(j))_{k+1}))$$

for all $(i,j) \in [n]\times[n^k]$. Theorem 1 says there exist $\pi'_1, \pi'_3 \in \mathsf{SES}_{S_{[n]},\{\mathsf{id}\}}$ and $\pi'_2 \in \mathsf{SES}_{\{\mathsf{id}\},S_{[n^k]}}$ such that $\pi' = \pi'_3 \circ \pi'_2 \circ \pi'_1$. Back to $\mathbb{S}_{[n]^k}$ and construct $\pi_1, \pi_2, \pi_3 \in \mathbb{S}_{[n]^k}$ where for all $i \in [n], \mathbf{j} \in [n]^k$,

$$\pi_\ell(i,\mathbf{j}) = (\pi'_\ell(i,\eta^{-1}(\mathbf{j}))_1, \eta(\pi'_\ell(i,\eta^{-1}(\mathbf{j})))).$$

It is easy to verify that $\pi = \pi_3 \circ \pi_2 \circ \pi_1$. Furthermore, $\pi'_1, \pi'_3 \in \mathsf{SES}_{S_{[n]},\{\mathsf{id}\}}$ and $\pi'_2 \in \mathsf{SES}_{\{\mathsf{id}\},S_{[n^k]}}$ imply that $\pi_1, \pi_3 \in \mathsf{SES}_{\mathbb{S}_{[n]},\{\mathsf{id}_{[n]^k}\}}$ and $\pi_2 \in \mathsf{SES}_{\{\mathsf{id}_{[n]}\},\mathbb{S}_{[n]^k}}$. So, we have

$$\mathsf{SEN}^{(k+1)\text{-dim}}_{\mathbb{S}_{[n]}}(\mathbb{S}_{[n]^{k+1}}) \leq 2\cdot \mathsf{SEN}^{(k+1)\text{-dim}}_{\mathbb{S}_{[n]}}(\mathsf{SES}_{\mathbb{S}_{[n]},\{\mathsf{id}_{[n]^k}\}}) + \mathsf{SEN}^{(k+1)\text{-dim}}_{\mathbb{S}_{[n]}}(\mathsf{SES}_{\{\mathsf{id}_{[n]}\},\mathbb{S}_{[n]^k}}).$$

Note that $\mathsf{SES}_{\mathbb{S}_{[n]},\{\mathsf{id}_{[n]^k}\}} \subseteq \mathsf{SES}^{(k+1)\text{-dim}}_{\mathbb{S}_{[n]}}$. Thus $\mathsf{SEN}^{(k+1)\text{-dim}}_{\mathbb{S}_{[n]}}(\mathsf{SES}_{\mathbb{S}_{[n]},\{\mathsf{id}_{[n]^k}\}}) = 1$. Besides, we will prove that $\mathsf{SEN}^{(k+1)\text{-dim}}_{\mathbb{S}_{[n]}}(\mathsf{SES}_{\{\mathsf{id}_{[n]}\},\mathbb{S}_{[n]^k}}) \leq \mathsf{SEN}^{k\text{-dim}}_{\mathbb{S}_{[n]}}(\mathbb{S}_{[n]^k})$ holds. Let $t = \mathsf{SEN}^{k\text{-dim}}_{\mathbb{S}_{[n]}}(\mathbb{S}_{[n]^k})$. For $\pi \in \mathsf{SES}_{\{\mathsf{id}_{[n]}\},\mathbb{S}_{[n]^k}}$ with $\pi = \bigoplus_{i=1}^n \mathsf{id}_{\{i\}} \otimes \pi_i$ where $\pi_1, \pi_2, \ldots, \pi_n \in \mathbb{S}_{[n]^k}$, there exists $\pi_{i,1}, \pi_{i,2}, \ldots, \pi_{i,t} \in \mathsf{SES}^{k\text{-dim}}_{\mathbb{S}_{[n]}}$ such

that $\pi_i = \pi_{i,t} \circ \ldots \circ \pi_{i,2} \circ \pi_{i,1}$ for all $i \in [n]$. Thus, we have

$$\begin{aligned}\pi &= \bigoplus_{i=1}^{n} \mathsf{id}_{\{i\}} \otimes \pi_{i,t} \circ \cdots \circ \pi_{i,2} \circ \pi_{i,1} \\ &= \left(\bigoplus_{i=1}^{n} \mathsf{id}_{\{i\}} \otimes \pi_{i,t}\right) \circ \cdots \circ \left(\bigoplus_{i=1}^{n} \mathsf{id}_{\{i\}} \otimes \pi_{i,2}\right) \circ \left(\bigoplus_{i=1}^{n} \mathsf{id}_{\{i\}} \otimes \pi_{i,t}\right).\end{aligned}$$

Since $\bigoplus_{i=1}^{n} \mathsf{id}_{\{i\}} \otimes \pi_{i,j} \in \mathsf{SES}_{\mathbb{S}_{[n]}}^{(k+1)\text{-dim}}$ for all $j \in [k]$,

$$\mathsf{SEN}_{\mathbb{S}_{[n]}}^{(k+1)\text{-dim}}(\mathsf{SES}_{\{\mathsf{id}_{[n]}\},\mathbb{S}_{[n]^k}}) \leq t.$$

Combining the induction assumption, we prove $\mathsf{SEN}_{\mathbb{S}_{[n]}}^{(k+1)\text{-dim}}(\mathbb{S}_{[n]^{k+1}}) \leq 2k+1$, which closes the induction and finishes the proof. □

Theorem 2 further implies an interesting result, which is equivalent to a low-depth *periodic switching network* investigated in [4]. It is also a key tool to prove some other results in this paper.

Lemma 1. *Given an integer n, let $N = 2^n$. For all $\sigma \in \mathbb{S}_{[N]}$, there exist integers $s_1, s_2, \ldots, s_{2n-1}$ and sets $T_1, T_2, \ldots, T_{2n-1} \subseteq [N]$ satisfying*

- *$t_1, t_2, t_1 + s_i, t_2 + s_i$ are distinct for all distinct $t_1, t_2 \in T_j$;*
- *and $\sigma = \prod_{i=1}^{2n-1} \left(\prod_{t \in T_i} (t \leftrightarrow (t + s_i))\right)$.*

Proof. First, we build a mapping $\eta : [2]^n \to [N]$ such that $\eta(x) = \sum_{i=1}^{n} (x_i - 1)2^{i-1} + 1$. It is trivial to verify that η is a bijection. ($\eta^{-1}(\cdot)$ could be treated as somehow a binary representation of integers. To keep the consistency of notations, we assume it is between $[2]^n$ and $[N]$, although it may be more standard to define it between $[0, N-1]$ and $\{0,1\}^n$.) Let π be the permutation in $\mathbb{S}_{[2]^n}$ obtained by σ with η conjugated, i.e. $\pi = \eta^{-1} \circ \sigma \circ \eta$. Note that Theorem 2 says $\mathsf{SES}_{\mathbb{S}_{[2]}}^{n\text{-dim}}(\pi) \leq 2n-1$ by a shuffle-exchange $\pi = \pi_{2n-1} \circ \cdots \circ \pi_2 \circ \pi_1$. Thus, σ can decomposed as

$$\begin{aligned}\sigma &= \eta \circ \pi \circ \eta^{-1} \\ &= \eta \circ \pi_{2n-1} \circ \cdots \circ \pi_2 \circ \pi_1 \circ \eta^{-1} \\ &= (\eta \circ \pi_{2n-1} \circ \eta^{-1}) \circ \cdots \circ (\eta \circ \pi_2 \circ \eta^{-1}) \circ (\eta \circ \pi_1 \circ \eta^{-1})\end{aligned}$$

Consider some term $(\eta \circ \pi_i \circ \eta^{-1})$, denoted as σ_i, for example. Assume $\pi_i \in$

$\mathbb{S}_{[2]^n}^{\left\{\{j\}_{-k_i}\times[2]\mid j_{-k_i}\in[2]\right\}}$ holds[1] for all $i \in [2n-1]$ since $\pi_i \in \mathsf{SES}^{n\text{-dim}}_{\mathbb{S}_{[2]}}$. Thus, σ_i can be decomposed as

$$\sigma_i = \prod_{j_{-k_i}\in[2]} \eta \circ \left((\pi_i)_{\{j_{-k_i}\}\times[2]} \oplus \mathsf{id}_{\overline{\{j_{-k_i}\}\times[2]}}\right) \circ \eta^{-1}.$$

Note that for all $j_{-k_i} \in [2]$, $\eta \circ \left((\pi_i)_{\{j\}_{-k_i}\times[2]} \oplus \mathsf{id}_{\overline{\{j\}_{-k_i}\times[2]}}\right) \circ \eta^{-1}$ is either id or $(\eta(j_{-k_i},1) \leftrightarrow \eta(j_{-k_i},2))$ since $|\{j\}_{-k_i} \times [2]| = 2$. Furthermore, the definition of η says that $\eta(j_{-k_i},2) - \eta(j_{-k_i},1) = 2^{k_i-1}$ always holds. To summarize, for all $i \in [2n-1]$, let $s_i = 2^{k_i-1}$ and $T_i \subseteq [N]$ gather all $\eta(j_{-k_i},1)$ for all $j_{-k_i} \in [2]$ satisfying $(\pi_i)_{\{j_{-k_i}\times[2]\}} \neq \mathsf{id}$. $(s_i, T_i)_{i\in[2n-1]}$ satisfies the conditions and finishes the proof. □

Lemma 1 constructs a "$(2\log_2(N)-1)$-depth" swap decomposition of $\pi \in \mathbb{S}_{[N]}$ when N is a power of 2. Actually, it can be extended to a construction for general N, with loss of a constant fact on the depth. Note that on any continuous range from $[N]$ with the length of a power of 2, Lemma 1 ensures that any permutations can be achieved with the depth of $O(\log(N))$. It is easy to see that constant number of appropriate "sub-permutations" achieve arbitrary $\sigma \in \mathbb{S}_{[N]}$, which leads to the following corollary.

