**Title**

Mixing Times of the Swap-or-Not and Overlapping Cycles Shuffles



**Author**

Oberschelp, Hans Fountain

**Publication Date**

2023

Peer reviewed|Thesis/dissertation

Mixing Times of the Swap-or-Not and Overlapping Cycles Shuffles

By

HANS OBERSCHELP

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Mathematics

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

---

Ben Morris, Chair

---

Phillip Rogaway

---

Dan Romik

Committee in Charge

2023

# Contents

**Abstract**

This dissertation analyzes two algorithms for shuffling cards: the swap-or-not shuffle and the overlapping cycles shuffle.

The swap-or-not shuffle was developed by Hoang, Morris, and Rogaway [4] as a building block for quick data encryption algorithms with concrete security bounds. In Chapter 1 we introduce concepts from cryptography using the language of probability. We also reproduce an important theorem from cryptography first shown by Maurer, Pietrzak, and Renner [8] which says that a random permutation with a total variation mixing time of $t$ steps can be used to create an encryption algorithm with strong security against chosen ciphertext attacks after $2t$ rounds. We also prove a new theorem that says that a random permutation with a separation mixing time of $t$ steps will create an algorithm with strong chosen ciphertext attack security after only $t$ rounds.

In Chapter 2 we show that for the swap-or-not shuffle on $n$ cards, the separation mixing time of $\sqrt{n}$ of the cards is about $\log_2(n)$. We combine this with our theorem from Chapter 1 to tighten the best known bound on the CCA security of the $n$ card swap-or-not shuffle in the special case of fewer than $\sqrt{n}$ queries.

In Chapter 3 we consider the overlapping cycles shuffle. In each step of the overlapping cycles shuffle on $n$ cards, a fair coin is flipped which determines whether the $m$th card or the $n$th card is moved to the top of the deck. Angel, Peres, and Wilson [1] showed the following interesting fact: If $m = \lfloor \alpha n \rfloor$ where $\alpha$ is rational, then the relaxation time of a single card in the overlapping cycles shuffle is $O(n^2)$. However if $\alpha$ is the inverse golden ratio, then the relaxation time is $O(n^{\frac{3}{2}})$. We show that the mixing time of the entire deck under the overlapping cycles shuffle matches these relaxation times for a single card up to a factor of $\log(n)^3$.

**Acknowledgments**

To begin, I would like to thank my advisor Ben Morris who both introduced me to stochastic processes as an undergraduate, and taught me all the specialized theory present in this thesis as a graduate student. It took many many incorrect proofs for me to grow as a mathematician, and I appreciate Prof. Morris's endless patience in listening to my ideas each week, and helping me filter out errors and refine my logic.

I would also like to thank Hamilton Samraj Santhakumar. When I joined Hamilton's research with Prof. Morris on big key encryption, I knew little about how mathematical research was conducted. At our weekly meetings I was able to observe how Hamilton and Prof. Morris brainstormed ideas, and thanks to Hamilton's excellent example I learned how to build intuition and construct my own proof ideas.

I may not have ever attended graduate school if it were not for the help of Elena Fuchs. During my third year of undergraduate, Prof. Fuchs guided me through the opaque process of graduate admissions, providing me with encouragement, information, advice, and a just-in-time suggestion to apply to REUs. I owe Prof. Fuchs enormous thanks for all her help.

Lastly, I want to thank all of my friends and family for their support throughout the years. I especially want to thank Maureen, Mason, and my parents. In my first week of graduate school, when I came face to face with the sobering reality of functional analysis and I considered dropping out, these four people reassured me and comforted me and got me through that tough first quarter. In Davis I met so many kind and friendly people, both inside and outside the math department, and I want to thank all of them for making Davis such a wonderful and comfortable place to live.

CHAPTER 1

# A Probabilistic Proof of the nCPA to CCA Bound

In this chapter we provide a new proof of a result in cryptography by Maurer, Pietrzak and Renner [8]. A cipher is a system for encrypting data so that two parties, Alice and Bob, can send messages to each other without fear of a third party eavesdropping. If Alice wants to send a message to Bob, Alice will use an encryption algorithm to encrypt her plaintext message into ciphertext, and then send the ciphertext to Bob. Then Bob will use a decryption algorithm to decrypt to cyphertext back to the original message.

For a cipher to be "good" it should have the property that, if an adversary were to steal the ciphertext in transit, they would have a hard time learning much of anything about the original message. To quantify this notion of security against an adversary, cryptographers use various statistics. Today we focus on two: non-adaptive plaintext attack advantage (nCPA advantage), and chosen ciphertext attack advantage (CCA advantage). In both of these standards, we allow an adversary to make $q$ requests for plaintext-ciphertext pairs, which we call "queries". We then ask if the adversary can tell the difference between an honest response to their queries (where we provide the true plaintext-ciphertext pairs) and a dishonest response (where we give them junk data that we randomly generate). If they can't tell which is which, then we can be confident the cipher is secure. This is because if they knew how to encrypt and decrypt even partial messages they could use that to distinguish between honest and dishonest responses.

The difference between the nCPA and CCA standards has to do with the rules describing how the adversary is allowed to ask for their queries. The CCA standard allows for strictly more powerful adversaries, as it gives the adversary all the allowances of the nCPA standard along with some additional ones. While the CCA standard is more strict, it can be significantly more difficult to show that a cipher is secure in the CCA standard versus the nCPA standard. However, Maurer

Pietrzak and Renner [8] showed the following useful result: Suppose $P$ and $Q$ are encryption algorithms which have strong security in the nCPA standard. Then, if $P^{-1}$ is the decryption algorithm associated with $P$, we can create a new encryption algorithm $Z = P^{-1} \circ Q$ by running $Q$ and then $P^{-1}$. In this case, $Z$ will have strong security in the CCA standard.

It turns out that nCPA advantage is closely related to *total variation distance*, a statistical distance from probability theory. To create a cipher with strong CCA security, it is enough to analyze the total variation distance of a cipher $H$. If the total variation distance is small, then we know that $H$ has strong nCPA security. Then, we can let $P$ and $Q$ be independent copies of $H$ (i.e. ciphers using the same algorithm as $H$ but with independently generated keys). By setting $Z = P^{-1} \circ Q$ we have a cipher with strong CCA security. The only problem is this $Z$ will have double the runtime of $H$, because $Z$ must go through both the encryption and decryption algorithms for $H$. This may not be necessary, because it might be the case that $H$ has strong CCA security from the start.

We provide a new proof of Maurer, Pietrzak's and Renner's [8] result. Our proof uses probability directly, as opposed to information theory, and has the advantage of providing an alternate sufficient condition of low CCA advantage. Namely, the CCA advantage of a random permutation can be bounded by its *separation distance* from the uniform distribution (much like how nCPA advantage is related to total variation distance). In practice this means that some ciphers, with small separation distance from uniform, can be run twice as quickly as previously known. The proof is roughly split into two parts. First we prove a technical lemma about Markov chains, which, when translated into the language of cryptography, states that separation distance is small when two ciphers with strong nCPA security are composed. Then we show that CCA security is strong when separation distance is small.

### 1.1. Definitions

We begin by introducing the definitions of nCPA and CCA security with a game. This will help to provide context for our mathematical definitions of nCPA and CCA security, which we provide afterwards.

Imagine you have two machines, Machine H and Machine T. Machine H generates a uniformly random permutation $U$ of the numbers $\{1, \ldots, n\}$. You can query machine H by inputting any one of the numbers $\{1, \ldots, n\}$. If you input 5 then machine $A$ will output $U(5)$ i.e. the number that $U$ permutes 5 to. Machine H only generates $U$ one time so if you input 5 again you will get the same output, and if you input 7 you will get a different output than the one from 5.

Machine T works exactly as Machine H except that it independently generates a random permutation $X$ according to some pre-established distribution of your choosing.

You play a game against an opponent we will call the "adversary", or $A$ for short. At the start of the game you flip a fair coin. Then, the adversary provides you will a sequence of $q$ queries, which are numbers they want to input into one of the machines. If you flipped Heads at the start of the game, input their queries into Machine H and tell the adversary the results. If you flipped Tails, input their queries into Machine T and tell the adversary the results. Now the adversary guesses if you flipped Heads or Tails. We say $\text{nCPA}_{q,A}(X)$ is the **non-adaptive chosen plaintext attack advantage** of $A$ against random permutation $X$ and we define this such that

$$2 \cdot \text{nCPA}_{q,A}(X) - 1$$

is the probability of $A$ winning the game. Note that $\text{nCPA}_{q,A}(X)$ is normalized so that $\text{nCPA}_{q,A}(X) = 1$ if $A$ always wins the game and $\text{nCPA}_{q,A}(X) = 0$ if $A$ utilizes the naive strategy of always guessing Heads. We say

$$\text{nCPA}_q(X) = \max_A \{\text{nCPA}_{q,A}(X)\}$$

Note that $\text{nCPA}_q(X)$ is close to 0 if the distribution of $X$ is close to the uniform distribution.

Now we define $\text{CCA}_q(X)$ or the **chosen ciphertext attack advantage** against $X$. This is defined in exactly the same way as $\text{nCPA})_q(X)$ except with two rule changes to the game:

- The adversary can make some or all of their queries to the inverse permutation. Specifically, the adversary can provide a number $c \in \{1, \ldots, n\}$ and specify that they want a "reverse query" and you must provide them with $U^{-1}(c)$ if you flipped Heads or $X^{-1}(c)$ if you flipped tails.

- The adversary is allowed to provide their queries one at a time, adapting their choice of next query based on the information they have received. For example, the adversary may first ask for a reverse query of the number 5. When they are provided with the number 3 as the response, they may use that to decide that they want to query the number 2 in the normal forwards direction as their second query. This continues until they have exhausted all $q$ of their queries.

Note that the CCA advantage against $X$ must be higher than the nCPA advantage against $X$. This is because the adversary has strictly more tools at their disposal in the CCA version of the game, and so an optimal adversary will have a better chance at distinguishing $X$ from $U$. In fact, there are examples of distributions for $X$ where the nCPA advantage is close to 0 and the CCA advantage is close to 1.

We now redefine both nCPA and CCA advantage in the language of probability. These definitions will be equivalent to the ones described above.

DEFINITION 1.1.1. *If two finite random variables $X$ and $Y$ take in the same set $\mathcal{V}$, they have* ***total variation distance*** *given by*

$$d_{\mathrm{TV}}(X,Y) = \sup_{A\subset\mathcal{V}} \Big(\mathbb{P}(X\in A) - \mathbb{P}(Y\in A)\Big).$$

*We can equivalently define total variation distance by*

$$\begin{aligned} d_{\mathrm{TV}}(X,Y) &= \frac{1}{2}\sum_{a\in\mathcal{V}} \Big|\mathbb{P}(X=a) - \mathbb{P}(Y=a)\Big| \\ &= \sum_{a\in\mathcal{V}} \Big(\mathbb{P}(X=a) - \mathbb{P}(Y=a)\Big)^{+}. \end{aligned}$$

Note that total variation distance is a metric, and in particular $d_{\mathrm{TV}}(X,Y) = d_{\mathrm{TV}}(Y,X)$.

DEFINITION 1.1.2. *If two finite random variables, $X$ and $Y$, are defined taking values in the same set $\mathcal{V}$, the* ***separation distance*** *from $X$ to $Y$ is given by*

$$d_{\mathrm{sep}}(X,Y) = \sup_{a\in\mathcal{V}} \left(1 - \frac{\mathbb{P}(X=a)}{\mathbb{P}(Y=a)}\right)$$

*where* $\frac{x}{0} := 1$.

Note that separation distance is not a metric, as $d_{\text{sep}}(X, Y)$ does not necessarily equal $d_{\text{sep}}(Y, X)$.

DEFINITION 1.1.3. *For two finite sets, $S$ and $\mathcal{V}$, let $\mathcal{X} = \{X(i) : i \in S\}$ be a collection of random variables, all taking values in $\mathcal{V}$, and let $Y$ be another random variable taking values in $\mathcal{V}$. Then we define*

$$d_{\text{TV}}(\mathcal{X}, Y) = \max_{i \in S} \ d_{\text{TV}}(X(i), Y).$$

*For separation distance we similarly define*

$$d_{\text{sep}}(\mathcal{X}, Y) = \max_{i \in S} \ d_{\text{sep}}(X(i), Y).$$

DEFINITION 1.1.4. *Let $X$ be a random permutation of length $n$. Let $S_q$ be the set of all ordered $q$-tuples of $\{1, \ldots, n\}$. For $p = (p_1, \ldots, p_q) \in S_q$, Let $X(p)$ be the random vector $\Big(X(p_1), \ldots, X(p_q)\Big)$. Let $\mu$ be a uniform random element of $S_q$. Let $\mathcal{X}$ be the set of all $X(p)$ for $p \in S_q$. The* ***nCPA-security*** *of $X$ with $q$ queries is defined by*

$$\text{nCPA}_\text{q}(\text{X}) = \text{d}_{\text{TV}}(\mathcal{X}, \mu).$$

Note $\text{nCPA}_n(X) = d_{\text{TV}}(X, U)$ where $U$ is the uniform random permutation.

We will encode CCA queries to a permutation as a string of the form "number, arrow, number". For example, the notation $3 \to 5$ will be used if an adversary queries the image of 3 and $\pi(3) = 5$. The notation $7 \leftarrow 2$ will be used if an adversary queries the preimage of 7 and $\pi(2) = 7$.

DEFINITION 1.1.5. *We will define $\mathcal{N}_n$ to be the space of CCA queries to a permutation of length $n$. Specifically, let $\mathcal{N}_n$ be the following set of 3-symbol strings,*

$$\mathcal{N}_n := \{aRb \ : \ a \in \{1, \ldots, n\}, R \in \{\to, \leftarrow\}, b \in \{1, \ldots, n\}\}.$$

*We call the first two symbols of $p \in \mathcal{N}_n$ the* ***input****, which we denote $I(p)$. For example, $I(3 \to 5) = 3 \to$. We call the last entry the* ***output****, which we denote $O(p)$. For example, $O(7 \leftarrow 2) = 2$.*

*For $a, b \in \{1, \ldots, n\}$ we say $a \to b$ and $b \to a$ are* ***reversals*** *of each other. We say two CCA queries $p_1$ and $p_2$ are* ***equivalent*** *(and we write $p_1 \sim p_2$) if $p_1 = p_2$ or $p_1$ and $p_2$ are reversals of*

*each other.*

Note that $\sim$ gives an equivalence relation on $\mathcal{N}_n$.

DEFINITION 1.1.6. *A function $f : S_n \to {\mathcal{N}_n}^q$ is called a $q$-query CCA* ***strategy*** *if for every $k \in \{1, \ldots, n\}$ and $\sigma, \tau \in S_n$ the following statements hold:*

(1) if $f(\sigma)_k \sim (a \to b)$ then $\sigma(a) = b$,

(2) $I(f(\cdot)_1)$ is constant, i.e. $I(f(\sigma)_1)$ does not depend on $\sigma$;

(3) if $(f(\sigma)_1, \ldots, f(\sigma)_{k-1}) = (f(\tau)_1, \ldots, f(\tau)_{k-1})$, then $I(f(\sigma)_k) = I(f(\tau)_k)$.

A strategy is a way an adversary might make $q$ queries to an unknown permutation. At first the adversary knows nothing, so the question of the first query does not depend on the permutation. The first question the adversary asks is "where does this permutation (or, if the adversary so chooses, the inverse of this permutation) send the element $a$?" The result of the first query tells the adversary the answer to this question. Then, the adversary's second question can be based on the information gained by the first query. The adversary's third question can be based on the information gained by the first two queries, and so forth.

DEFINITION 1.1.7. *Let $X$ be a random permutation of length $n$. The* ***CCA-security*** *of $X$ with $q$ queries is given by*

$$CCA_q(X) = \max_{f \text{ is a } q\text{-query strategy}} d_{\mathrm{TV}}(f(X), f(U)),$$

*where $U$ is the uniform random permutation of length $n$.*

## 1.2. Main Theorem

Now that we have defined nCPA and CCA security, we can state the main theorem of this chapter. This is part of Maurer Pietrzak and Renner's [8] Corollary 2 on page 2.

THEOREM 1.2.1. *Let $X, Y$ be random permutations of length $n$. Let $q \in \{1, \ldots, n\}$. Then*

$$\mathrm{CCA}_q(X^{-1} \circ Y) \leq \mathrm{nCPA}_q(X) + \mathrm{nCPA}_q(Y)$$

The rest of this chapter will be devoted to proving this theorem. Along the way we will also prove another upper bound on the CCA security, which is the following:

THEOREM 1.2.2. *Let $X$ be a random permutation of length $n$. Let $S$ be the set of all ordered $q$-tuples of $\{1, \ldots, n\}$. Let $\mathcal{X} := \{X(p)\}_{p \in S}$. Let $\mu_q$ be the uniform distribution on $S$. Then,*

$$CCA_q(X) \leq d_{\text{sep}}(\mathcal{X}, \mu).$$

## 1.3. Technical Lemmas

In this section we prove some technical lemmas regarding Markov chains and random permutations. In the first two results we show an upper bound for separation distance of the composition of two Markov chains in terms of the total variation distances of the individual chains.

LEMMA 1.3.1. *Let $P, Q$ be Markov chains on state space $S$ where $S$ is finite, and suppose $P, Q$ both have stationary distribution $\pi$. Let $\overleftarrow{P}$ be the time reversal of $P$. Then for all $i, j \in S$,*

$$1 - \frac{Q\overleftarrow{P}(i,j)}{\pi(j)} \leq d_{\text{TV}}(P(j,\cdot), \pi) + d_{\text{TV}}(Q(i,\cdot), \pi).$$

PROOF. Fix any $i, j \in S$. Then,

$$Q\overleftarrow{P}(i,j) = \sum_{z \in S} Q(i,z) \cdot \overleftarrow{P}(z,j) \tag{1.1}$$

$$= \sum_{z \in S} Q(i,z) \cdot \frac{\pi(j)}{\pi(z)} \cdot P(j,z) \tag{1.2}$$

$$= \pi(j) \sum_{z \in S} \frac{P(j,z)}{\pi(z)} \cdot \frac{Q(i,z)}{\pi(z)} \cdot \pi(z),$$

where (1.1) comes from conditioning on the state after the Q step, and (1.2) uses the definition of the time reversal. Let

$$\Delta_P(z) := \frac{P(j,z) - \pi(z)}{\pi(z)},$$

$$\Delta_Q(z) := \frac{Q(i,z) - \pi(z)}{\pi(z)}.$$

Then

$$\frac{P(j,z)}{\pi(z)} = 1 + \Delta_P(z),$$
$$\frac{Q(i,z)}{\pi(z)} = 1 + \Delta_Q(z),$$

and hence

$$\begin{aligned}\frac{Q\overleftarrow{P}(i,j)}{\pi(j)} &= \sum_{z\in S}(1+\Delta_P(z))\cdot(1+\Delta_Q(z))\cdot\pi(z)\\ &= \sum_{z\in S}\pi(z) + \sum_{z\in S}\Delta_P(z)\pi(z) + \sum_{z\in S}\Delta_Q(z)\pi(z) + \sum_{z\in S}\Delta_P(z)\Delta_Q(z)\pi(z). \qquad (1.3)\end{aligned}$$

Since $\pi(z)$ is a probability vector we have $\sum_{z\in S}\pi(z) = 1$. Furthermore, $P(j,\cdot)$ is also a probability vector so

$$\sum_{z\in S}\Delta_P(z)\pi(z) = \sum_{z\in S}(P(j,z) - \pi(z)) = 0.$$

Similarly, $\sum_{z\in S}\Delta_Q(z)\pi(z) = 0$. To bound the final sum in (1.3) note that

$$\sum_{z\in S}\Delta_P(z)\Delta_Q(z)\pi(z) \geq \sum_{z\in S} -\Big(\Delta_P(z)\Delta_Q(z)\pi(z)\Big)^{-}. \qquad (1.4)$$

For every nonzero term on the right hand side of (1.4), either $\Delta_P(z) > 0$ and $\Delta_Q(z) < 0$, or $\Delta_Q(z) > 0$ and $\Delta_P(z) < 0$. This gives us

$$\begin{aligned}\sum_{z\in S} -\Big(\Delta_P(z)\Delta_Q(z)\pi(z)\Big)^{-} &= \sum_{\substack{\Delta_P(z)>0\\ \Delta_Q(z)<0}}\Delta_P(z)\Delta_Q(z)\pi(z) \;+ \sum_{\substack{\Delta_Q(z)>0\\ \Delta_P(z)<0}}\Delta_P(z)\Delta_Q(z)\pi(z)\\ &\geq \sum_{\substack{\Delta_P(z)>0\\ \Delta_Q(z)<0}} -\Delta_P(z)\pi(z) \;+ \sum_{\substack{\Delta_Q(z)>0\\ \Delta_P(z)<0}} -\Delta_Q(z)\pi(z), \qquad (1.5)\end{aligned}$$

where (1.5) comes from the fact that $\Delta_Q(z) \geq -1$ and $\Delta_P(z) \geq -1$ for all $z$. Finally, note that

$$\begin{aligned}\sum_{\substack{\Delta_P(z)>0\\ \Delta_Q(z)<0}} -\Delta_P(z)\pi(z) &\geq \sum_{\Delta_P(z)>0} -\Delta_P(z)\pi(z)\\ &= \sum_{P(j,z)>\pi(z)} -(P(j,z) - \pi(z))\\ &= -d_{\mathrm{TV}}(P(j,\cdot),\pi). \qquad (1.6)\end{aligned}$$

A similar argument shows that

$$\sum_{\substack{\Delta_Q(z)>0\\ \Delta_P(z)<0}} -\Delta_Q(z)\pi(z) \geq -d_{\mathrm{TV}}(Q(i,\cdot),\pi) \tag{1.7}$$

Combining (1.6) and (1.7) with (1.4) and (1.5) gives,

$$\frac{Q\overleftarrow{P}(i,j)}{\pi(j)} \geq 1 - d_{\mathrm{TV}}(P(j,\cdot),\pi) - d_{\mathrm{TV}}(Q(i,\cdot),\pi)$$

and the lemma follows. □

COROLLARY 1.3.2. *Let $P, Q$ be Markov chains on a finite state space $S$, both with the stationary distribution $\pi$. Let $\overleftarrow{P}$ be the time reversal of $P$. Let $\mathcal{P} := \{P(i,\cdot)\}_{i\in S}$, $\mathcal{Q} := \{Q(i,\cdot)\}_{i\in S}$, and $\mathcal{Q}\overleftarrow{\mathcal{P}} := \{Q\overleftarrow{P}(i,\cdot)\}_{i\in S}$. Then,*

(1) for all $i \in S$ we have $d_{\mathrm{sep}}(Q\overleftarrow{P}(i,\cdot),\pi) \leq d_{\mathrm{TV}}(\mathcal{P},\pi) + d_{\mathrm{TV}}(Q(i),\pi)$

(2) $d_{\mathrm{sep}}(\mathcal{Q}\overleftarrow{\mathcal{P}},\pi) \leq d_{\mathrm{TV}}(\mathcal{P},\pi) + d_{\mathrm{TV}}(\mathcal{Q},\pi)$.

PROOF. By Lemma 8, for all $i, j \in S$,

$$1 - \frac{Q\overleftarrow{P}(i,j)}{\pi(j)} \leq d_{\mathrm{TV}}(P(j,\cdot),\pi) + d_{\mathrm{TV}}(Q(i,\cdot),\pi).$$

Taking the maximum of both sides over $j$ gives

$$\max_{j\in S}\left[1 - \frac{Q\overleftarrow{P}(i,j)}{\pi(j)}\right] \leq \max_{j\in S}\Big[d_{\mathrm{TV}}(P(j,\cdot),\pi) + d_{\mathrm{TV}}(Q(i,\cdot),\pi)\Big],$$

$$d_{\mathrm{sep}}(Q\overleftarrow{P}(i,\cdot),\pi) \leq d_{\mathrm{TV}}(\mathcal{P},\pi) + d_{\mathrm{TV}}(Q(i,\cdot),\pi).$$

This is our first result. Now we take the maximum of both sides over $i$ and have

$$\max_{i\in S}\ d_{\mathrm{sep}}(Q\overleftarrow{P}(i,\cdot),\pi) \leq \max_{i\in S}\Big[d_{\mathrm{TV}}(\mathcal{P},\pi) + d_{\mathrm{TV}}(Q(i,\cdot),\pi)\Big],$$

$$d_{\mathrm{sep}}(\mathcal{Q}\overleftarrow{\mathcal{P}},\pi) \leq d_{\mathrm{TV}}(\mathcal{P},\pi) + d_{\mathrm{TV}}(\mathcal{Q},\pi).$$

□

The next three results show that CCA advantage is bounded above by separation distance. We will later combine this fact with the prior results from this section to achieve a bound on CCA security

in terms of nCPA security. It is also a useful fact in its own right because it gives us a tight bound on CCA security using a well-studied metric from probability.