Corollary 1. *Given an integer n, for all $\pi \in \mathbb{S}_{[n]}$, there exist $\ell = O(\log n)$, $s_1, s_2, \ldots, s_\ell$ and sets $T_1, T_2, \ldots, T_\ell \subseteq [n]$ satisfying*

- *$t_1, t_2, t_1 + s_i, t_2 + s_i$ are distinct for all distinct $t_1, t_2 \in T_i$ for $i \in [\ell]$,*
- *and $\pi = \prod_{i=1}^{\ell} \left(\prod_{t\in T_i}(t \leftrightarrow (t+s_i))\right)$.*

4. Shuffle-exchange over cyclic permutation groups

In this section, we focus on the shuffle-exchanges over cyclic permutation groups. A straightforward parity analysis says an odd permutation in $\mathbb{S}_{[n]\times[m]}$ cannot be achieved by a $(\mathbb{C}_n, \mathbb{C}_m)$-shuffle-exchange with odd n, m.

Fact 1. *If n and m are odd, $\sigma \notin \langle \mathsf{SES}_{\mathbb{C}_n,\mathbb{C}_m} \rangle$ for all $\sigma \in \mathbb{S}_{[n]\times[m]} \setminus \mathbb{A}_{[n]\times[m]}$.*

In the other cases, we can upper-bound the shuffle-exchange norm as the following theorems say.

[1] We simplify expressions $[a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_\ell]$, $[a_1, \ldots, a_{i-1}, t, a_{i+1}, \ldots, a_\ell]$ and $[\{a_1\} \times \cdots \times \{a_{i-1}\} \times T \times \{a_{i+1}\} \times \cdots \times \{a_\ell\}]$ with $[a_{-i}]$, $[a_{-i}, t]$ and $[\{a\}_{-i} \times T]$ respectively, when the number ℓ is clear in context. Here, symbols a, i, t, T can be arbitrary.

Theorem 3. *If n or m is even,*

$$\mathsf{SEN}_{\mathbb{C}_n,\mathbb{C}_m}\left(\mathbb{S}_{[n]\times[m]}\right) = O(\min(n+m, n\log m, m\log n)).$$

Theorem 4. $\mathsf{SEN}_{\mathbb{C}_n,\mathbb{C}_m}\left(\mathbb{A}_{[n]\times[m]}\right) = O(\min(n+m, n\log m, m\log n))$.

The main idea to decompose a given permutation σ is to decompose it first into a $(\mathbb{S}_{[n]}, \mathbb{S}_{[m]})$-shuffle-exchange, which allows rows or columns permuted arbitrarily instead of cyclically. Theorem 1 shows that to achieve an arbitrary permutation, a 3-norm shuffle-exchange is sufficient (no matter the parity of n, m or the target permutation). Thus, our task is converted to design cyclic shuffle-exchanges achieving permutations in $\mathsf{SES}_{\mathbb{S}_{[n]},\mathbb{S}_{[m]}}$. For the case where n or m is even first, the methods to achieve $\mathsf{SES}_{\{\mathsf{id}\},\mathbb{S}_{[m]}}$ are summarized as follows:

m is even	$O(m)$-norm	by Corollary 2
	$O(n\log m)$-norm	by Corollary 3
n is even	$O(m+\log n)$-norm	by Corollary 4
	$O(n\log m)$-norm	by Corollary 5

Note that $\mathsf{SES}_{\mathbb{S}_{[n]},\{\mathsf{id}\}}$ can be achieved in the same way. Thus, there are 3 ways to achieve a constant $\mathbb{S}_{[n]} \times \mathbb{S}_{[m]}$ shuffle-exchange, which leads to Theorem 3:

- $O(m)$-norm for $\mathsf{SES}_{\{\mathsf{id}\},\mathbb{S}_{[m]}}$, $O(n+\log m)$-norm for $\mathsf{SES}_{\mathbb{S}_{[n]},\{\mathsf{id}\}}$ $\Rightarrow O(n+m)$-norm;
- $O(n\log m)$-norm for $\mathsf{SES}_{\{\mathsf{id}\},\mathbb{S}_{[m]}}$, $O(n+\log m)$-norm for $\mathsf{SES}_{\mathbb{S}_{[n]},\{\mathsf{id}\}}$ $\Rightarrow O(n\log m)$-norm;
- $O(m)$-norm for $\mathsf{SES}_{\{\mathsf{id}\},\mathbb{S}_{[m]}}$, $O(m\log n)$-norm for $\mathsf{SES}_{\mathbb{S}_{[n]},\{\mathsf{id}\}}$ $\Rightarrow O(m\log n)$-norm.

The basic idea to handle the cases with odd n, m is similar, which is described in Subsection 4.3.

4.1. Cases with even m

Our methods are based on the following observation. As that shown in Figure 4, a cyclic shuffle-exchange "↓1 ←2 ↑1 →2 ↓1" achieves an swap between ♔ and ♖.

The strategy in Figure 4, called *the basic strategy* in the following text, is a starting point of our shuffle-exchange design. Obverse that:

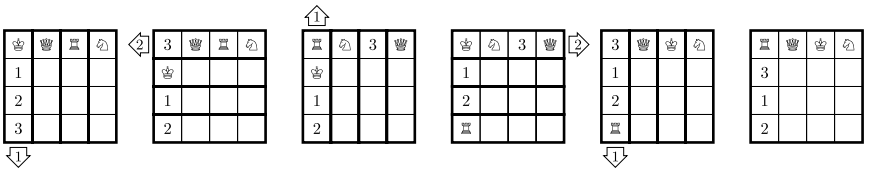

Figure 4: Swap ♔ and ♖ via a cyclic shuffle-exchange.

- Via a similar shuffle-exchange, any swap in row-1 can be achieved actually, with a pair of conjugate shifts to move the target piece to the column-1 first and "undo" the shift in the end, if necessary. For example, "$\boxed{1}\ \boxed{1}\ \boxed{1}\ \boxed{1}\ \boxed{1}\ \boxed{1}\ \boxed{1}$" swaps ♔ and ♖.
- Besides achieving a swap in row-1, the shuffle-exchange also rotates column-1 (excluding the first position) from 123 to 231.
- "$\boxed{1}\ \boxed{2}\ \boxed{1}\ \boxed{2}\ \boxed{1}$" actually does the same thing in row-1, while rotates 123 reversely to 231.

To achieve a permutation on a single row, for example row-1, we decompose it $O(m)$ swaps. The key idea is that a pair of swaps can be achieved via the basic strategy, with opposite direction to rotate column-1, which cancels the change on column-1. This observation directly implies $\mathsf{SES}_{\{\mathsf{id}\},\mathbb{A}_{[m]}}$ can be achieved by a cyclic shuffle-exchange, i.e., $\mathsf{SES}_{\{\mathsf{id}\},\mathbb{A}_{[m]}} \leq \langle \mathsf{SEN}_{\mathbb{C}_n,\mathbb{C}_m} \rangle$. However, a trivial method requires to achieve the n permutations row by row, and achieve at most m pairs of swaps for each one of them, which leads to a $O(nm)$-norm cyclic shuffle-exchange.

We will introduce two different parallelizing tricks to reduce the norm down to $O(n+m)$ and $O(n \log m)$ respectively. The first way is to achieve permutations on non-adjacent rows in parallel. The trick is that $\tau_1, \tau_2, \ldots, \tau_k$ can be achieved in parallel by a 5-norm cyclic shuffle-exchange, where each permutation swaps 2 positions in the same row and the rows for them are distinct and pair-wise non-adjacent. In fact, it is just a simple variant of the basic strategy, with the only difference that all target rows is shifted while just row-1 shifted originally at step 2 and step 4.

Lemma 2. $\mathsf{SEN}_{\mathbb{C}_n,\mathbb{C}_m}\left(\mathsf{SES}_{\{\mathsf{id}\},\mathbb{A}_{[m]}}\right) = O(m)$.

Proof. Divide $[n]$ into $[n] = U \sqcup V \sqcup W$ such that i and $i +_n 1$ are not contained in the same subset for all $i \in [n]$. Decompose σ into $\sigma = \sigma_U \circ \sigma_V \circ \sigma_W$ where $\sigma_U := \prod_{i \in U} (\sigma_i)^{\text{row-}i}$, and σ_V, σ_W are defined in the same way. Consider σ_U as an example. Note that every permutation in $\mathbb{A}_{[m]}$ can be decomposed into $2m$ swaps. Assume $\sigma_i = (j_{i,2m} \leftrightarrow k_{i,2m}) \circ \cdots \circ (j_{i,1} \leftrightarrow k_{i,1})$ for all $i \in [n]$.