LEMMA 1.3.3. *Let $\sigma$ be a permutation of length $n$. Let $f$ be a CCA strategy with $q$ queries. Let $p = (p_1, \ldots, p_q) \sim (a_1 \to b_1, \ldots, a_q \to b_q) \in \mathcal{N}^q$, and suppose that $p$ is in the image of $f$. Then,*

$$f(\sigma) = p \text{ if and only if } \sigma(a_1, \ldots, a_q) = (b_1, \ldots, b_q)$$

PROOF. First we assume $f(\sigma) = p$. Then $f(\sigma) \sim (a_1 \to b_1, \ldots, a_q \to b_q)$. The definition of a strategy requires $\sigma(a_i) = b_i$ for all $i$. So if $f(\sigma) = p$ then $\sigma(a_1, \ldots, a_q) = (b_1, \ldots, b_q)$.

Now we assume $f(\sigma) = \ell = (\ell_1, \ldots, \ell_q) \neq p = (p_1, \ldots, p_q)$. Let $m = \min\{i : \ell_i \neq p_i\}$. Note that for all $j < m$ we have $\ell_j = p_j$. This, along with $p$ and $\ell$ being in the image of the same strategy, means $I(\ell_m) = I(p_m)$. This implies $O(\ell_m) \neq O(p_m)$. We now consider two cases:

- **Case 1**

  If $p_m = a_m \to b_m$, then $\ell_m = a_m \to c_m$ where $c_m \neq b_m$. So $\sigma(a_m) = c_m \neq b_m$.
- **Case 2**

  If $p_m = b_m \leftarrow a_m$, then $\ell_m = b_m \leftarrow d_m$ where $d_m \neq a_m$. So $\sigma(a_m) \neq \sigma(d_m) = b_m$ because $\sigma$ is a permutation.

Either way $\sigma(a_m) \neq b_m$, hence if $f(\sigma) \neq p$ then $\sigma(a_1, \ldots, a_q) \neq (b_1, \ldots, b_q)$. □

COROLLARY 1.3.4. *Let $f$ be a CCA strategy with $q$ queries. Let $\Phi \subset \mathcal{N}_n{}^q$ be the image of $f$. Let $S$ be the set of all ordered $q$-tuples of distinct elements of $\{1, \ldots, n\}$. Then there is a one-to-one correspondence between $\Phi$ and a subset $H_f \subset S^2$ where each $p \in \Phi$ is matched with $(a, b) \in H_f$ such that*

$$f(\sigma) = p \text{ if and only if } \sigma(s_1) = s_2.$$

PROOF. By Lemma 1.3.3 we already know that for each $p \in \Phi$ there exists $(a, b) \in S^2$ such that

$$f(\sigma) = p \text{ if and only if } \sigma(s_1) = s_2.$$

All that remains is to show that this mapping is injective. Suppose $p, p' \in \Phi$ such that $p \neq p'$. Let $k$ be the minimal value of $\{1, \ldots, q\}$ such that $p_k \neq p'_k$. Since $p, p'$ are both in the image of $f$ we have that $I(p_1) = I(p'_1)$. In addition, if $k \geq 2$ then we know that $(p_1, \ldots, p_{k-1}) = (p'_1, \ldots, p'_{k-1})$ and since $p, p'$ are in the image of $f$ we have $I(p_k) = I(p'_k)$. Without loss of generality assume that $I(p_k), I(p'_k)$ both take the form $a \to$ for some $a \in \{1, \ldots, n\}$. Since $p_k \neq p'_k$ there must exist $b, b' \in \{1, \ldots, b\}$ such that $b \neq b'$ and

$$p_k = a \to b \text{ and } p'_k = a \to b'.$$

So the following statements hold:

- if $f(\sigma) = p$ then $\sigma(a) = b$,
- if $f(\sigma) = p'$ then $\sigma(a) = b'$.

Therefore $s, s' \in S^2$ associated with $p, p'$ respectively cannot be the same. □

THEOREM 1.2.2. *Let $X$ be a random permutation of length $n$. Let $S$ be the set of all ordered $q$-tuples of $\{1, \ldots, n\}$. Let $\mathcal{X} := \{X(p)\}_{p \in S}$. Let $\mu_q$ be the uniform distribution on $S$. Then,*

$$CCA_q(X) \leq d_{\text{sep}}(\mathcal{X}, \mu).$$

PROOF. Fix some $q$-query strategy $f$. Assume $f$ is optimal (total variation distance maximizing). Let $\Phi \subset {\mathcal{N}_n}^q$ be the image of $f$. By the optimality of $f$ we can assume that there does not exist $p \in \Phi$ such that $p_i \sim p_j$ for any $i \neq j$. (This is because no optimal strategy would ever ask a question it already knows the answer to. In other words, if $p_i \sim (3 \to 5)$, then no optimal strategy would ask $(3 \to)$ or $(5 \leftarrow)$ as $I(p_j)$ for $j > i$.)

First we compute $|\Phi|$. We can count all $p \in \Phi$ as follows: There is,

$$\begin{aligned}
1 \quad & \text{possible value of } I(p_1), \\
n \quad & \text{possible values of } O(p_1), \\
1 \quad & \text{possible value of } I(p_2), \text{ given } p_1 \\
(n-1) \quad & \text{possible values of } O(p_2) \text{ given } p_1 \text{ and } I(p_2), \\
& \vdots \\
1 \quad & \text{possible value of } I(p_q) \text{ given } p_1, \ldots, p_{q-1}, \\
(n-q+1) \quad & \text{possible values of } O(p_q) \text{ given } p_1, \ldots, p_{q-1} \text{ and } I(p_q).
\end{aligned}$$

So $|\Phi| = n(n-1)\ldots(n-q+1) =: (n)_q$. We will set this result aside for now.

Using the definition of total variation distance,

$$d_{\mathrm{TV}}(f(X), f(U)) = \sum_{p \in \Phi} \Big[\mathbb{P}(f(U) = p) - \mathbb{P}(f(X) = p)\Big]^+.$$

Lemma 1.3.3 tells us that for each $p \in \Phi$ there exists $(a, b) = ((a_1, \ldots, a_q), (b_1, \ldots, b_q)) \in S^2$ such that

$$\Big[\mathbb{P}(f(U) = p) - \mathbb{P}(f(X) = p)\Big] = \Big[\mathbb{P}(U(a) = b) - \mathbb{P}(X(a) = b)\Big].$$

Let $H_f$ be the set of all such $(a, b)$. Then by Corollary 1.3.4 we have $|H_f| = |\Phi| = (n)_q$ and

$$\begin{aligned}
d_{\mathrm{TV}}(f(X), f(U)) &= \sum_{(a,b) \in H_f} \Big[\mathbb{P}(U(a) = b) - \mathbb{P}(X(a) = b)\Big]^+ \\
&= \sum_{(a,b) \in H_f} \left[\frac{1}{(n)_q} - \mathbb{P}(X(a) = b)\right]^+ \\
&= \frac{1}{(n)_q} \sum_{(a,b) \in H_f} \left[1 - \frac{\mathbb{P}(X(a) = b)}{{(n)_q}^{-1}}\right]^+.
\end{aligned}$$

If we replace each term in the sum with the maximum over all $(a,b) \in H_f$, we get the inequality

$$\begin{aligned} d_{\mathrm{TV}}(f(X), f(U)) &\leq \frac{1}{(n)_q} |H_f| \max_{(a,b)\in H_f} \left| 1 - \frac{\mathbb{P}(X(a)=b)}{(n)_q{}^{-1}} \right| \\ &= \max_{(a,b)\in H_f} \left| 1 - \frac{\mathbb{P}(X(a)=b)}{(n)_q{}^{-1}} \right| \\ &\leq \max_{(a,b)\in S^2} \left| 1 - \frac{\mathbb{P}(X(a)=b)}{(n)_q{}^{-1}} \right|. \end{aligned} \tag{1.8}$$

Using the definition of separation distance we can rewrite (1.8) as

$$\begin{aligned} d_{\mathrm{TV}}(f(X), f(U)) &\leq \max_{a\in S} d_{\mathrm{sep}}(X(a), U(a)) \\ &= d_{\mathrm{sep}}(\mathcal{X}, \mu) \end{aligned}$$

Since this inequality holds for all strategies $f$, we get

$$\mathrm{CCA}_q(X) \leq d_{\mathrm{sep}}(\mathcal{X}, \mu)$$

□

### 1.4. Proof of Main Theorem

We now have all the tools necessary to prove the nCPA to CCA bound, the main result of this chapter.

THEOREM 1.2.1. *Let $X, Y$ be random permutations of length $n$. Let $q \in \{1, \ldots, n\}$. Then*

$$\mathrm{CCA}_q(X^{-1} \circ Y) \leq \mathrm{nCPA}_q(X) + \mathrm{nCPA}_q(Y)$$

PROOF. This is a straightforward application of Corollary 1.3.2 and Theorem 1.2.2. Let $S$ be the set of all ordered $q$-tuples of distinct elements of $\{1, \ldots, n\}$. Let $\mathcal{X} := \{X(p)\}_{p\in S}$ and $\mathcal{Y} := \{Y(p)\}_{p\in S}$ and $\mathcal{X}^{-1}\mathcal{Y} := \{X^{-1} \circ Y(p)\}_{p\in S}$. Let $\mu_q$ be the uniform distribution on $S$. Then from Theorem 1.3.2 we have

$$\mathrm{CCA}_q(X^{-1} \circ Y) \leq d_{\mathrm{sep}}(\mathcal{X}^{-1}\mathcal{Y}, \mu_q). \tag{1.9}$$

We can think of $X$ and $Y$ each as one step of a Markov Chain on $S_n$. Then by Corollary 1.3.2 we have,

$$d_{\text{sep}}(\mathcal{X}^{-1}\mathcal{Y}, \mu) \leq d_{\text{TV}}(\mathcal{X}, \mu_q) + d_{\text{TV}}(\mathcal{Y}, \mu_q). \tag{1.10}$$

By applying the definition of $\text{nCPA}_q$ the right hand side of (1.10) we have

$$d_{\text{sep}}(\mathcal{X}^{-1}\mathcal{Y}, \mu) \leq \text{nCPA}_q(X) + \text{nCPA}_q(Y). \tag{1.11}$$

Combining (1.9) and (1.11) completes the theorem. □

CHAPTER 2

# Mixing Time of the Swap-or-Not Shuffle

In the previously chapter we showed how certain ciphers, that have a small separation distance from the uniform distribution, can be run twice as fast as previously known. In this chapter we apply this result to the swap-or-not shuffle.

The swap-or-not shuffle was created by Hoang, Morris, and Rogaway [4] as a card-shuffling algorithm that lends itself to quickly encrypting messages. In particular, Hoang, Morris, and Rogaway built a cipher based on the swap-or-not shuffle that achieved the best known security bounds for a fast cipher on medium sized message domains, like credit card numbers. They determined the total variation distance of the swap-or-not shuffle and used it to show that their cipher had strong nCPA security. Then they used the result by Maurer, Pietrzak and Renner [8] to show that after running the cipher for twice as long, it had strong CCA security.

Their argument showed that, in a message space of size $N$, the swap-or-not shuffle can achieve strong CCA security after approximately $r = 6\log_2(N)$ rounds when the number of queries is less than $N^{1-\epsilon}$. In 2017, Dai, Hoang, and Tessaro [2] improved the bound, and showed that only $r = 4\log_2(N)$ rounds are required. In this section, we show that approximately $r = \log_2(N)$ rounds is sufficient provided that the number of queries is less than $\sqrt{N}$. Our bound comes from analyzing the separation distance of the swap-or-not shuffle, and applying Theorem 1.2.2. Our upper bound on the number of rounds required for strong security is tight when the number of queries is more than $\log_2(N)$. This demonstrates the usefulness of Theorem 1.2.2, because such a tight bound would not be possible with the runtime-doubling technique of composing two copies of the shuffle.

## 2.1. Definition of the Swap on Not Shuffle

The swap-or-not shuffle is a random permutation defined as follows: We start with a deck of $N = 2^d$ cards, and a collection of vectors $\mathcal{K}_1, \ldots, \mathcal{K}_r \in \mathbb{Z}_2^d$ which we call *round keys*. First, label

each card as a unique element of $\mathbb{Z}_2^d$. The specific labeling is not important to the security of the shuffle, so for simplicity we will label cards by their initial position in the deck, in binary. So in a shuffle of 16 cards, card 0010 is initially in position 0010 (or the 3rd topmost card). In round $j$, let the cards in positions $x$ and $y$ be "paired" with respect to round key $\mathcal{K}_j$ if $x + y = \mathcal{K}_j$ (where addition is done in $\mathbb{Z}_2^d$). Then for each pair flip an independent coin, and if Heads swap the positions of the cards in the pair, and if Tails do nothing. Repeat this for $r$ independent rounds.