Then, we claim $\sigma_U = \tau_{2m} \circ \tau_{2m-1} \circ \cdots \circ \tau_1$ where for all $t \in [2m]$, $\tau_t \in \mathbb{S}_{[n]\times[m]}$, satisfying $\mathsf{SEN}_{\mathbb{C}_n,\mathbb{C}_m}(\tau_t) = O(1)$, is defined as

$$\tau_t := \prod_{i\in U} \left(\underrightarrow{j_{i,t} - 1}\right)^{\text{row-}i} \circ \left(\underrightarrow{(-1)^t}\right)^{\text{col-1}} \circ \prod_{i\in U} \left(\underrightarrow{k_{i,t} - j_{i,t}}\right)^{\text{row-}i} \circ \left(\underrightarrow{(-1)^{t+1}}\right)^{\text{col-1}} \circ \prod_{i\in U} \left(\underrightarrow{j_{i,t} - k_{i,t}}\right)^{\text{row-}i} \circ \left(\underrightarrow{(-1)^t}\right)^{\text{col-1}} \circ \prod_{i\in U} \left(\underrightarrow{1 - j_{i,t}}\right)^{\text{row-}i}.$$

First, $\tau_{2m} \circ \cdots \circ \tau_1$ achieves $(\sigma_i)^{\text{row-}i}$ for all $i \in U$, i.e., $\sigma_U(p) = \tau_{2m} \circ \cdots \circ \tau_1(p)$ for any position p on some row-i where $i \in U$. It can be proved by showing τ_t achieves $(j_{i,t} \leftrightarrow k_{i,t})^{\text{row-}i}$ for all $i \in U$.

- Let $p = (i, j)$ be a position on row-i which is not $(i, j_{i,t})$ or $(i, k_{i,t})$. Note that $(i, j +_m 1 -_m j_{i,t})$ and $(i, j +_m 1 -_m k_{i,t})$ are not on col-1. Thus,

$$\tau_t(p) = \prod_{i\in U} \left(\underrightarrow{j_{i,t} - 1}\right)^{\text{row-}i} \circ \prod_{i\in U} \left(\underrightarrow{k_{i,t} - j_{i,t}}\right)^{\text{row-}i} \circ \prod_{i\in U} \left(\underrightarrow{j_{i,t} - k_{i,t}}\right)^{\text{row-}i} \circ \prod_{i\in U} \left(\underrightarrow{1 - j_{i,t}}\right)^{\text{row-}i} (p) = p.$$

- Without loss of generality, assume t is even. Note that $i +_n 1$ and $i -_n 1$ are not in U. Trace $(i, j_{i,t})$ and $(i, k_{i,t})$ during τ_t:

$$(i, j_{i,t}) \to (i, 1) \to (i +_n 1, 1) \to (i +_n 1, 1) \to (i, 1) \to (i, 1 +_m k_{i,t} -_m j_{i,t}) \to (i, 1 +_m k_{i,t} -_m j_{i,t}) \to (i, k_{i,t})$$

$$(i, k_{i,t}) \to (i, k_{i,t} +_m 1 -_m j_{i,t}) \to (i, k_{i,t} +_m 1 -_m j_{i,t}) \to (i, 1) \to (i -_n 1, 1) \to (i -_n 1, 1) \to (i, 1) \to (i, j_{i,t})$$

Second, if p is not on any row-i where $i \in U$, $\tau_{2m} \circ \cdots \circ \tau_1$ keeps p invariant. Specifically, if p is not on col-1, it keeps invariant obviously. Besides, it is easy to check the following fact for all $i \in [n] \setminus U$:

$$\tau_t(i, 1) = \begin{cases} (i +_n 1, 1) & \text{if } t \text{ is odd and } i +_n 1 \notin U, \\ (i -_n 1, 1) & \text{if } t \text{ is even and } i -_n 1 \notin U, \\ (i +_n 2, 1) & \text{if } t \text{ is odd and } i +_n 1 \in U, \\ (i -_n 2, 1) & \text{if } t \text{ is even and } i -_n 1 \in U. \end{cases}$$

Combining the property of U, that $j \in U$ implies $j +_n 1, j -_n 1 \notin U$, we have that $\tau_{t+1} \circ \tau_t(i, 1) = (i, 1)$ for all even t and $i \in [n] \setminus U$, and that $\tau_{2m} \circ \cdots \circ \tau_1$ also keeps such $(i, 1)$ invariant. Thus, $\mathsf{SEN}_{\mathbb{C}_n,\mathbb{C}_m}(\sigma_U)$, $\mathsf{SEN}_{\mathbb{C}_n,\mathbb{C}_m}(\sigma_V)$, $\mathsf{SEN}_{\mathbb{C}_n,\mathbb{C}_m}(\sigma_W)$ are $O(m)$, which implies $\mathsf{SEN}_{\mathbb{C}_n,\mathbb{C}_m}(\sigma) = O(m)$ as the result. □

Corollary 2. *If m is even,* $\mathsf{SEN}_{\mathbb{C}_n,\mathbb{C}_m}\left(\mathsf{SES}_{\{\mathsf{id}\},\mathbb{S}_{[m]}}\right) = O(m)$.

Proof. Let σ be an arbitrary permutation in $\mathsf{SES}_{\{\mathsf{id}\},\mathbb{S}_{[m]}}$ which can be represented as a composition of permutations on each row, i.e., $\sigma = \prod_{i\in[n]} \sigma_i^{\text{row-}i}$ where $\sigma_i \in \mathbb{S}_{[m]}$ for all $i \in [n]$. Construct $\pi = \prod_{i\in[n]} \pi_i^{\text{row-}i} \in \mathsf{SES}_{\{\mathsf{id}\},\mathbb{C}_m}$ satisfying $\pi_i := (\underrightarrow{1})$ if σ_i is odd, and $\pi_i = \mathsf{id}$ otherwise. Since m is even and $(\underrightarrow{1})$ is an odd permutation, $\sigma' := \pi \circ \sigma$ is in $\mathsf{SES}_{\{\mathsf{id}\},\mathbb{A}_{[m]}}$. Lemma 2 says $\mathsf{SEN}_{\mathbb{C}_n,\mathbb{C}_m}(\sigma') = O(m)$. Note that $\sigma = \pi^{-1} \circ \sigma'$ where π^{-1} is in $\mathsf{SES}_{\mathbb{C}_n,\mathbb{C}_m}$. Thus, $\mathsf{SEN}_{\mathbb{C}_n,\mathbb{C}_m}(\sigma) = O(m)$. □

The second way is to achieve swaps on the same row in parallel. Now, we focus on a single even permutation on, for example, row-1. The trick is that swaps with the same spanning, i.e. $(i_1 \leftrightarrow i_1 +_m \ell), (i_2 \leftrightarrow i_2 +_m \ell), \ldots, (i_k \leftrightarrow i_k +_m \ell)$, can be achieved in parallel using another variant of the base strategy with also norm of 5, as that shown in the following figure.

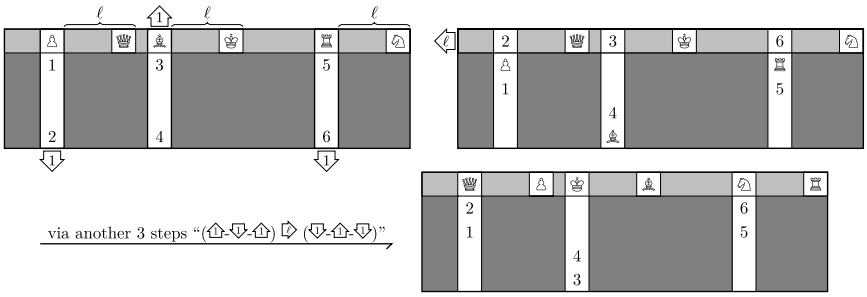

Figure 5: Achieve 3 swaps in parallel.

Previous works show a special swap decomposition that an arbitrary permutation can be decomposed into a few "good" components, of which each one is a composition of several swaps with the same spanning, as Corollary 1 says. Thus, an $O(\log m)$-norm cyclic shuffle-exchange is sufficient to achieve σ in row-1. Note that this shuffle-exchange may change the other rows. Recall the strategy in Figure 5. The direction of row shifts is arbitrary, which means if even number of swaps are on col-i, the change on col-i can be recovered by alternating shift direction selection. So, we design a "plus version" decomposition in Lemma 3 for even permutations, in which there are even number of swaps on each column.

Lemma 3. *For all $\pi \in \mathbb{A}_{[n]}$, there exist $\ell = O(\log n)$, $s_j \in [n]$ and sets $T_j \subseteq [n]$ for $j \in [\ell]$ satisfying*

- $i_1, i_2, i_1 + s_j, i_2 + s_j$ *are distinct for all distinct* $i_1, i_2 \in T_j$*;*
- $\pi = \prod_{i \in T_\ell}(i \leftrightarrow (i +_n s_\ell)) \circ \cdots \circ \prod_{i \in T_1}(i \leftrightarrow (i +_n s_1))$*;*
- *and each* $i \in [n]$ *is contained in even sets among* $T_1, T_2, \ldots, T_\ell$*.*

Proof. Corollary 1 says there exist $\ell', (s_j)_{j\in[\ell']}, (T_j)_{j\in[\ell']}$ satisfying the first 2 constraints. Our aim is to append $s_{\ell'+1}, s_{\ell'+2}, \ldots, s_\ell$ and $T_{\ell'+1}, T_{\ell'+2}, \ldots, T_\ell$ to satisfy the third constraint. We say $i \in [n]$ contained by an odd number of sets among $T_1, T_2, ..., T_j$ is an *odd position up to* j, gathered by a set denoted as S_j. For positive integer k let $r_k = \max S_{\ell'+2k-2}$. Define

- $s_{\ell'+2k-1} = -\lfloor r_k/2 \rfloor$, $s_{\ell'+2k} = \lfloor r_k/2 \rfloor$;
- $T_{\ell'+2k-1} = S_{\ell'+2k-2} \cap [r_k - \lfloor r_k/2 \rfloor + 1, r_k]$, $T_{\ell'+2k} = T_{\ell'+2k-1} - \lfloor r_k/2 \rfloor$.

Note that the odd positions up to $\ell' + 2k - 2$ in $[r_k - \lfloor r_k/2 \rfloor + 1, r_k]$ are contained in $T_{\ell'+2k-1}$ and not in $T_{\ell'+2k}$, which implies they are not in $S_{\ell'+2k}$ and $r_{k+1} \le \lceil r_k/2 \rceil$. So, there exists $k = O(\log n)$ such that $r_k \le 1$, due to the initial case $r_1 \le n$. Let $\ell = \ell' + 2k - 1$ where $\ell = O(\log n)$. We claim $\ell, (s_j)_{j\in[\ell]}, (T_j)_{j\in[\ell]}$ satisfy all the 3 constraints.

Recalling these definitions, the first constraint is met obviously. For pairs $(s_{\ell'+2k-1}, T_{\ell'+2k-1})$ and $(s_{\ell'+2k}, T_{\ell'+2k})$, note that $(i \leftrightarrow i +_n s_{\ell'+2k-1}) = (i - \lfloor r_k/2 \rfloor \leftrightarrow i - \lfloor r_k/2 \rfloor +_n s_{\ell'+2k})$ holds, and the swaps in some groups, i.e. $(i \leftrightarrow i + s_j)$ for $i \in T_j$, are pair-wise commutative since the first constraint satisfied. So, we have

$$\prod_{i \in T_{\ell'+2k}} (i \leftrightarrow (i +_n s_{\ell'+2k})) \circ \prod_{i \in T_{\ell'+2k-1}} (i \leftrightarrow (i +_n s_{\ell'+2k-1})) = \mathsf{id},$$

i.e. the appended swaps actually do nothing except changing the parity of the "hitting number", which means the first constraint is satisfied. Since π is even, the total number of swaps is also even, which implies the number of odd positions up to ℓ is even. Then $\max S_\ell = r_k \le 1$ means $\max S_\ell = 0$, i.e. there is no odd position finally, which satisfies the third constraint. □

Lemma 4. *For all* $\pi \in \mathbb{A}_{[m]}$ *and* $r \in [n]$, $\mathsf{SEN}_{\mathbb{C}_n, \mathbb{C}_m}(\pi^{\text{row-}r}) = O(\log m)$.

Proof. Let π be an arbitrary permutation in $\mathbb{A}_{[m]}$. Decompose π in the way of Lemma 3 and get ℓ, $(s_j)_{j\in[\ell]}$ and $(T_j)_{j\in[\ell]}$. Define $a_{1,i}, a_{2,i}, \ldots, a_{\ell,i}$ for $i \in [m]$ as a list with alternating ± 1 on the position j where $i \in T_j$, and 0 on the others, or formally,

$$a_{j,i} = \begin{cases} 1 & \text{if } j = 1 \text{ and } i \in T_1, \\ (-1)^{a_{1,i} + a_{2,i} + \cdots + a_{j-1,i}} & \text{if } j > 1 \text{ and } i \in T_j, \\ 0 & \text{if } i \notin T_j. \end{cases}$$

We claim $\pi^{\text{row-}r} = \tau_\ell \circ \cdots \circ \tau_2 \circ \tau_1$ where τ_j for all $j \in [\ell]$, with the norm of $O(1)$, is defined as

$$\tau_j = \prod_{i \in T_j} \left(\xrightarrow{a_{j,i}}\right)^{\text{col-}i} \circ \left(\xrightarrow{s_j}\right)^{\text{row-}r} \circ \prod_{i \in T_j} \left(\xrightarrow{-a_{j,i}}\right)^{\text{col-}i} \circ \left(\xrightarrow{-s_j}\right)^{\text{row-}r} \circ \prod_{i \in T_j} \left(\xrightarrow{a_{j,i}}\right)^{\text{col-}i}.$$

First, τ_j achieves $\prod_{i \in T_j}(i \leftrightarrow i + s_j)$ on row-r, i.e., for all positions p on row-r, $\tau_j(p) = \prod_{i \in T_j}(i \leftrightarrow i + s_j)^{\text{row-}r}(p)$. Trace all positions $p = (r, i)$ on row-r during τ_j as follows.

$$(r,i) \to \begin{cases} (r +_n a_{j,i}, i) \to (r +_n a_{j,i}, i) \to \\ \qquad (r,i) \to (r, i +_m s_j) \to (r, i +_m s_j) & \text{if } i \in T_j, \\ (r,i) \to (r, i -_m s_j) \to (r -_n a_{j,i}, i -_m s_j) \to \\ \qquad (r -_n a_{j,i}, i -_m s_j) \to (r, i -_m s_j) & \text{if } i -_m s_j \in T_j, \\ (r,i) \to (r, i -_m s_j) \to (r, i -_m s_j) \to (r,i) \to (r,i) & \text{otherwise.} \end{cases}$$

Note that $i \in T_j$ and $i -_n s_j \in T_j$ cannot be satisfied at the same time due to the definition of T_j.

Then, define permutation $\sigma \in \mathbb{S}_{[n]}$ as

$$\sigma(i) = \begin{cases} i & \text{if } i = r, \\ i +_n 2 & \text{if } i = r -_n 1 \text{ and} \\ i +_n 1 & \text{otherwise,} \end{cases}$$

i.e., σ shifts the elements in $[n]$ except r. With this notation, we have τ_j achieves $\prod_{i \in [m]} (\sigma^{a_{j,i}})^{\text{col-}i}$ on the other rows except row-r, since τ_j moves (k, i) for some $i \in T_j$ and $k \neq r$ as follows:

$$(k,i) \to \begin{cases} (k +_n a_{j,i}, i) \to (k +_n a_{j,i}, i -_n s_j) \to \\ \quad (k +_n a_{j,i}, i -_n s_j) \to (k +_n a_{j,i}, i) \to (k +_n 2a_{j,i}, i) & \text{if } k +_n a_{j,i} = r, \\ (k +_n a_{j,i}, i) \to (k +_n a_{j,i}, i) \to \\ \qquad (k,i) \to (k,i) \to (k +_n a_{j,i}, i) & \text{if } k +_n a_{j,i} \neq r. \end{cases}$$

Decompose τ_j into the action on row-r and that on the others as $\tau_j = \prod_{i \in T_j} (\sigma^{a_{j,i}})^{\text{col-}i} \circ \prod_{i \in T_j}(i \leftrightarrow i + s_j)$. It is obvious that for any $j_1, j_2 \in [\ell]$, $\prod_{i \in T_{j_1}} (\sigma^{a_{j_1,i}})^{\text{col-}i}$ and $\prod_{i \in T_{j_2}}(i \leftrightarrow i + s_{j_2})$ are commutative. So, we have $\tau_\ell \circ \cdots \circ \tau_2 \circ \tau_1$ equals to

$$\begin{aligned} &\prod_{i \in T_\ell} (\sigma^{a_{\ell,i}})^{\text{col-}i} \circ \prod_{i \in T_\ell} (i \leftrightarrow i + s_\ell) \circ \cdots \circ \prod_{i \in T_1} (\sigma^{a_{1,i}})^{\text{col-}i} \circ \prod_{i \in T_1} (i \leftrightarrow i + s_1) \\ &= \left(\prod_{i \in T_\ell} (\sigma^{a_{\ell,i}})^{\text{col-}i} \circ \cdots \circ \prod_{i \in T_1} (\sigma^{a_{1,i}})^{\text{col-}i} \right) \circ \left(\prod_{i \in T_\ell} (i \leftrightarrow i + s_\ell) \circ \cdots \circ \prod_{i \in T_1} (i \leftrightarrow i + s_1) \right) \\ &= \prod_{i \in [m]} \left(\sigma^{\sum_{j \in [\ell]} a_{j,i}} \right)^{\text{col-}i} \circ \pi^{\text{row-}r}. \end{aligned}$$

Recall the second constraint in Lemma 3 says each $i \in [m]$ is contained in even number of sets among $T_1, T_2, \ldots, T_\ell$. So, $\sum_{j\in[\ell]} a_{j,i} = 0$ according to the definition, and $\tau_\ell \circ \cdots \circ \tau_2 \circ \tau_1 = \pi^{\text{row-}r}$ as the result. □

Corollary 3. *If m is even, for all $\pi \in \mathbb{S}_{[m]}$ and $r \in [n]$, $\mathsf{SEN}_{\mathbb{C}_n,\mathbb{C}_m}(\pi^{\text{row-}r}) = O(\log m)$.*

Proof. If π is even, Lemma 4 says $\mathsf{SEN}_{\mathbb{C}_n,\mathbb{C}_m}(\pi^{\text{row-}r}) = O(\log m)$. Otherwise, we have

$$\mathsf{SEN}_{\mathbb{C}_n,\mathbb{C}_m}(\pi^{\text{row-}r}) \le \mathsf{SEN}_{\mathbb{C}_n,\mathbb{C}_m}\left((\pi \circ (\underrightarrow{1})_m)^{\text{row-}r}\right) + \mathsf{SEN}_{\mathbb{C}_n,\mathbb{C}_m}\left((\underrightarrow{-1})_m^{\text{row-}r}\right) = O(\log m)$$

since $\pi \circ (\underrightarrow{1})_m$ is odd and $(\underrightarrow{-1})_m^{\text{row-}r}$ is in $\mathsf{SES}_{\mathbb{C}_n,\mathbb{C}_m}$. □

4.2. Cases with even n

Our solution to the cases with even n is to pre-process the given π. In detail, we append an extra swap $(1 \leftrightarrow 2)$ on each odd row-i, where π_i is odd, to obtain $\pi' \in \mathsf{SES}_{\{\mathsf{id}\},\mathbb{A}_{[m]}}$. Lemma 5 shows how these swaps on different rows be achieved in parallel using the first kind of parallelization in Subsection 4.1. Recalling that in the proof for Lemma 2, the parity of permutations ensures col-1 recovered, which is not satisfied now. Here, we utilize Corollary 3, which shows the existence of a $O(\log n)$-norm cyclic shuffle-exchange achieving a permutation on a single row, to recover col-1.