Denote $x^t(\mathcal{K}_1, \ldots, \mathcal{K}_t)$ as the random element of $\mathbb{Z}_2^d$ which is the position of card $x$ (i.e. the card initially in position $x$) after $t$ steps of the shuffle using round keys $\mathcal{K}_1, \ldots, \mathcal{K}_t$. Let $x^t := x^t(K_1, \ldots, K_t)$ where $K_1, \ldots, K_t$ are iid uniformly sampled from $\mathbb{Z}_2^d$. So $x^0 = x$. For $y \in \mathbb{Z}_2^d$ we write $x \to y$ for the event $x^r = y$.

We say the swap-or-not shuffle on $N$ cards with $r$ rounds is the random permutation given by $(x_1^r, \ldots, x_N^r)$.

### 2.2. The Security of the Swap-or-Not Shuffle

The goal of this chapter is to show that the swap-or-not shuffle on $N$ cards has a strong CCA security after about $\log_2(N)$ rounds. Our main result is as follows:

THEOREM 2.2.1. *Fix any $d \geq 2$. Let $X$ be the swap-or-not shuffle with $N = 2^d$ cards, and $r$ rounds. Consider a CCA adversary equipped with $q$ queries up against this swap-or-not shuffle. The security of $X$ against this adversary is bounded by*

$$CCA_q(X) \leq \frac{q^2}{N} + 2^{-r+d} + \frac{rq(q-1)(9 + 48 \cdot 2^{-r+d})}{4(N-2)}$$

We claim that this result means that, as long as the number of queries is fewer than $\sqrt{N}$, it will only take slightly longer than $\log_2(N)$ rounds for the CCA advantage to be small. One way to see this is with some examples. In the following graph we consider the cases $d = 64, r = 96$ and $d = 96, r = 128$ and $d = 128, r = 160$. Note that in each case $r = d + 32$. We graph the logarithm of $CCA_q(X)$ against the logarithm of the number of queries. In each case we can see that the security is strong until the number of queries starts to approach $2^{\frac{1}{2}d} = \sqrt{N}$.

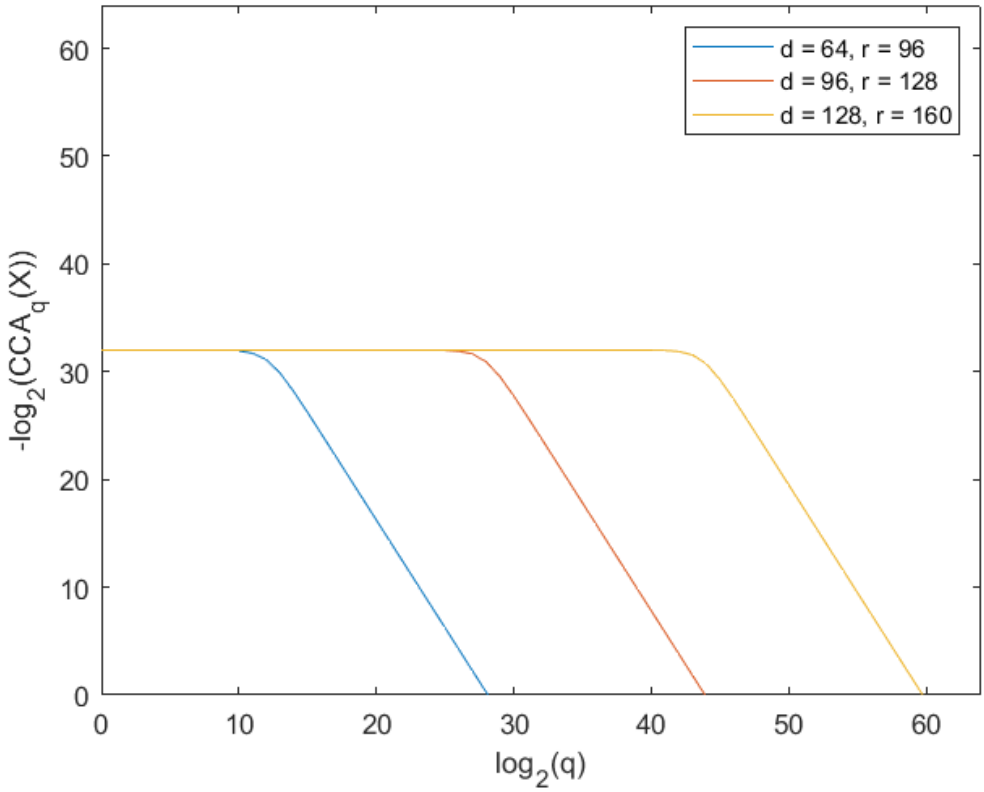


FIGURE 1. Comparison of swap-or-not CCA advantage to number of queries

If a specific advantage, $\epsilon$, is desired, we offer the following corollary to determine the optimal number of rounds and the maximum allowable queries.

THEOREM 2.2.2. *Fix any $\epsilon \in (0,1)$ and $d \geq 2$. Let $X$ be the swap-or-not shuffle with $N = 2^d$ cards, and $r \geq d - \log_2(\epsilon)$ rounds. Consider a CCA adversary equipped with $q \leq \sqrt{\epsilon \cdot \frac{N-2}{r}}$ queries up against this swap-or-not shuffle. The security of $X$ against this adversary is bounded by*

$$CCA_q(X) \leq \frac{13}{4}\epsilon + 12\epsilon^2$$

PROOF. By Theorem 2.2.1 we have that

$$CCA_q(X) \leq \frac{q^2}{N} + 2^{-r+d} + \frac{rq(q-1)(9+48\cdot 2^{-r+d})}{4(N-2)}.$$

Note that as $r \geq d - \log_2(\epsilon)$ we have that $2^{-r+d} \leq \epsilon$. Also note that as $q \leq \sqrt{\epsilon \cdot \frac{N-2}{r}}$ we have that $q^2 \leq \frac{N-2}{r}$. So,

$$\begin{aligned} CCA_q(X) &\leq \frac{q^2}{N} + 2^{-r+d} + \frac{rq(q-1)(9+48\cdot 2^{-r+d})}{4(N-2)} \\ &\leq \frac{q^2}{N-2} + 2^{-r+d} + \frac{rq^2(9+48\cdot 2^{-r+d})}{4(N-2)} \\ &\leq \frac{\epsilon(N-2)}{r(N-2)} + \epsilon + \frac{r\epsilon(N-2)(9+48\cdot\epsilon)}{4r(N-2)} \\ &\leq \frac{13}{4}\epsilon - 12\epsilon^2. \end{aligned}$$

□

The remainder of this chapter is devoted to proving the main theorem.

### 2.3. Collisions and the Tilde Process

We will fix a set of $q$ cards, with initial positions $x_1, x_2, \ldots, x_q$. So $(x_1^r, x_2^r, \ldots, x_q^r)$ is the random vector of positions of these $q$ cards after $r$ rounds. Define coins $c_{i,j}$ as follows:

$$c_{i,j} = \begin{cases} 1 \text{ if card } i \text{ if swapped in round } j \\ 0 \text{ otherwise} \end{cases}$$

Then for round keys $\mathcal{K}_1, \ldots, \mathcal{K}_r$ we have

$$\begin{aligned} x_1^r &= x_1 + c_{1,1}\mathcal{K}_1 + c_{1,2}\mathcal{K}_2 + \cdots + c_{1,r}K_r \\ x_2^r &= x_2 + c_{2,1}\mathcal{K}_1 + c_{2,2}\mathcal{K}_2 + \cdots + c_{2,r}\mathcal{K}_r \\ &\vdots \\ x_q^r &= x_q + c_{q,1}\mathcal{K}_1 + c_{q,2}\mathcal{K}_2 + \cdots + c_{q,r}\mathcal{K}_r \end{aligned}$$

Note that the coins $c_{i,j}$ are not independent. In particular, if $x_i^{t-1} + \mathcal{K}_t = x_j^{t-1}$ then $c_{i,t} = c_{j,t}$. We can see that the round keys $\mathcal{K}_1, \ldots, \mathcal{K}_r$ need to span $\mathbb{Z}_2^d$ to make $x_1^r, \ldots, x_q^r$ close to uniform. Otherwise each $x_i^r - x_i$ will be in the same subspace of $\mathbb{Z}_2^d$, which would be very unlikely for a uniform random permutation. Fortunately, it is very likely $K_1, \ldots, K_r$ span $\mathbb{Z}_2^d$ as long as $r$ is slightly larger than $d$. We will now make this precise.

LEMMA 2.3.1. *Fix $r \geq d$. Let $A_r$ be the event that $K_1, \ldots, K_r$ span $\mathbb{Z}_2^d$. Then*

$$\mathbb{P}(A_r) \geq 1 - 2^{d-r}$$

PROOF. For any $v \in \mathbb{Z}_2^d$ let $H_v$ be the event that $v$ is orthogonal to each of $K_1, \ldots, K_r$. Then,

$$A_r^C = \bigcup_{v \neq 0} H_v.$$

So,

$$\mathbb{P}(A_r^C) = \mathbb{P}\left(\bigcup_{v \neq 0} H_v\right) \leq \sum_{v \neq 0} \mathbb{P}(H_v).$$

For each $v \neq 0$, we have $\mathbb{P}(H_v) = 2^{-r}$, as each $K_i$ is independently in or out of the plane $v^\perp$ with probability $\frac{1}{2}$ each. Since there are $2^d - 1$ different vectors in the sum,

$$\mathbb{P}(A_r^C) \leq \sum_{v \in \mathbb{Z}_2^d} 2^{-r} \leq 2^{d-r}$$

□

The fact that the round keys are likely to span $\mathbb{Z}_2^d$ after $r$ rounds when $r$ is larger than $d$ should give us hope that the swap-or-not shuffle will be well-mixed after $r$ rounds. Indeed, we would know that the swap-or-not shuffle was perfectly mixed if only the coins $c_{i,j}$ were all independent. With this idea in mind, our strategy to prove the swap-or-not shuffle is well-mixed will proceed as follows:

- First we define a new process, which is similar to swap-or-not shuffle but has independent coins.
- Then we show that this new process is uniform as long as the round keys span $\mathbb{Z}_2^d$
- Finally we couple the swap-or-not shuffle to this new process in such a way that it is likely to stay coupled for all $r$ rounds.

Our new process will be a variation on the swap-or-not shuffle, and is not strictly speaking a shuffle (that is, it is not a random permutation). Start with a deck of $n = 2^d$ cards, labeled by their initial

positions in the deck. As before, if round keys $\mathcal{K}_1, \ldots, \mathcal{K}_r \in \mathbb{Z}_2^d$, then let

$$\widetilde{x_1^r}(\mathcal{K}_1, \ldots, \mathcal{K}_t) = x_1 + \widetilde{c_{1,1}}\mathcal{K}_1 + \widetilde{c_{1,2}}\mathcal{K}_2 + \cdots + \widetilde{c_{1,r}}\mathcal{K}_r$$

$$\widetilde{x_2^r}(\mathcal{K}_1, \ldots, \mathcal{K}_t) = x_2 + \widetilde{c_{2,1}}\mathcal{K}_1 + \widetilde{c_{2,2}}\mathcal{K}_2 + \cdots + \widetilde{c_{2,r}}\mathcal{K}_r$$

$$\vdots$$

$$\widetilde{x_q^r}(\mathcal{K}_1, \ldots, \mathcal{K}_t) = x_q + \widetilde{c_{q,1}}\mathcal{K}_1 + \widetilde{c_{q,2}}\mathcal{K}_2 + \cdots + \widetilde{c_{q,r}}\mathcal{K}_r$$

where $\widetilde{c_{i,j}}$ are iid Bernoulli($\frac{1}{2}$) random variables. In other words, if $x$ and $x + \mathcal{K}_j$ are paired, then instead of swapping places or remaining put with probability $\frac{1}{2}$ each, now $x$ and $x + \mathcal{K}_j$ will *both* go to $x$, or *both* go to $x + \mathcal{K}_j$, or swap, or stay put, with probability $\frac{1}{4}$ each. We call this process the tilde process (and we keep in mind it is not a random permutation because it is not necessarily injective). As before, we write $\widetilde{x^t}(\mathcal{K}_1, \ldots, \mathcal{K}_t)$ as the (random) position of card $x$ under the tilde process after $t$ steps using round keys $\mathcal{K}_1, \ldots, \mathcal{K}_t$. Let $\widetilde{x^t}$ be defined similarly but with iid uniform round keys. We write $x \widetilde{\to} y$ for the event $\widetilde{x^r} = y$.