Lemma 5. *If n is even, for all $\pi \in \mathsf{SES}_{\{\mathsf{id}\},\mathbb{S}_{[m]}}$, there exists $\tau \in \mathbb{S}_{[n]\times[m]}$ with $\mathsf{SEN}_{\mathbb{C}_n,\mathbb{C}_m}(\tau) = O(\log n)$ such that $\pi \circ \tau \in \mathsf{SES}_{\{\mathsf{id}\},\mathbb{A}_{[m]}}$.*

Proof. Define $\tau := \prod_{i\in[n]:\pi_i\notin\mathbb{A}_{[m]}} (1 \leftrightarrow 2)^{\text{row-}i}$ which satisfies $\pi\circ\tau \in \mathsf{SES}_{\{\mathsf{id}\},\mathbb{A}_{[m]}}$ obviously. Partition $[n]$ into $[n] = U \sqcup V \sqcup W$ and define τ_U, τ_V and τ_W as Lemma 2. We claim $\mathsf{SEN}_{\mathbb{C}_n,\mathbb{C}_m}(\tau_U) = O(\log n)$ as well as τ_V and τ_W. Consider the following 5-norm cyclic shuffle-exchange

$$\tau'_U := \left(\underrightarrow{1}\right)^{\text{col-1}} \circ \prod_{i\in U:\pi_i \text{ is odd}} \left(\underrightarrow{1}\right)^{\text{row-}i} \circ \left(\underrightarrow{-1}\right)^{\text{col-1}} \circ \prod_{i\in U:\pi_i \text{ is odd}} \left(\underrightarrow{-1}\right)^{\text{row-}i} \circ \left(\underrightarrow{1}\right)^{\text{col-1}}.$$

$\tau'_U \circ \tau_U$ will not move anyone out of col-1 eventually, which is shown by tracing $(i,j) \in [n] \times [2,m]$:

$$(i,j) \to \begin{cases} (i,j-1) \to (i +_n 1, j-1) \to \\ \quad (i +_n 1, j-1) \to (i,j-1) \to (i,j) \to (i,j) & \text{if } i \in U,\ \pi_i \text{ is odd and } j = 2, \\ (i,j) \to (i,j) \to (i, j -_m 1) \to \\ \quad (i, j -_m 1) \to (i,j) \to (i,j) & \text{if } i \in U,\ \pi_i \text{ is odd and } j > 2, \\ \qquad\qquad \text{invariant} \to (i,j) & \text{otherwise.} \end{cases}$$

Then, Corollary 3 says $\mathsf{SEN}_{\mathbb{C}_n,\mathbb{C}_m}(\tau'_U \circ \tau_U) = \log(n)$, which also implies that $\mathsf{SEN}_{\mathbb{C}_n,\mathbb{C}_m}(\tau_U) = \log(n)$. Similarly, τ_V, τ_W can also be achieved by $O(\log n)$-norm shuffle-exchanges, and $\mathsf{SEN}_{\mathbb{C}_n,\mathbb{C}_m}(\tau) = \log(n)$ as the result. □

Combined with Corollary 2 and Corollary 3, Lemma 5 leads to the following corollaries.

Corollary 4. *If n is even,* $\mathsf{SEN}_{\mathbb{C}_n,\mathbb{C}_m}\left(\mathsf{SES}_{\{\mathsf{id}\},\mathbb{S}_{[m]}}\right) = O(m + \log n)$.

Corollary 5. *If n is even,* $\mathsf{SEN}_{\mathbb{C}_n,\mathbb{C}_m}\left(\mathsf{SES}_{\{\mathsf{id}\},\mathbb{S}_{[m]}}\right) = O(n \log m)$.

4.3. Cases with odd n, m

Theorem 1 says π equals to a composition of $\pi_1, \pi_2, \pi_3 \in \mathsf{SEN}_{\mathbb{S}_{[n]},\mathbb{S}_{[m]}}$. However, although assuming π is even, someone among π_1, π_2, π_3 could still be odd, which can not be achieved by a cyclic shuffle-exchange with odd n, m. Our solution is to switch the parity, utilizing the following simple equalities

$$(1 \leftrightarrow 2)^{\text{row-1}} \circ ((1 \leftrightarrow 2)^{\text{row-1}} \circ (1 \leftrightarrow 2)^{\text{col-1}}) \circ (1 \leftrightarrow 2)^{\text{col-1}} = \mathsf{id}$$
$$(1 \leftrightarrow 2)^{\text{col-1}} \circ ((1 \leftrightarrow 2)^{\text{col-1}} \circ (1 \leftrightarrow 2)^{\text{row-1}}) \circ (1 \leftrightarrow 2)^{\text{row-1}} = \mathsf{id}$$

with a negligible cost due to the following fact.

Fact 2. $(1 \leftrightarrow 2)^{\text{row-1}} \circ (1 \leftrightarrow 2)^{\text{col-1}}$ *and* $(1 \leftrightarrow 2)^{\text{col-1}} \circ (1 \leftrightarrow 2)^{\text{row-1}}$ *can be achieved by* 4*-norm cyclic shuffle-exchanges.*

Proof. It can be verified that

$$(1 \leftrightarrow 2)^{\text{row-1}} \circ (1 \leftrightarrow 2)^{\text{col-1}} = \left(\underset{\rightarrow}{1}\right)^{\text{col-1}} \circ \left(\underset{\rightarrow}{1}\right)^{\text{row-1}} \circ \left(\underset{\rightarrow}{-1}\right)^{\text{col-1}} \circ \left(\underset{\rightarrow}{-1}\right)^{\text{row-1}}.$$

Note that $(1 \leftrightarrow 2)^{\text{col-1}} \circ (1 \leftrightarrow 2)^{\text{row-1}} = ((1 \leftrightarrow 2)^{\text{row-1}} \circ (1 \leftrightarrow 2)^{\text{col-1}})^{-1}$. So, $\mathsf{SEN}_{\mathbb{C}_n,\mathbb{C}_m}((1 \leftrightarrow 2)^{\text{col-1}} \circ (1 \leftrightarrow 2)^{\text{row-1}}) \leq 4$ also holds. □

Using a similar method as that for Lemma 5, an even permutation in $\mathsf{SES}_{\{\mathsf{id}\},\mathbb{S}_{[m]}}$ is switched into a permutation in $\mathsf{SES}_{\{\mathsf{id}\},\mathbb{A}_{[m]}}$ with $O(\log n)$ cost. Then, Lemma 2 and Lemma 4 help us deal with the rest of the process.

Lemma 6. *For odd n, m and $\pi \in \mathsf{SES}_{\{\mathsf{id}\},\mathbb{S}_{[m]}} \cap \mathbb{A}_{[n]\times[m]}$, there exists $\tau \in \mathbb{S}_{[n]\times[m]}$ with $\mathsf{SEN}_{\mathbb{C}_n,\mathbb{C}_m}(\tau) = O(\log n)$ such that $\pi \circ \tau \in \mathsf{SES}_{\{\mathsf{id}\},\mathbb{A}_{[m]}}$.*

Proof. Assume $\pi = \prod_{i\in[n]} \sigma_i^{\text{row-}i}$ where $\sigma_1, \sigma_2, \ldots, \sigma_n \in \mathbb{S}_{[m]}$. Let $S \subseteq [n]$ gather all $i \in [n]$ with σ_i odd. Define $\tau = \prod_{i\in S}(1 \leftrightarrow 2)^{\text{row-}i}$ then $\pi \circ \tau \in \mathsf{SES}_{\{\mathsf{id}\},\mathbb{A}_{[m]}}$ obviously. Note that $|S|$ is even since π is even. Thus, there is a

partition $[n] = U \sqcup V \sqcup W$ satisfying U, V, W are non-adjacent sets and all $|U \cap S|, |V \cap S|, |W \cap S|$ are even. We claim $\tau_U = \prod_{i \in U \cap S} (1 \leftrightarrow 2)^{\text{row-}i}$ can be achieve by a $O(\log n)$-norm cyclic shuffle-exchange. Note the argument can also be applied to τ_V and τ_W.

Via a similar analysis as that in the proof of Lemma 5, there exists τ'_U with norm of 5 such that $\tau'_U \circ \tau_U$ will not move anyone out of col-1 eventually. Recall that $|U \cap S|$ is even, which implies τ_U is even. Besides, n, m are odd implies $\tau'_U \in \langle \mathsf{SES}_{\mathbb{C}_n, \mathbb{C}_m} \rangle \leq \mathbb{A}_{[n] \times [m]}$ is even, which is followed by that $\tau'_U \circ \tau_U$ is even. Thus, Lemma 4 says $\mathsf{SEN}_{\mathbb{C}_n, \mathbb{C}_m}(\tau'_U \circ \tau_U) = O(\log n)$. So, $\mathsf{SEN}_{\mathbb{C}_n, \mathbb{C}_m}(\tau_U) = O(\log n)$, as well as τ_V, τ_W and τ. □

Proof of Theorem 4. If n or m is even, Theorem 3 can be directly applied. So, we assume n, m are odd in the proof.

Let π be an arbitrary even permutation in $\mathbb{A}_{[n] \times [m]}$. Theorem 1 says there exist $\pi_1, \pi_3 \in \mathsf{SES}_{\{\mathsf{id}\}, \mathbb{S}_{[m]}}$ and $\pi_2 \in \mathsf{SES}_{\mathbb{S}_{[n]}, \{\mathsf{id}\}}$ such that $\pi = \pi_3 \circ \pi_2 \circ \pi_1$. Consider the parity of π_1, π_2 and π_3.

- If π_1 and π_2 are odd, define $\pi'_1 = (1 \leftrightarrow 2)^{\text{row-1}} \circ \pi_1, \tau_1 = (1 \leftrightarrow 2)^{\text{col-1}} \circ (1 \leftrightarrow 2)^{\text{row-1}}$, $\pi'_2 = \pi_2 \circ (1 \leftrightarrow 2)^{\text{col-1}}, \tau_2 = \mathsf{id}$ and $\pi'_3 = \pi_3$;
- If π_2 and π_3 are odd, define $\pi'_1 = \pi_1, \tau_1 = \mathsf{id}, \pi'_2 = (1 \leftrightarrow 2)^{\text{col-1}} \circ \pi_2, \tau_2 = (1 \leftrightarrow 2)^{\text{row-1}} \circ (1 \leftrightarrow 2)^{\text{col-1}}$ and $\pi'_3 = \pi_3 \circ (1 \leftrightarrow 2)^{\text{row-1}}$;
- If π_1 and π_3 are odd, define $\pi'_1 = (1 \leftrightarrow 2)^{\text{row-1}} \circ \pi_1, \tau_1 = (1 \leftrightarrow 2)^{\text{col-1}} \circ (1 \leftrightarrow 2)^{\text{row-1}}, \pi'_2 = (1 \leftrightarrow 2)^{\text{col-1}} \circ \pi_2 \circ (1 \leftrightarrow 2)^{\text{col-1}}, \tau_2 = (1 \leftrightarrow 2)^{\text{row-1}} \circ (1 \leftrightarrow 2)^{\text{col-1}}$ and $\pi'_3 = \pi_3 \circ (1 \leftrightarrow 2)^{\text{row-1}}$;
- If all π_1, π_2 and π_3 are even, define $\pi'_1 = \pi_1, \tau_1 = \mathsf{id}, \pi'_2 = \pi_2, \tau_2 = \mathsf{id}$ and $\pi'_3 = \pi_3$.

Note that in all these cases, $\pi = \pi'_3 \circ \tau_2 \circ \pi'_2 \circ \tau_1 \circ \pi'_1$ where π'_1, π'_2, π'_3 are even and τ_1, τ_2 are of constant-norm. Since n, m are odd, Lemma 6 says there exist $\sigma_1, \sigma_2, \sigma_3$ with $O(\log n + \log m)$-norm, satisfying $\pi'_1 \circ \sigma_1, \pi'_3 \circ \sigma_3 \in \mathsf{SES}_{\{\mathsf{id}\}, \mathbb{A}_{[m]}}$ and $\pi'_2 \circ \sigma_2 \in \mathsf{SES}_{\mathbb{A}_{[n]}, \{\mathsf{id}\}}$. Using Lemma 2 and Lemma 4, the analysis for Theorem 3 also leads to $\mathsf{SEN}_{\mathbb{C}_n, \mathbb{C}_m}(\mathsf{SES}_{\mathbb{A}_{[n]}, \mathbb{A}_{[m]}}) = O(\min(n + m, n \log m, m \log n))$. Thus, we have

$$\begin{aligned}
\mathsf{SEN}_{\mathbb{C}_n, \mathbb{C}_m}(\pi) =& \mathsf{SEN}_{\mathbb{C}_n, \mathbb{C}_m}(\pi'_3 \circ \tau_2 \circ \pi'_2 \circ \tau_1 \circ \pi'_1) \\
=& \mathsf{SEN}_{\mathbb{C}_n, \mathbb{C}_m}((\pi'_3 \circ \sigma_3) \circ \sigma_3^{-1} \circ \tau_2 \circ (\pi'_2 \circ \sigma_2) \circ \sigma_2^{-1} \circ \tau_1 \circ (\pi'_1 \circ \sigma_1) \circ \sigma_1^{-1}) \\
\leq& 3\mathsf{SEN}_{\mathbb{C}_n, \mathbb{C}_m}(\mathsf{SES}_{\mathbb{A}_{[n]}, \mathbb{A}_{[m]}}) + O(\log n + \log m) \\
\leq& O(\min(n + m, n \log m, m \log n)).
\end{aligned}$$

□

5. Shuffle-exchange over general groups

In this section, we discuss a generalized version, the shuffle-exchange over arbitrary given groups $G \leq \mathbb{S}_{[n]}, H \leq \mathbb{S}_{[m]}$. Note that $\mathsf{SES}_{G,H}$ may not achieve all permutations in $\mathbb{S}_{[n]\times[m]}$. We have given an example in Section 3: $\langle \mathsf{SES}_{\mathbb{C}_n,\mathbb{C}_m} \rangle = \mathbb{A}_{[n]\times[m]} \lneq \mathbb{S}_{[n]\times[m]}$ when n, m are odd. Another example is that $\langle \mathsf{SES}_{G,H} \rangle$ cannot be transitive when G or H is not transitive, and extremely, $\langle \mathsf{SES}_{\{\mathsf{id}\},\{\mathsf{id}\}} \rangle = \{\mathsf{id}\}$. Our main result on the shuffle-exchange over general groups is a characterization of $\langle \mathsf{SES}_{G,H} \rangle$. We further show that once π is achievable, π can be achieved by an $O(nm)$-norm shuffle-exchange.

For σ in some permutation group G, our characterization cares about the parity of σ on the orbits of G. We introduce a linear space to describe such local parity as follows. Given a partition $\mathcal{P}$ of a set T, define the *orbit-wise parity* $\zeta^{\mathcal{P}} : \mathbb{S}_T^{\mathcal{P}} \to \mathbb{F}_2^{\mathcal{P}}$ as $(\zeta^{\mathcal{P}}(\sigma))_P = 1$ if and only if σ_P is odd for all $P \in \mathcal{P}$.[2] Given subspace $\mathcal{U} \leq \mathbb{F}_2^S, \mathcal{V} \leq \mathbb{F}_2^T$, define the *row-column spanning space*

$$\mathcal{M}_{\mathcal{U},\mathcal{V}} := \operatorname{span}\left(\mathcal{U} \otimes \mathbb{F}_2^T + \mathbb{F}_2^S \otimes \mathcal{V}\right) \leq \mathbb{F}_2^{S\times T},$$

i.e. $\mathcal{M}_{\mathcal{U},\mathcal{V}}$ is spanned by all $M = u \otimes x + y \otimes v$ for $u \in \mathcal{U}, v \in \mathcal{V}, x \in \mathbb{F}_2^T, y \in \mathbb{F}_2^S$.

Theorem 5. *Given $G \leq \mathbb{S}_{[n]}, H \leq \mathbb{S}_{[m]}$ and $\pi \in \mathbb{S}_{[n]\times[m]}$, $\mathsf{SEN}_{G,H}(\pi) = O(nm)$ if the following conditions are met. Otherwise, $\pi \notin \langle \mathsf{SES}_{G,H} \rangle$.*

1. *π is in $\mathbb{S}_{([n]/G)\times([m]/H)}$;*
2. *$\pi_{U\times[m]}$ is in $\{\mathsf{id}_U\} \otimes H$ for all fix-points $U \in [n]/G$;*
3. *$\pi_{[n]\times V}$ is in $G \otimes \{\mathsf{id}_V\}$ for all fix-points $V \in [m]/H$;*
4. *$\zeta^{([n]/G)\times([m]/H)}(\pi)$ is in $\mathcal{M}_{\zeta^{[n]/G}(G),\zeta^{[m]/H}(H)}$.*

We first focus on a single orbit $U \times V$ of $\langle \mathsf{SES}_{G,H} \rangle$ where $U \in [n]/G, V \in [m]/H$. Lemma 7 provide a useful gadget, which achieves 3-switching in $U \times V$ with 4-norm. Note that an even permutation can be decomposition into 3-cycles. Thus, with Lemma 7, a (G, H)-shuffle-exchange can achieve any even permutation on $U \times V$ and keep the other part fixed, as Corollary 6 says.