LEMMA 2.3.2. *Fix any $x_1, \ldots, x_q, y_1, \ldots, y_q \in \mathbb{Z}_2^d$. Also fix any $\mathcal{K}_1, \ldots, \mathcal{K}_r \in \mathbb{Z}_2^d$ with $r \geq d$ such that $\mathcal{K}_1, \ldots, \mathcal{K}_r$ span $\mathbb{Z}_2^d$. Consider the tilde process on $\mathbb{Z}_2^d$ with $r$ rounds. Let $K_1, \ldots, K_d$ be the iid uniform round keys, and let $\widetilde{c_{i,j}}$ be the coins. Then,*

*(1) For all $t$ the distribution of $\left(\widetilde{x_1^t}(\mathcal{K}_1, \ldots, \mathcal{K}_t) + x_1, \ldots, \widetilde{x_q^t}(\mathcal{K}_1, \ldots, \mathcal{K}_t) + x_q\right)$ is uniform over $(\text{span}(\mathcal{K}_1, \ldots, \mathcal{K}_t))^q$.*

*(2) $\mathbb{P}(x_1 \widetilde{\to} y_1, \ldots x_q \widetilde{\to} y_q \mid K_1 = \mathcal{K}_1, \ldots, K_r = \mathcal{K}_r) = 2^{-qd}$*

*(3) $\mathbb{P}(\widetilde{c_{i,j}} = \mathcal{C}_{i,j}$ for all $i, j \mid x_1 \widetilde{\to} y_1, \ldots x_q \widetilde{\to} y_q,\ K_1 = \mathcal{K}_1, \ldots, K_r = \mathcal{K}_r) = 2^{q(d-r)}$ for all $(\mathcal{C}_{i,j})$ where $x_i + \mathcal{C}_{i,1}\mathcal{K}_1 + \cdots + \mathcal{C}_{i,r}\mathcal{K}_r = y_i$ for all $i$. In other words if the round keys span $\mathbb{Z}_2^d$ then the coins are uniformly distributed across all "valid" choices that take each $x_i$ to $y_i$.*

PROOF. We prove (1) induction. For the base case, note that by definition, each $x_i^1 + x_i = \mathcal{K}_1^{c_{1,i}}$. Since $c_{1,1}, \ldots, c_{1,q}$ are independent, each $x_i^1 + x_i$ is independently equally likey to equal 0 or $\mathcal{K}_1$.

For the inductive step, assume $\left(\widetilde{x_1^t}(\mathcal{K}_1, \ldots, \mathcal{K}_t) + x_1, \ldots, \widetilde{x_q^t}(\mathcal{K}_1, \ldots, \mathcal{K}_t) + x_q\right)$ is distributed uniformly across $\text{span}(\mathcal{K}_1, \ldots, \mathcal{K}_t)$. In the case that $\mathcal{K}_{t+1} \in \text{span}(\mathcal{K}_1, \ldots, \mathcal{K}_t)$, adding $c_{i,t+1}\mathcal{K}_{t+1}$ to

each $\widetilde{x_i^t}$ amounts to adding a vector in the subspace $\text{span}(\mathcal{K}_1, \ldots, \mathcal{K}_t)^q$ to a uniform random element of that subspace, and so the distribution will remain uniform. In the case that that $\mathcal{K}_{t+1} \notin \text{span}(\mathcal{K}_1, \ldots, \mathcal{K}_t)$, the $\mathcal{K}_{t+1}$ component of each $\widetilde{x_i^{t+1}}$ will equally likely be present or absent independently and the component orthogonal to $\mathcal{K}_{t+1}$ will remain uniform, so the distribution of $\left(\widetilde{x_1^{t+1}}(\mathcal{K}_1, \ldots, \mathcal{K}_{t+1}) + x_1, \ldots, \widetilde{x_q^{t+1}}(\mathcal{K}_1, \ldots, \mathcal{K}_{t+1}) + x_q\right)$ will be uniform over $\text{span}(\mathcal{K}_1, \ldots, \mathcal{K}_t)$

Now (2) follows immediately from (1) after setting $t = r$ and recalling that by assumption that $|\text{span}(\mathcal{K}_1, \ldots, \mathcal{K}_r)| = |\mathbb{Z}_2^d| = 2^{-d}$.

To show (3), fix any $\mathcal{C}_{i,j} \in \{0,1\}$ such that $\widetilde{c_{i,j}} = \mathcal{C}_{i,j}$ and $K_j = \mathcal{K}_j$ for all $i, j$ imply $x_i \widetilde{\to} y_i$ for all $i$. Then,

$$\begin{aligned}
&\mathbb{P}(\widetilde{c_{i,j}} = \mathcal{C}_{i,j} \text{ for all } i,j \mid x_1 \widetilde{\to} y_1, \ldots x_q \widetilde{\to} y_q,\ K_1 = \mathcal{K}_1, \ldots, K_r = \mathcal{K}_r) \\
&= \frac{\mathbb{P}(\widetilde{c_{i,j}} = \mathcal{C}_{i,j} \text{ for all } i,j \mid K_1 = \mathcal{K}_1, \ldots, K_r = \mathcal{K}_r)}{\mathbb{P}(x_1 \widetilde{\to} y_1, x_2 \widetilde{\to} y_2, \ldots, x_q \widetilde{\to} y_q \mid K_1 = \mathcal{K}_1, \ldots, K_r = \mathcal{K}_r)} \qquad (2.1) \\
&= \frac{2^{-qr}}{2^{-qd}} \qquad (2.2)
\end{aligned}$$

where we have used that $\widetilde{c_{i,j}} = \mathcal{C}_{i,j}$ for all $i, j$ implies $x_1 \widetilde{\to} y_1, \ldots x_q \widetilde{\to} y_q$ in line (2.1), and that the coins are independent of the round keys to compute the numerator of (2.2). This completes the lemma. □

We showed earlier that when the round keys are chosen uniformly, they are likely to span $\mathbb{Z}_2^d$. This fact combined with the above lemma means that the tilde process has a near-uniform distribution. So, if we can show that the swap-or-not shuffle has a distribution similar to that of the tilde process, we can show that the swap-or-not shuffle is close to the uniform distribution. To do this we couple the tilde process with the swap-or-not shuffle as follows:

Fix $x_1, \ldots, x_q$ and $\mathcal{K}_1, \ldots, \mathcal{K}_r$. Set $x_i^0 = \widetilde{x_i^0} = x_i$. Then generate $\{\widetilde{c_{i,t}}\}$ as iid Bernoulli($\frac{1}{2}$) random variables. This defines the tilde process as described above. Now inductively define

$$c_{i,t} = \begin{cases} c_{j,t} \text{ if } x_j^{t-1} + x_i^{t-1} = \mathcal{K}_t \text{ for some } j < i \\ \widetilde{c_{i,t}} \text{ otherwise} \end{cases}$$

The $c_{i,j}$ define the swap-or-not shuffle, as $c_{i,t} = c_{j,t}$ if cards $x_i$ and $x_j$ are paired in round $t$ as required, and otherwise they are independent Bernoulli($\frac{1}{2}$) random variables.

DEFINITION 2.3.3. *In the tilde process, we say cards $x_i$ and $x_j$ have a* ***collision at time*** *$t$ if $K_t = \widetilde{x_i^{t-1}} + \widetilde{x_j^{t-1}}$ and $(\widetilde{c_{i,t}}, \widetilde{c_{j,t}}) \in \{(1,0),(0,1)\}$. In a tilde process with $r$ rounds, we say $x_i$ and $x_j$ have a* ***collision*** *if they have a collision at time $t$ for any $1 \le t \le r$.*

That is, $x_i$ and $x_j$ have a collision at time $t$ if $x_i$ "moves" to the same position as $x_j$ or vice versa. Note that it is possible to have two cards occupy the same position at time $t$ without a collision at time $t$ if $\widetilde{x_i^{t-1}} = \widetilde{x_j^{t-1}}$ and $\widetilde{c_{i,t}} = \widetilde{c_{j,t}}$. However, if $\widetilde{x_i^t} = \widetilde{x_j^t}$ then we know that at some time up to and including $t$ the cards $i$ and $j$ collided.

Collisions are important because they are the result of a non-injective step and cause the tilde process to "decouple" from the swap-or-not shuffle. We can show that in the absence of collisions the swap-or-not shuffle will stay coupled to the tilde process.

LEMMA 2.3.4. *Consider the coupled swap-or-not shuffle and tilde process. Fix cards $x_1, \ldots, x_q$. Let $M$ be the event that in the tilde process there is at least one collision involving any of these $q$ cards. Then*

$$\text{on the event } M^C \text{ we have } c_{i,t} = \widetilde{c_{i,t}} \text{ for all } 1 \le i \le q,\ 0 \le t \le r$$

*where $c_{i,t}$ and $\widetilde{c_{i,t}}$ are the coins used in the swap-or-not shuffle and tilde process respectively.*

PROOF. Suppose there exists some $i, t$ such that $c_{i,t} \neq \widetilde{c_{i,t}}$. Then let $i', t'$ be chosen so that $c_{i',t'} \neq \widetilde{c_{i',t'}}$ and so that $t'$ is minimal. Then for all $j \in \{1, \ldots, q\}$ and all times $s < t'$ we have $c_{j,s} = \widetilde{c_{j,s}}$. In particular this means that $x_j^{t'-1} = \widetilde{x_j^{t'-1}}$ for all cards $x_j$.

Since $c_{i',t'} \neq \widetilde{c_{i',t'}}$ there must exist some $j' < i'$ such that $x_{j'}^{t'-1} + x_{i'}^{t'-1} = \mathcal{K}_{t'}$. Since $j'$ is paired with $i'$, and $j'$ is the lesser of the pair, we know that $c_{j',t'} = \widetilde{c_{j',t'}}$. According to our coupling we have $c_{i',t'} = \widetilde{c_{j',t'}}$. Since $c_{i',t'} \neq \widetilde{c_{i',t'}}$ we know

$$\widetilde{c_{i',t'}} \neq \widetilde{c_{j',t'}}.$$

In addition, as $x_{i'}^{t'-1} = \widetilde{x_{i'}^{t'-1}}$ and $x_{j'}^{t'-1} = \widetilde{x_{j'}^{t'-1}}$ we have

$$\widetilde{x_{j'}^{t'-1}} + \widetilde{x_{i'}^{t'-1}} = \mathcal{K}_{t'}.$$

So cards $x_{i'}$ and $x_{j'}$ collide in round $t'$ of the tilde process. □

COROLLARY 2.3.5. *Consider the coupled swap-or-not shuffle and tilde process. Fix cards $x_1, \ldots, x_q$. Let $M$ be the event that the tilde process has any pairwise collisions between any of these $q$ cards. Then,*

$$\mathbb{P}(x_1 \to y_1, \ldots, x_q \to y_q) \geq \mathbb{P}(x_1 \widetilde{\to} y_1, \ldots, x_q \widetilde{\to} y_q, M^C)$$

PROOF. As we showed in Lemma 2.3.4, the event $M^C$ implies that $c_{i,t} = \widetilde{c_{i,t}}$ for all $i, t$. So, $M^C$ also implies that for all $i$ we have

$$\widetilde{x_i^r} = x_1 + K_1^{\widetilde{c_{i,1}}} + K_2^{\widetilde{c_{i,2}}} + \cdots + K_r^{\widetilde{c_{i,r}}} = x_i + K_1^{c_{i,1}} + K_2^{c_{i,2}} + \cdots + K_r^{c_{i,r}} = x_i^r.$$

So,

$$\mathbb{P}(x_1 \to y_1, \ldots, x_q \to y_q, M^C) = \mathbb{P}(x_1 \widetilde{\to} y_1, \ldots, x_q \widetilde{\to} y_q, M^C).$$

Hence

$$\begin{aligned}\mathbb{P}(x_1 \to y_1, \ldots, x_q \to y_q) &\geq \mathbb{P}(x_1 \to y_1, \ldots, x_q \to y_q, M^C) \\ &= \mathbb{P}(x_1 \widetilde{\to} y_1, \ldots, x_q \widetilde{\to} y_q, M^C)\end{aligned}$$

□

## 2.4. Collisions are Unlikely

We have shown that if collisions are unlikely then the distribution of the swap-or-not shuffle is close to the distribution of the well-mixed tilde process. This section is devoted to showing that collisions are in fact unlikely.