Lemma 7. *Let $G \leq \mathbb{S}_{[n]}, H \leq \mathbb{S}_{[m]}$ and $U \in [n]/G, V \in [m]/H$ with $|U|, |V| \geq 2$. For all distinct $p_1, p_2, p_3 \in U \times V$, $\mathsf{SEN}_{G,H}((p_1\ p_2\ p_3)) = O(1)$.*

Proof. Assume $p_1 = (y_1, x_1), p_2 = (y_2, x_2)$ and $p_3 = (y_3, x_3)$. Define σ as follows.

- If $y_1 = y_2$ and $x_1 = x_3$, let $\sigma = \mathsf{id}$.

[2] Given a set S (or an object, like a partition for example, which could be treated as a set), the notation $\mathbb{F}_2^S$ here stands for a space gathering all $|S|$-dimension $\mathbb{F}_2$ vector, in which components are labeled by elements in S.

- If $y_1 = y_2 = y_3$, select $\sigma_1 \in G$ s.t. $\sigma_1(y_3) \neq y_3$ which exists since $|U| \geq 2$. Furthermore, select $\sigma_2 \in H$ s.t. $\sigma_2(x_3) = x_1$, where such σ_2 exists since V is an orbit on H. Define $\sigma = \sigma_2^{\text{row-}\sigma_1(y_3)} \circ \sigma_1^{\text{col-}x_3}$.
- If $y_1 = y_2 \neq y_3$, select $\sigma_1 \in H$ satisfying $\sigma_1(x_3) = x_1$. Define $\sigma = \sigma_1^{\text{row-}y_3}$.
- If y_1, y_2, y_3 are distinct and x_1, x_2, x_3 are distinct, select $\sigma_1 \in G, \sigma_2 \in H$ s.t. $\sigma_1(x_2) = x_1$ and $\sigma_2(y_3) = y_1$. Define $\sigma = \sigma_2^{\text{row-}y_3} \circ \sigma_1^{\text{col-}x_2}$.

The detail shuffle-exchange design depends on the relative position of the 3 positions. We do not list all cases since the others can be reduced to the listed ones directly.

Note that $\sigma(p_1), \sigma(p_2)$ are on the same row, and $\sigma(p_1), \sigma(p_3)$ are on the same column. Select $\tau_1 \in G, \tau_2 \in H$ such that $\tau_1^{\text{row-}y_1}(\sigma(p_2)) = \sigma(p_1)$ and $\tau_2^{\text{col-}x_1}(\sigma(p_3)) = \sigma(p_1)$. It can be verified that

$$(\tau_2^{\text{col-}x_1})^{-1} \circ (\tau_1^{\text{row-}y_1})^{-1} \circ \tau_2^{\text{col-}x_1} \circ \tau_1^{\text{row-}y_1} = (\sigma(p_1)\ \sigma(p_3)\ \sigma(p_2)).$$

Conjugating by σ, we have

$$\sigma^{-1} \circ (\tau_2^{\text{col-}x_1})^{-1} \circ (\tau_1^{\text{row-}y_1})^{-1} \circ \tau_2^{\text{col-}x_1} \circ \tau_1^{\text{row-}y_1} \circ \sigma = (p_1\ p_3\ p_2) = (p_1\ p_2\ p_3)^{-1},$$

which finishes the proof. □

Note all $\pi \in \mathbb{A}_T$ can be decomposed as a composition of $O(|T|)$ 3-cycles. Lemma 7 says a 3-cycle can be achieved with a constant-norm shuffle-exchange, which directly leads to the following corollary.

Corollary 6. *Given $G \leq \mathbb{S}_{[n]}, H \leq \mathbb{S}_{[m]}$ and $U \in [n]/G, V \in [m]/H$ with $|U|, |V| \geq 2$, $\mathsf{SEN}_{G,H}(\pi \oplus \mathsf{id}_{\overline{U \times V}}) = O(|U| \cdot |V|)$ for all $\pi \in \mathbb{A}_{U \times V}$.*

Obviously, $\langle \mathsf{SES}_{G,H} \rangle$ is a subgroup of $\mathbb{S}_{([n]/G) \times ([m]/H)}$. Assume G, H are fix-point-free. Corollary 6 also shows that orbit-wise even π can be archived via performing algorithms used in Corollary 6 on each orbit. Next, we describe orbit-wise parity of a permutation in $\mathbb{S}_{([n]/G) \times ([m]/H)}$ as a vector from $\mathbb{F}_2^{([n]/G) \times ([m]/H)}$, and define a linear space $\mathcal{M}_{\zeta(G),\zeta(H)}$, which is spanned by orbit-wise parity vectors of all permutations in $\mathsf{SES}_{G,H}$. The definition of $\mathcal{M}_{\zeta(G),\zeta(H)}$ leads to the following conclusion: π can be cancel to a permutation even on all orbits, with a (G, H)-shuffle-exchange, if and only if the orbit-wise parity vector $\zeta(\pi)$ is in $\mathcal{M}_{\zeta(G),\zeta(H)}$. Thus, for some π with $\zeta(\pi) \in \mathcal{M}_{\zeta(G),\zeta(H)}$, use such (G, H)-shuffle-exchange to transform it into an orbit-wise even permutation, which can be archived orbit-by-orbit by Corollary 6. For some π with $\zeta(\pi) \notin \mathcal{M}_{\zeta(G),\zeta(H)}$, no such (G, H)-shuffle-exchange works to transform π into an orbit-wise even one, which also says $\pi \notin \mathsf{SES}_{G,H}$ since id is orbit-wise even! We believe the insight about parity fundamentally illusions the construction of $\langle \mathsf{SES}_{G,H} \rangle$ and also leads to the characterization in Theorem 5.

Proof of Theorem 5. In the first step, we assume for π, these conditions are met. We construct a $O(nm)$-norm (G, H)-shuffle-exchange achieving π. At beginning, recall that condition 1 says $\pi \in \mathbb{S}_{([n]/G)\times([m]/H)}$ and condition 4 says $\zeta^{([n]/G)\times([m]/H)}(\pi)$ is in $\mathcal{M}_{\zeta^{[n]/G}(G),\zeta^{[m]/H}(H)}$. Thus, there exist $(v_\sigma)_{\sigma\in G}$ and $(w_\tau)_{\tau\in H}$ such that

$$\zeta(\pi) = \sum_{\sigma\in G} \zeta(\sigma) \otimes v_\sigma + \sum_{\tau\in H} w_\tau \otimes \zeta(\tau).$$

Note that it can be simplified since that G, H are groups and ζ is a linear mapping. That is, there exists $\sigma^{(V)} \in G$ for all $V \in [m]/H$ and $(\tau^{(U)} \in H)_{U\in[n]/G}$ such that

$$\zeta(\pi) = \sum_{V\in[m]/H} \zeta(\sigma^{(V)}) \otimes e_V + \sum_{U\in[n]/G} e_U \otimes \zeta(\tau^{(U)}) \tag{1}$$

where e_U, e_V are indicators (with dimension notation omitted for convenience) s.t. $(e_U)_S = 1$ if and only if $U = S$. Let

$$\pi' = \prod_{U\in[n]/G} (\tau^{(U)})^{\text{row-}U_1} \circ \prod_{V\in[m]/H} (\sigma^{(V)})^{\text{col-}V_1} \circ \pi$$

where U_1, V_1 means arbitrary elements in U, V. Decompose π and obtain

$$\begin{aligned}\pi' &= \prod_{U\in[n]/G} (\tau^{(U)})^{\text{row-}U_1} \circ \prod_{V\in[m]/H} (\sigma^{(V)})^{\text{col-}V_1} \circ \prod_{\substack{U\in[n]/G\\V\in[m]/H}} \pi_{U\times V} \oplus \mathsf{id}_{\overline{U\times V}} \\ &= \prod_{\substack{U\in[n]/G\\V\in[m]/H}} \left((\tau_V^{(U)})^{\text{row-}U_1} \circ (\sigma_U^{(V)})^{\text{col-}V_1} \circ \pi_{U\times V}\right) \oplus \mathsf{id}_{\overline{U\times V}}.\end{aligned}$$

Recall that (1) implies

$$\zeta(\pi)_{U,V} = \zeta(\tau_V^{(U)}) + \zeta(\sigma_U^{(V)}).$$

Meanwhile,

$$\begin{aligned}\zeta\left((\tau_V^{(U)})^{\text{row-}U_1} \circ (\sigma_U^{(V)})^{\text{col-}V_1}\right)_{U,V} &= \zeta\left((\tau_V^{(U)})^{\text{row-}U_1}\right)_{U,V} + \zeta\left((\sigma_U^{(V)})^{\text{col-}V_1}\right)_{U,V} \\ &= \zeta(\tau_V^{(U)}) + \zeta(\sigma_U^{(V)}).\end{aligned}$$

Thus, $\pi'_{U\times V}$ is always even, and it suffices to show $\mathsf{SEN}_{G,H}(\pi') = O(nm)$.

Note that $\pi' \in \mathbb{S}_{([n]/G)\times([m]/H)}$. For any fix-point $U \in [n]/G, V \in [m]/H$ with $|U|, |V| = 1$, $\pi'_{[n]\times V}$ and $\pi'_{U\times[m]}$ are commutative because $U \times V$ is a fix-point of $\mathbb{S}_{([n]/G)\times([m]/H)}$. Thus, we have

$$\pi' = \prod_{\substack{U\in[n]/G, V\in[m]/H \\ |U|,|V|\geq 2}} \pi'_{U\times V} \circ \prod_{\substack{V\in[m]/H \\ |V|=1}} \pi'_{[n]\times V} \circ \prod_{\substack{U\in[n]/G \\ |U|=1}} \pi'_{U\times[m]}.$$

Conditions 2 and 3 ensure that the second and the third term can be achieved within 1-norm, and Corollary 6 says the first term on each orbit can be achieved within $O(|U| \cdot |V|)$-norm. Totally, a $O(nm)$-norm (G, H)-shuffle-exchange achieves π as desired.

In the second step, we show a permutation cannot be achieved if someone among these conditions is not met. When condition 1 not met, there exists $(y, x) \in [n] \times [m]$ mapped to (z, w) such that either y, z are not in the same orbit from $[n]/G$, or x, w are not in the same orbit from $[m]/H$. In this first case for example, $\mathsf{SES}_{\{\mathsf{id}\},H}$ does not change the first coordinate while $\mathsf{SES}_{G,\{\mathsf{id}\}}$ cannot move y out of y-orbit. When condition 2 (which is well-defined when condition 1 met) not met, $\pi_{[n]\times V} \notin G \otimes \{\mathsf{id}_V\}$ for some fix-point $V \in [m]/H$. Since V is a fix-point, $\sigma_{[n]\times V} = \mathsf{id}$ for all $\sigma \in \mathsf{SES}_{\{\mathsf{id}\},H}$, which implies $\langle \mathsf{SES}_{G,H}\rangle_{[n]\times V} = G \otimes \{\mathsf{id}_V\}$. Thus, $\pi_{[n]\times V} \notin G \otimes \{\mathsf{id}_V\}$ implies $\pi \notin \langle \mathsf{SES}_{G,H}\rangle$. For the same reason, $\pi \notin \langle \mathsf{SES}_{G,H}\rangle$ when condition 3 not met. Note that

$$\zeta(\mathsf{SES}_{G,H}) = \zeta(\mathsf{SES}_{G,\{\mathsf{id}\}}) + \zeta(\mathsf{SES}_{\{\mathsf{id}\},H}) = \zeta(G) \otimes \mathbb{F}_2^{[m]/H} + \mathbb{F}_2^{[n]/G} \otimes \zeta(H).$$

Thus, $\zeta(\langle \mathsf{SES}_{G,H}\rangle) = \mathrm{span}(\zeta(\mathsf{SES}_{G,H})) = \mathcal{M}_{\zeta(G),\zeta(H)}$ and $\zeta(\pi) \in \mathcal{M}_{\zeta(G),\zeta(H)}$ is a necessary condition of $\pi \in \langle \mathsf{SES}_{G,H}\rangle$, as condition 4 says. □

6. Conclusion

Throughout this paper, we discuss the problem of achieving a permutation via a low-norm shuffle-exchange from the perspective of the Cayley graph. We give a clean and computationally efficient characterization for connectivity and reachability, as well as a polynomial upper bound for diameter for the shuffle-exchange over general groups. We also focus that on the cyclic group base, and provide a nearly optimal low-norm shuffle-exchange. The work somehow gives a general framework to research the construction of generated groups over 2-dimension set, and explore the properties the Cayley graph in a new direction.

We recognize that our work is only the beginning of the research about the shuffle-exchange. The problems listed as follows could be further directions on this topic.

- For the cyclic group case, there still exists a gap between our upper bound and the counting lower bound. We conjecture our construction may be optimal and the lower bound can be improved by a more careful analysis.
- We believe the cyclic group is the weakest one with respect to the shuffle-exchange. So, we conjecture

$$\mathsf{SEN}_{G,H}(\langle \mathsf{SES}_{G,H} \rangle) = O(\mathsf{SEN}_{\mathbb{C}_n,\mathbb{C}_m}(\langle \mathsf{SES}_{\mathbb{C}_n,\mathbb{C}_m} \rangle))$$

  for all $G \leq \mathbb{S}_{[n]}$ and $H \leq \mathbb{S}_{[m]}$. For the general case, we only provide an upper bound of $O(nm)$, which is so far from optimal in our opinion.
- Furthermore, we wonder whether our model can be realized in reality to accelerate algorithms for permutation rearrangement. In parallel computing, for example, under the memristor model one can perform the same permutation on each row (or column) of a matrix [21]. Currently, there is technique barrier on implementing different permutations on each row (or column) in parallel.

Acknowledgement

We thank Yi Deng, Zhihan Jin, and Mingji Xia for inspiring discussion.

References

[1] László Babai and Ákos Seress. On the diameter of permutation groups. *European Journal of Combinatorics*, **13**(4):231–243, 1992. MR1179520

[2] Xuewen Bao, Frank K. Hwang, and Qiao Li. Rearrangeability of bit permutation networks. *Theoretical Computer Science*, **352**(1-3):197–214, 2006. MR2207517

[3] Václad E. Beneš. Optimal rearrangeable multistage connecting networks. *Bell System Technical Journal*, **43**(4):1641–1656, 1964. MR0167367

[4] V. E. Beneš. Proving the rearrangeability of connecting networks by group calculations. *Bell System Technical Journal*, **54**(2):421–434, 1975. MR0371511

[5] Hasan Cam. Rearrangeability of (2n-1)-stage shuffle-exchange networks. *SIAM Journal on Computing*, **32**(3):557–585, 2003. MR2001744

[6] Arthur Cayley. Desiderata and suggestions: No. 2. the theory of groups: graphical representation. *American Journal of Mathematics*, **1**(2):174–176, 1878. MR1505159

[7] Erik D. Demaine, Martin L. Demaine, Sarah Eisenstat, Anna Lubiw, and Andrew Winslow. Algorithms for solving rubik's cubes. In *Algorithms – ESA 2011 – 19th Annual European Symposium, Saarbrücken, Germany, September 5–9, 2011*. 2011. MR2893242

[8] Erik D. Demaine, Sarah Eisenstat, and Mikhail Rudoy. Solving the rubik's cube optimally is np-complete. In *35th Symposium on Theoretical Aspects of Computer Science*, 2018. MR3779305

[9] Oksana Firman, Philipp Kindermann, Alexander Ravsky, Alexander Wolff, and Johannes Zink. Computing optimal tangles faster. *arXiv preprint* arXiv:1901.06548, 2019. MR4063851

[10] Robert A. Hearn and Erik D. Demaine. *Games, Puzzles, and Computation*. AK Peters/CRC Press, 2009. MR2743249

[11] Harald A. Helfgott and Ákos Seress. On the diameter of permutation groups. *Annals of Mathematics*, **179**:611–658, 2014. MR3152942

[12] Derek F. Holt, Bettina Eick, and Eamonn A. O'Brien. *Handbook of Computational Group Theory*. Chapman and Hall/CRC, 2005. MR2129747

[13] Vadim Lioubimov. Shuffle-exchange conjecture. http://www.openproblemgarden.org/op/shuffle_exchange_conjecture.

[14] Wilhelm Magnus, Abraham Karrass, and Donald Solitar. *Combinatorial Group Theory: Presentations of Groups in Terms of Generators and Relations*. Courier Corporation, 2004. MR2109550

[15] Maya Olszewski, Jeff Meder, Emmanuel Kieffer, Raphaël Bleuse, Martin Rosalie, Grégoire Danoy, and Pascal Bouvry. Visualizing the template of a chaotic attractor. In *International Symposium on Graph Drawing and Network Visualization*, pages 106–119. Springer, 2018. MR3901453

[16] George Pólya. Kombinatorische anzahlbestimmungen für gruppen, graphen und chemische verbindungen. *Acta Mathematica*, **68**(1):145–254, 1937. MR1577579

[17] Janis Pritzkau. Loopover. https://loopover.xyz/.

[18] Daniel Ratner and Manfred Warmuth. The (n^2-1)-puzzle and related relocation problems. *Journal of Symbolic Computation*, **10**(2):111–137, 1990. MR1080669

[19] Tomas Rokicki, Herbert Kociemba, Morley Davidson, and John Dethridge. The diameter of the rubik's cube group is twenty. *SIAM Journal on Discrete Mathematics*, **27**(2):1082–1105, 2013. MR3068558

[20] Charles C Sims. Computational methods in the study of permutation groups. In *Computational Problems in Abstract Algebra*, pages 169–183. Elsevier, 1970. MR0257203

[21] Nishil Talati, Saransh Gupta, Pravin Mane, and Shahar Kvatinsky. Logic design within memristive memories using memristor-aided logic (magic). *IEEE Transactions on Nanotechnology*, **15**(4):635–650, 2016.

Xiaoming Sun
State Key Lab of Processors, Institute of Computing Technology
Chinese Academy of Sciences
Beijing
China
University of Chinese Academy of Sciences
Beijing
China
E-mail: sunxiaoming@ict.ac.cn

Yuan Sun
State Key Lab of Processors, Institute of Computing Technology
Chinese Academy of Sciences
Beijing
China
University of Chinese Academy of Sciences
Beijing
China
E-mail: sunyuan2016@ict.ac.cn

Kewen Wu
University of California at Berkeley
Berkeley
United States
E-mail: shlw_kevin@hotmail.com

Zhiyu Xia
State Key Lab of Processors, Institute of Computing Technology
Chinese Academy of Sciences
Beijing
China
University of Chinese Academy of Sciences
Beijing
China
E-mail: xiazhiyu@ict.ac.cn