PROPOSITION 2.4.1. *Consider the tilde process on $N = 2^d$ cards with $r \geq d$ rounds. Fix any $x_i, x_j \in \mathbb{Z}_2^d$. Let $M_{i,j}$ be the event that $x_i$ and $x_j$ have a collision. Then for all $y_i, y_j \in \mathbb{Z}_2^d$ such that*

*$y_i \neq y_j$ we have,*

$$\mathbb{P}(x_i \widetilde{\to} y_i, x_j \widetilde{\to} y_j \mid M_{i,j}) \leq \frac{7 + 48 \cdot 2^{d-r}}{2(N-1)(N-2)}.$$

PROOF. Let $\tau$ be the final time that $x_i$ and $x_j$ collide in the first $r$ rounds. If $x_i$ and $x_j$ do not collide in the first $r$ round, set $\tau = \infty$. A collision between $x_i$ and $x_j$ at time $t$, given the values of $\widetilde{x_i^{t-1}}$ and $\widetilde{x_j^{t-1}}$ happens when $K_t = \widetilde{x_i^{t-1}} + \widetilde{x_j^{t-1}}$ and $\widetilde{c_{i,t}} \neq \widetilde{c_{j,t}}$ with probability $\frac{1}{2} \cdot 2^{-d}$. Note that this probability is the same regardless of the values of $\widetilde{x_i^{t-1}}$ and $\widetilde{x_j^{t-1}}$, a collision at time $t$ is independent of $K_1, \ldots, K_{t-1}$ and independent of all coins before time $t$. Let $R_t$ be the filtration recording $K_1, \ldots, K_t$ and all $c_{k,s}$ with $s \leq t$. Then,

$$\mathbb{P}(x_i \text{ and } x_j \text{ collide in round } t \mid R_{t-1}) = \frac{1}{2} \cdot 2^{-d}.$$

independent of $R_{t-1}, x_i, x_j$. Thus, if we condition on $\tau = T$ for some $T \leq r$ then the trajectory of $x_i$ and $x_j$ can be described as follows:

- From round 1 to $T-1$, the round keys and coins for $x_i$ and $x_j$ are chosen uniformly and independently.
- In round $T$, the round key is set equal to $\widetilde{x_i^{T-1}} + \widetilde{x_j^{T-1}}$. The coin for $x_i$ in round $T$ is still equally likely to flip heads or tails, but the coin for $x_j$ is fixed to be the opposite. This guarantees $\widetilde{x_i^T} = \widetilde{x_j^T}$.
- For a round $s$ between $T+1$ and $r$, the round key and coins are chosen uniformly from all options except $\big(K = \widetilde{x_i^{s-1}} + \widetilde{x_j^{s-1}}$ and $(\widetilde{c_{i,s}}, \widetilde{c_{j,s}}) \in \{(1,0),(0,1)\}\big)$.

We can break the possible trajectories into cases.

Let $B_i$ be the event that $x_i$'s coins flip tails in all rounds strictly *before* $T$, with a similar definition for $B_j$. Let $F_i$ be the event that $x_i$'s coins flip tails in all rounds strictly *after* $T$, with a similar definition for $F_j$. We are concerned with finding upper bounds for the probability $\mathbb{P}(x_i \widetilde{\to} y_i, x_j \widetilde{\to} y_j \mid M_{i,j})$ for all $y_i \neq y_j$, and we will do this by considering the following cases:

(1) $E_1 = F_i \cap F_j$

In this case, neither $x_i$ nor $x_j$ move from their shared position after $T$, so for all $y_i \neq y_j$,

$$\mathbb{P}(x_i \widetilde{\to} y_i, x_j \widetilde{\to} y_j \mid E_1, \tau = T) = 0.$$

(2) $E_2 = B_i^C \cap B_j^C$

In this case, both $x_i$ and $x_j$ move before $T$. Since the shuffle before the collision uses uniform keys, the locations $x_i$ and $x_j$ move to are uniform (although not necessarily independent, as they may have moved in the same round). Therefore, when they collide at time $T$, their shared position will be uniform, regardless of if $x_i$ or $x_j$ is the one to flip heads. So, due to symmetry, all values of $(y_i, y_j)$ with $y_i \neq y_j$ are equally likely outcomes for $(\widetilde{x_i^r}, \widetilde{x_j^r})$. Therefore, for all $y_i \neq y_j$,

$$\mathbb{P}(x_i \widetilde{\to} y_i, x_j \widetilde{\to} y_j \mid E_2, \tau = T) = \frac{\mathbb{P}\left(\widetilde{x_i^r} \neq \widetilde{x_j^r} \mid E_2, \tau = T\right)}{N(N-1)} \leq \frac{1}{N(N-1)}$$

(3) $E_3 = \left((F_i^C \cap F_j) \cup (F_i \cap F_j^C)\right) \cap (B_i \cup B_j)$

On the event $E_3$ exactly one of the cards exclusively flips tails after $T$. If $\widetilde{x_i^T} = \widetilde{x_j^T} = v$ then on the event $F_i^C \cap F_j$ we have $\widetilde{x_j^r} = v$. By symmetry, all positions other than $v$ are equally likely values for $\widetilde{x_i^r}$. So for all $y_i \neq v$ we have,

$$\begin{aligned}
&\mathbb{P}\left(x_i \widetilde{\to} y_i, x_j \widetilde{\to} v \mid \widetilde{x_i^T} = \widetilde{x_j^T} = v, F_i^C \cap F_j, B_i \cup B_j, \tau = T\right) \\
&= \mathbb{P}\left(x_i \widetilde{\to} y_i \mid \widetilde{x_i^T} = \widetilde{x_j^T} = v, F_i^C \cap F_j, B_i \cup B_j, \tau = T\right) \\
&\leq \frac{\mathbb{P}\left(\widetilde{x_i^r} \neq v \mid \widetilde{x_i^T} = \widetilde{x_j^T} = v, F_i^C \cap F_j, B_i \cup B_j, \tau = T\right)}{N-1} \leq \frac{1}{N-1}.
\end{aligned}$$

Furthermore, note that on the events $x_j \widetilde{\to} y_j$ and $F_j$ and $\tau = T$, it must be the case that $\widetilde{x_i^T} = \widetilde{x_j^T} = y_j$. Therefore for all $y_i \neq y_j$ we have

$$\begin{aligned}
&\mathbb{P}\left(x_i \widetilde{\to} y_i, x_j \widetilde{\to} y_j \mid F_i^C \cap F_j, B_i \cup B_j, \tau = T\right) \\
&\leq \mathbb{P}\left(x_i \widetilde{\to} y_i, x_j \widetilde{\to} y_j \mid \widetilde{x_i^T} = \widetilde{x_j^T} = y_j, F_i^C \cap F_j, B_i \cup B_j, \tau = T\right) \qquad (2.3) \\
&\leq \frac{1}{N-1}
\end{aligned}$$

where line (2.3) comes from the fact that if $A, Z$ are events with $A \subset Z$ then $\mathbb{P}(A) \leq \mathbb{P}(A \mid Z)$. The same argument works for $F_i \cap F_j^C$ so we have

$$\mathbb{P}\left(x_i \widetilde{\to} y_i, x_j \widetilde{\to} y_j \mid F_i^C \cap F_j, B_i \cup B_j, \tau = T\right) \leq \frac{1}{N-1} \text{ for all } y_i \neq y_j.$$

Therefore, by the union bound,

$$\mathbb{P}\left(x_i \widetilde{\to} y_i, x_j \widetilde{\to} y_j \mid E_3, \tau = T\right) \leq \frac{2}{N-1} \text{ for all } y_i \neq y_j.$$

(4) $E_4 = F_i^C \cap F_j^C$

In this case, both cards flip heads at some point after round $T$. We split this case into three subcases. Let $G$ be the event that $x_i$ and $x_j$ have their first post-$T$ head flip at the same time. Let $H$ be the event that, at the the first time after $T$ a card moves, the round key is the zero vector. Condition on $\widetilde{x_i^T} = \widetilde{x_j^T} = v$. Consider the following subcases:

(a) Conditioning on $E_4 \cap G$

In this subcase, there exists a round $L > T$, where $x_i$ and $x_j$ both flip tails for all rounds between $T$ and $L$, and then both flips heads in round $L$. Since both $x_i$ and $x_j$ flip heads in round $L$, the distribution of the round $L$ key is uniform even after conditioning on $\tau = T$. This effectively puts us in Case 2, as $x_i$ and $x_j$ move together using the round $L$ key to a uniform random position. As in Case 2, due to symmetry, for all $y_i \neq y_j$,

$$\mathbb{P}(x_i \widetilde{\to} y_i, x_j \widetilde{\to} y_j \mid E_4, G, \widetilde{x_i^T} = \widetilde{x_j^T} = v, \tau = T) \leq \frac{1}{N(N-1)}.$$

(b) Conditioning on $E_4 \cap G^C \cap H$

In this subcase we note that due to symmetry, all targets of the form $(y_i, y_j)$ where $y_i = v$ or $y_j = v$ are equally likely. Similarly, all targets of the form $(y_i, y_j)$ where $y_i \neq y_j$ and $y_i, y_j \neq v$ are equally likely. So for all $y_i \neq y_j$ with $y_i = v$ or $y_j = v$,

$$\begin{aligned}
&\mathbb{P}(x_i \widetilde{\to} y_i, x_j \widetilde{\to} y_j \mid E_4, G^C, H, \widetilde{x_i^T} = \widetilde{x_j^T} = v, \tau = T\ ) \\
&= \frac{\mathbb{P}\left(\widetilde{x_i^r} = v \text{ or } \widetilde{x_j^r} = v, \widetilde{x_i^r} \neq \widetilde{x_j^r} \mid E_4, G^C, H, \widetilde{x_i^T} = \widetilde{x_j^T} = v, \tau = T\right)}{2(N-1)} \\
&\leq \frac{1}{2(N-1)}, \qquad (2.4)
\end{aligned}$$

and for all $y_i \neq y_j$ with $y_i, y_j \neq v$,

$$\begin{aligned}
&\mathbb{P}(x_i \widetilde{\to} y_i, x_j \widetilde{\to} y_j \mid E_4, G^C, H, \widetilde{x_i^T} = \widetilde{x_j^T} = v, \tau = T\ ) \\
&= \frac{\mathbb{P}\left(\widetilde{x_i^r} \neq v \text{ and } \widetilde{x_j^r} \neq v, \widetilde{x_i^r} \neq \widetilde{x_j^r} \mid E_4, G^C, H, \widetilde{x_i^T} = \widetilde{x_j^T} = v, \tau = T\right)}{(N-1)(N-2)} \\
&\leq \frac{1}{(N-1)(N-2)}.
\end{aligned}$$

Since line (2.4) provides the higher upper bound, we have for all $y_i \neq y_j$ that,

$$\mathbb{P}(x_i \widetilde{\to} y_i, x_j \widetilde{\to} y_j \mid E_4, G^C, H, \widetilde{x_i^T} = \widetilde{x_j^T} = v, \tau = T\ ) \leq \frac{1}{2(N-1)}.$$

For future reference, note that

$$\mathbb{P}\left(H, G^C \mid E_4, \widetilde{x_i^T} = \widetilde{x_j^T} = v, \tau = T\right) \leq \mathbb{P}\left(H \mid E_4, \widetilde{x_i^T} = \widetilde{x_j^T} = v, \tau = T\right) = \frac{1}{N}.$$

(c) Conditioning on $E_4 \cap G^C \cap H^C$

In this subcase, there exists a round $L > T$ where $x_i$ and $x_j$ both flip tails between rounds $T$ and $L$, and in round $L$ either $x_i$ flips heads and $x_j$ flips tails or vice versa. Suppose first that $x_i$ flips heads in round $L$. Then $x_i$ is sent to a uniform position other than $v$. Let $u = \widetilde{x_i^L} \neq v$. It may be the case that $x_i$ flips heads some more times before $x_j$ flips heads, and further moves around, but its position will still be uniform amongst states other than $v$, so without loss of generality assume $x_i$ is still at $u$ in the round before $x_j$ flips heads. When $x_j$ flips heads, all positions other than $u$ are equally likely destinations for $x_j$. Since we are conditioning on no collisions, $x_j$ is half as likely to be sent to $u$ as anywhere else, and if $x_j$ is in fact sent to $u$, we know $x_i$ will swap with $x_j$ and be sent back to $v$. Therefore at time $L$, all positions with $x_i \neq v$ are equally likely, and positions with $x_i = v$ are half as likely.

If we instead suppose that $x_j$ flips heads in round $L$, then by the same argument all positions with $x_j \neq v$ are equally likely and positions with $x_j = v$ are half as likely.

Overall this means that, after $x_i$ and $x_j$ have each had their turn to flip heads,

all positions with $x_i, x_j \neq v$ are equally likely, and positions with $x_i = v$ or $x_j = v$ are less likely. Now that after the "flipping heads after $T$" condition has been met for both $x_i$ and $x_j$, the rest of the shuffle is a standard tilde process except for continuing to condition on no collisions. For the rest of the shuffle, our only bias in round keys is against those that pair $x_i$ and $x_j$'s positions, and force $x_i$ and $x_j$ to swap when they are paired. However, we already have symmetry in probability between states $(a, b)$ and $(b, a)$, regardless of if $a$ or $b$ equals $v$. Therefore, just as in the standard tilde process with collisions allowed, starting with states of the form $(v, b)$ and $(a, v)$ less likely means these states will still be less likely after the final round $r$. So for all $y_i \neq y_j$,

$$\begin{aligned}
&\mathbb{P}\left(x_i \widetilde{\to} y_i, x_j \widetilde{\to} y_j \mid E_4, G^C, H^C, \widetilde{x_i^T} = \widetilde{x_j^T} = v, \tau = T\right) \\
\leq\ &\frac{\mathbb{P}\left(v \neq \widetilde{x_i^r} \neq \widetilde{x_j^r} \neq v \mid E_4, G^C, H^C, \widetilde{x_i^T} = \widetilde{x_j^T} = v, \tau = T\right)}{(N-1)(N-2)} \\
\leq\ &\frac{1}{(N-1)(N-2)}
\end{aligned}$$

Now we can combine the three subcases.

$$\begin{aligned}
&\mathbb{P}\left(x_i \widetilde{\to} y_i, x_j \widetilde{\to} y_j \mid E_4, \widetilde{x_i^T} = \widetilde{x_j^T} = v, \tau = T\right) \\
=\ &\mathbb{P}\left(x_i \widetilde{\to} y_i, x_j \widetilde{\to} y_j \mid E_4, F, \widetilde{x_i^T} = \widetilde{x_j^T} = v, \tau = T\right) \cdot \mathbb{P}\left(F \mid E_4, \widetilde{x_i^T} = \widetilde{x_j^T} = v, \tau = T\right) \\
+\ &\mathbb{P}\left(x_i \widetilde{\to} y_i, x_j \widetilde{\to} y_j \mid E_4, G^C, H, \widetilde{x_i^T} = \widetilde{x_j^T} = v, \tau = T\right) \cdot \mathbb{P}\left(G^C, H \mid E_4, \widetilde{x_i^T} = \widetilde{x_j^T} = v, \tau = T\right) \\
+\ &\mathbb{P}\left(x_i \widetilde{\to} y_i, x_j \widetilde{\to} y_j \mid E_4, G^C, H^C, \widetilde{x_i^T} = \widetilde{x_j^T} = v, \tau = T\right) \cdot \mathbb{P}\left(G^C, H^C \mid E_4, \widetilde{x_i^T} = \widetilde{x_j^T} = v, \tau = T\right) \\
\leq\ &\frac{1}{N(N-1)} \cdot 1 + \frac{1}{2(N-1)} \cdot \frac{1}{N} + \frac{1}{(N-1)(N-2)} \cdot 1 < \frac{5}{2(N-1)(N-2)}
\end{aligned}$$

Since this bound does not depend on $v$, we have

$$\mathbb{P}\left(x_i \widetilde{\to} y_i, x_j \widetilde{\to} y_j \mid E_4, \tau = T\right) \leq \frac{5}{2(N-1)(N-2)}$$

Now it is time to combine our four cases. The bounds in cases $E_2$ and $E_4$ are already sufficiently small, but to make the bound in $E_3$ useful we need to incorporate the fact that $E_3$ is unlikely. First

note that

$$\begin{aligned} E_3 = \quad & (B_i \cap B_j \cap F_i^C \cap F_j) \cup (B_i \cap B_j \cap F_i \cap F_j^C) \\ & \cup (B_i^C \cap B_j \cap F_i^C \cap F_j) \cup (B_i^C \cap B_j \cap F_i \cap F_j^C) \\ & \cup (B_i \cap B_j^C \cap F_i^C \cap F_j) \cup (B_i \cap B_j^C \cap F_i \cap F_j^C) \end{aligned}$$

In other words $E_3$ is given by the union of 6 events, encompassing the outcomes where exactly one card flips all tails after $T$, and at least one card flips all tails before $T$. Note that the probability of any particular card flipping all Tails before $T$ is $2^{-(T-1)}$. The probability of any particular card flipping all Tails after $T$ is $2^{-(r-T)}$. (Heads and Tails are still equally likely, as there is still symmetry after excluding Heads, Tails and Tails, Heads flips with pairing round keys to avoid collision.) Therefore, each of these 6 events has probability bounded above by $2^{-(T-1)} \cdot 2^{-(r-T)} = 2^{-(r-1)}$. Since there are 6 events, taking the union bound gives us

$$\mathbb{P}(E_3) \leq 6 \cdot 2^{-(r-1)} = 12 \cdot 2^{-r} = \frac{12 \cdot 2^{d-r}}{N}$$

Now we compute,

$$\begin{aligned} \mathbb{P}(x_i \widetilde{\rightarrow} y_i, x_j \widetilde{\rightarrow} y_j \mid \tau = T) = & \mathbb{P}(x_i \widetilde{\rightarrow} y_i, x_j \widetilde{\rightarrow} y_j, E_1 \cup E_2 \cup E_3 \cup E_4 \mid \tau = T) \\ & \leq \mathbb{P}(x_i \widetilde{\rightarrow} y_i, x_j \widetilde{\rightarrow} y_j \mid E_1, \tau = T) \cdot \mathbb{P}(E_1 \mid \tau = T) \\ & \quad + \mathbb{P}(x_i \widetilde{\rightarrow} y_i, x_j \widetilde{\rightarrow} y_j \mid E_2, \tau = T) \cdot \mathbb{P}(E_2 \mid \tau = T) \\ & \quad + \mathbb{P}(x_i \widetilde{\rightarrow} y_i, x_j \widetilde{\rightarrow} y_j \mid E_3, \tau = T) \cdot \mathbb{P}(E_3 \mid \tau = T) \\ & \quad + \mathbb{P}(x_i \widetilde{\rightarrow} y_i, x_j \widetilde{\rightarrow} y_j \mid E_4, \tau = T) \cdot \mathbb{P}(E_4 \mid \tau = T) \\ & \leq 0 + \frac{1}{N(N-1)} \cdot 1 + \frac{2}{N-1} \cdot \frac{12 \cdot 2^{d-r}}{N} + \frac{5}{2(N-1)(N-2)} \cdot 1 \\ & \leq \frac{7 + 48 \cdot 2^{d-r}}{2(N-1)(N-2)} \end{aligned}$$

Note that this bound is the same regardless of the value of $T \in \{1, \ldots, r\}$. Since $\tau \leq r$ is equivalent to $M_{i,j}$ we have,

$$\mathbb{P}(x_i \widetilde{\rightarrow} y_i, x_j \widetilde{\rightarrow} y_j \mid M_{i,j}) \leq \frac{7 + 48 \cdot 2^{d-r}}{2(N-1)(N-2)}$$

which is the statement of the theorem. □

### 2.5. Uniformity of the Swap-or-Not Shuffle

We are now ready to prove that the swap-or-not shuffle has a distribution that is close to uniform. We begin by defining a new construction of the tilde process.

PROPOSITION 2.5.1. *The tilde process can be defined as follows:*
*Fix a subset of $q$ distinct cards $x_1, \ldots, x_q \in \mathbb{Z}_2^d$. As before, generate uniform independent $K_1, \ldots, K_r \in \mathbb{Z}_2^d$. Additionally, generate a uniform $W \in (\mathbb{Z}_2^d)^q$. Now, for any $1 \le \ell \le r$, we let*

$$\begin{aligned}
\widetilde{x_1^\ell} &= x_1 + \widehat{c_{1,1}}K_1 + \widehat{c_{1,2}}K_2 + \cdots + \widehat{c_{1,k}}K_\ell \\
\widetilde{x_2^\ell} &= x_2 + \widehat{c_{2,1}}K_1 + \widehat{c_{2,2}}K_2 + \cdots + \widehat{c_{2,r}}K_\ell \\
&\vdots \\
\widetilde{x_q^\ell} &= x_q + \widehat{c_{q,1}}K_1 + \widehat{c_{q,2}}K_2 + \cdots + \widehat{c_{q,r}}K_\ell
\end{aligned}$$

*where $\widehat{c_{i,j}}$ are random elements of $\{0,1\}$ defined as follows:*

- *If $K_1, \ldots, K_r$ span $\mathbb{Z}_2^d$, then, conditioned on the values of $K_1, \ldots, K_r$, the coins $\widehat{c_{i,1}}, \widehat{c_{i,2}}, \ldots, \widehat{c_{i,r}}$ are chosen uniformly from all choices such that $\widehat{x_i^r} = W_i$ for all $i$, and independently of all $\widehat{c_{m,j}}$ where $m \neq i$.*
- *If $K_1, \ldots, K_r$ do not span $\mathbb{Z}_2^d$, then $\widehat{c_{i,j}}$ are all chosen independently and uniformly from $\{0,1\}$.*

PROOF. In the case that $K_1, \ldots, K_r$ do not span $\mathbb{Z}_2^d$ we have by definition that all $\widehat{c_{i,j}}$ are independent, which is consistent with the tilde process. In the case that $K_1, \ldots, K_r$ do span $\mathbb{Z}_2^d$ then the disribution of the coins matches that of statement (3) in Lemma 2.3.2, so the distribution of the coins is consistent with the tilde process. □

Now that we have shown this new construction for the tilde process, we will from now on assume that the tilde process is generated using $W$.

LEMMA 2.5.2. *In the tilde process with $r$ rounds, generated using $W$ we have,*

$$\mathbb{P}(W_i = y_i, W_j = y_j \mid M_{i,j}) \le \frac{9 + 48 \cdot 2^{-r+d}}{2(N-1)(N-2)}.$$

PROOF. This inequality is similar to the one in Proposition 2.5.1. To relate the two inequalities, we must condition on $A_r$, as $A_r$ determines if $W$ fixes the final positions of the cards, or if $W$ is ignored completely. We can decompose $\mathbb{P}(W_i = y_i, W_j = y_j \mid M_{i,j})$ as

$$\begin{aligned}\mathbb{P}(W_i = y_i, W_j = y_j \mid M_{i,j}) &= \mathbb{P}(W_i = y_i, W_j = y_j, A_r \mid M_{i,j}) + \mathbb{P}(W_i = y_i, W_j = y_j, A_r^C \mid M_{i,j}) \\ &= \mathbb{P}(x_i \widetilde{\to} y_i, x_j \widetilde{\to} y_j, A_r \mid M_{i,j}) + \mathbb{P}(W_i = y_i, W_j = y_j, A_r^C \mid M_{i,j}) \\ &\leq \mathbb{P}(x_i \widetilde{\to} y_i, x_j \widetilde{\to} y_j \mid M_{i,j}) + \mathbb{P}(W_i = y_i, W_j = y_j) \cdot \mathbb{P}(A_r^C \mid M_{i,j}). \qquad (2.5)\end{aligned}$$

In line (2.5) we used $\mathbb{P}(W_i = y_i, W_j = y_j \mid A_r^C, M_{i,j}) = \mathbb{P}(W_i = y_i, W_j = y_j)$, which is true because on the event $A_r^C$, the value of $W$ is independent of the trajectories of the cards. Since $W$ is uniform, we have $\mathbb{P}(W_i = y_i, W_j = y_j) = \frac{1}{N^2}$. Also recall that Proposition 2.4.1 gave us

$$\mathbb{P}(x_i \widetilde{\to} y_i, x_j \widetilde{\to} y_j \mid M_{i,j}) \leq \frac{7 + 48 \cdot 2^{-r+d}}{2(N-1)(N-2)}.$$

Putting this together with (2.5), we get,

$$\mathbb{P}(W_i = y_i, W_j = y_j \mid M_{i,j}) \leq \frac{7 + 48 \cdot 2^{-r+d}}{2(N-1)(N-2)} + \frac{1}{N^2} \cdot 1 \leq \frac{9 + 48 \cdot 2^{-r+d}}{2(N-1)(N-2)}$$

which completes the lemma. □

LEMMA 2.5.3. *Consider the tilde process with $r$ rounds, generated using $W = (w_1, \ldots, w_q)$. Let $M$ be the event that there are no pairwise collisions between any of $x_1, \ldots, x_q$. Then,*

$$\mathbb{P}(W = (y_1, \ldots, y_q), M) < \frac{rq(q-1)(9 + 48 \cdot 2^{-r+d})}{4(N-2)N^q}.$$

PROOF. To start, we will use the union bound to break up $M$ into it's specific collisions:

$$\begin{aligned}\mathbb{P}(W = (y_1, \ldots, y_q), M) &= \mathbb{P}\left(\bigcup_{1 \leq i < j \leq q} \{W = (y_1, \ldots, y_q), M_{i,j}\}\right) \\ &\leq \sum_{1 \leq i < j \leq q} \mathbb{P}(W = (y_1, \ldots, y_q), M_{i,j})\end{aligned}$$

We break the terms in the sum into

$$\mathbb{P}(W = (y_1, \ldots, y_q), M_{i,j}) = \mathbb{P}(M_{i,j} \mid W = (y_1, \ldots, y_q)) \cdot \mathbb{P}(W = (y_1, \ldots, y_q)). \qquad (2.6)$$

Note that $M_{i.j}$ depends only on the trajectories of $x_i$ and $x_j$, and is independent of other cards. So,

$$\mathbb{P}(M_{i,j} \mid W = (y_1, \ldots, y_q)) = \mathbb{P}(M_{i,j} \mid W_i = y_i, W_j = y_j). \tag{2.7}$$

To compute $\mathbb{P}(M_{i,j} \mid W_i = y_i, W_j = y_j)$ we use Bayes' formula:

$$\mathbb{P}(M_{i,j} \mid W_i = y_i, W_j = y_j) = \frac{\mathbb{P}(M_{i,j})}{\mathbb{P}(W_i = y_i, W_j = y_j)} \cdot \mathbb{P}(W_i = y_i, W_j = y_j \mid M_{i,j}). \tag{2.8}$$

Now we need to bound the three probabilities on the RHS of (2.8). Since $W$ is uniform,

$$\mathbb{P}(W_i = y_i, W_j = y_j) = \frac{1}{N^2}.$$

In round $t$ of the shuffle, there is a collision if $K_t = \widetilde{x_i^{t-1}} + \widetilde{x_i^{t-1}}$ and $\widetilde{c_{i,t}} \neq \widetilde{c_{j,t}}$. The round keys and coins are chosen independently and uniformly. There is a $\frac{1}{N}$ chance the round key is chosen to pair $x_i$ and $x_j$, and a $\frac{1}{2}$ chance afterwards that the coins cause a collision. Therefore, the probability of collision in round $t$ is $\frac{1}{2N}$ for all $t$. Using the union bound, we see that the probability of having at least one collision across all the rounds has

$$\mathbb{P}(M_{i,j}) \leq \frac{r}{2N}.$$

Finally, we use the bound for $\mathbb{P}(W_i = y_i, W_j = y_j \mid M_{i,j})$, we calculated in Lemma 2.5.2:

$$\mathbb{P}(W_i = y_i, W_j = y_j \mid M_{i,j}) \leq \frac{9 + 48 \cdot 2^{-r+d}}{2(N-1)(N-2)}.$$

Together, we get

$$\mathbb{P}(M_{i,j} \mid W_i = y_i, W_j = y_j) \leq \frac{r}{2N} \cdot N^2 \cdot \frac{9 + 48 \cdot 2^{-r+d}}{2(N-1)(N-2)} \leq \frac{r(9 + 48 \cdot 2^{-r+d})}{2(N-2)}$$

where in the second inequality we used that $\frac{N}{N-1} \leq 2$ as $N \geq 2$. Now we combine this with lines (2.6) and (2.7) to get

$$\begin{aligned}
\mathbb{P}(W = (y_1, \ldots, y_q), M_{i,j}) &= \mathbb{P}(M_{i,j} \mid W = (y_1, \ldots, y_q)) \cdot \mathbb{P}(W = (y_1, \ldots, y_q)) \\
&= \mathbb{P}(M_{i,j} \mid W_i = y_i, W_j = y_j) \cdot \mathbb{P}(W = (y_1, \ldots, y_q)) \\
&\leq \frac{r(9 + 48 \cdot 2^{-r+d})}{2(N-2)} \cdot \frac{1}{N^q},
\end{aligned}$$

where we have used that $\mathbb{P}(W = (y_1, \ldots, y_q)) = \frac{1}{N^q}$ due to uniformity. Finally, we sum over all $i \neq j$:

$$\mathbb{P}(W = (y_1, \ldots, y_q), M) \leq \sum_{1 \leq i < j \leq q} \frac{r(9 + 48 \cdot 2^{-r+d})}{2(N-1)} \cdot \frac{1}{N^q} \leq \frac{rq(q-1)(9 + 48 \cdot 2^{-r+d})}{4(N-2)N^q}.$$

□

THEOREM 2.5.4. *Fix $d \in \mathbb{N}$, and $r \geq d$. Fix $x_1, \ldots, x_q, y_1, \ldots, y_q \in \mathbb{Z}_2^d$. Then, in a swap-or-not shuffle with $r$ rounds and $N = 2^d$ cards,*

$$\mathbb{P}(x_1 \to y_1, \ldots, x_q \to y_q) \geq \frac{1}{(N)_q} \cdot \left(1 - \frac{q^2}{N} - 2^{-r+d} - \frac{rq(q-1)(9 + 48 \cdot 2^{-r+d})}{4(N-2)}\right)$$

PROOF. We begin considering the coupled tilde process, generated with $W$, and applying Corollary 2.3.5:

$$\begin{aligned}
\mathbb{P}(x_1 \to y_1, x_2 \to y_2, \ldots, x_q \to y_q) &\geq \mathbb{P}(x_1 \widetilde{\to} y_1, x_2 \widetilde{\to} y_2, \ldots, x_q \widetilde{\to} y_q, M^C) \\
&\geq \mathbb{P}(x_1 \widetilde{\to} y_1, x_2 \widetilde{\to} y_2, \ldots, x_q \widetilde{\to} y_q, M^C, A_r) \\
&= \mathbb{P}(W = (y_1, \ldots, y_q), M^C, A_r)
\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{P}(W = (y_1, \ldots, y_q), M^C, A_r) \\
&\geq \mathbb{P}(W = (y_1, \ldots, y_q)) - \mathbb{P}(W = (y_1, \ldots, y_q), M) - \mathbb{P}(W = (y_1, \ldots, y_q), A_r^C). \qquad (2.9)
\end{aligned}$$

We now need to bound the three probabilities in (2.9). Since $W$ is uniform, we have $\mathbb{P}(W = (y_1, \ldots, y_q)) = \frac{1}{N^q}$. We know from Lemma 2.3.1 that $\mathbb{P}(A_r^C) = 2^{-r+d}$. Since $W$ is independent of the round keys, we have

$$\mathbb{P}(W = (y_1, \ldots, y_q), A_r^C) = \frac{1}{N^q} \cdot 2^{-r+d}$$

Combining this with our bound for $\mathbb{P}(W = (y_1, \ldots, y_q), M)$ from Lemma 2.5.3, we get

$$\begin{aligned}
\mathbb{P}(W = (y_1, \ldots, y_q), M^C \cap A_r) &\geq \frac{1}{N^q} - \frac{1}{N^q} \cdot 2^{-r+d} - \frac{rq(q-1)(9 + 48 \cdot 2^{-r+d})}{4(N-2)N^q} \\
&= \frac{1}{N^q} \cdot \left(1 - 2^{-r+d} - \frac{rq(q-1)(9 + 48 \cdot 2^{-r+d})}{4(N-2)}\right)
\end{aligned}$$

To show small separation distance, our goal is to prove that $\mathbb{P}(x_1 \to y_1, x_2 \to y_2, \ldots, x_q \to y_q) \geq (1-\epsilon)\frac{1}{(N)_q}$ for a small $\epsilon$, so it remains to show that $\frac{1}{(N)_q}$ approximately equals $\frac{1}{N^q}$ for sufficiently small $q$. Note that

$$\frac{1}{(N)_q} = \frac{1}{N(N-1)\ldots(N-q+1)} \leq \frac{1}{(N-q)^q}. \tag{2.10}$$

Note that for any $a > 1, b \in \mathbb{N}$,

$$(a-1)^b = a^b \left(1 - \frac{1}{a}\right)^b \geq a^b \left(1 - \frac{b}{a}\right),$$

hence

$$\begin{aligned}
\frac{1}{(N-q)^q} &= \frac{1}{(q(\frac{N}{q}-1))^q} \\
&= \frac{1}{q^q(\frac{N}{q}-1)^q} \\
&\leq \frac{1}{q^q(\frac{N}{q})^q(1-\frac{q^2}{N})} \\
&= \frac{N^{-q}}{1-\frac{q^2}{N}}.
\end{aligned}$$

Combining with (2.10) gives

$$\frac{1}{(N)_q} \cdot \left(1 - \frac{q^2}{N}\right) \leq N^{-q}.$$

Going back to the bound on mixing, we get

$$\begin{aligned}
\mathbb{P}(x_1 \to y_1, \ldots, x_q \to y_q) &\geq \frac{1}{(N)_q} \cdot \left(1 - \frac{q^2}{N}\right) \cdot \left(1 - 2^{-r+d} - \frac{rq(q-1)(9+48\cdot 2^{-r+d})}{4(N-2)}\right) \\
&\geq \frac{1}{(N)_q} \cdot \left(1 - \frac{q^2}{N} - 2^{-r+d} - \frac{rq(q-1)(9+48\cdot 2^{-r+d})}{4(N-2)}\right)
\end{aligned}$$

□

### 2.6. Upper Bound on Advantage

In Chapter 1 we showed that small separation distance leads to good CCA security. In this section, we will use that theorem to prove the main result of this chapter. That is, we will show that the

swap-or-not shuffle has good CCA security as long as the number of queries is a bit lower than the square root of the number of cards.

THEOREM 2.2.1. *Fix any $d \geq 2$. Let $X$ be the swap-or-not shuffle with $N = 2^d$ cards, and $r$ rounds. Consider a CCA adversary equipped with $q$ queries up against this swap-or-not shuffle. The security of $X$ against this adversary is bounded by*

$$CCA_q(X) \leq \frac{q^2}{N} + 2^{-r+d} + \frac{rq(q-1)(9+48\cdot 2^{-r+d})}{4(N-2)}$$

PROOF. The proof is a direct result of Theorem 2.5.4 and Theorem 1.2.2. Consider any distinct $x_1, \ldots, x_q$ and distinct $y_1, \ldots, y_q$ in $\mathbb{Z}_2^d$. By Theorem 2.5.4 we have,

$$\mathbb{P}(x_1 \to y_1, \ldots, x_q \to y_q) \geq \frac{1}{(N)_q} \cdot \left(1 - \frac{q^2}{N} - 2^{-r+d} - \frac{rq(q-1)(9+48\cdot 2^{-r+d})}{4(N-2)}\right).$$

Note that under a uniform random permutation, the probability of $(x_1, \ldots, x_q)$ being sent to $(y_1, \ldots, y_q)$ is $\frac{1}{(N)_q}$. So,

$$d_{\text{sep}}(X, \mu) \leq \frac{q^2}{N} + 2^{-r+d} + \frac{rq(q-1)(9+48\cdot 2^{-r+d})}{4(N-2)}$$

where $\mu$ is the uniform random permutation. Since this holds for all distinct choices of $q$ queries, we have, by Theorem 1.2.2,

$$CCA_q(X) \leq \frac{q^2}{N} + 2^{-r+d} + \frac{rq(q-1)(9+48\cdot 2^{-r+d})}{4(N-2)}$$

□

This shows that about $\log_2(N)$ rounds is sufficient for the swap-or-not shuffle on $N$ cards to achieve strong CCA security against an adversary with fewer than $\sqrt{N}$ queries. This lower bound on the number of rounds is tight. To be specific, suppose $Y$ is the swap-or-not shuffle on $N = 2^d$ cards with $d-1$ rounds. Then as long as an adversary has $q > d + \varepsilon$ queries the nCPA security (and therefore the CCA security) of $Y$ is very weak. This is because with $d-1$ rounds the round keys will not span $\mathbb{Z}_2^d$. This means that for each queried card $x_1, \ldots, x_q$ the adversary will notice $Y(x_1) - x_1, \ldots, Y(x_q) - x_q$ are all in the same subspace. This behavior is unlikely under the uniform random permutation when $q > d$ so $Y$ will have high total variation distance from uniform.

CHAPTER 3

# Mixing Time of the Overlapping Cycles Shuffle

The overlapping cycles shuffle was first described by Jonasson [6] and takes two parameters, $n \in \mathbb{N}$ and $m \in \{2, \ldots, n-1\}$. We define the shuffle as follows: Begin with a deck of $n$ cards. In each round, flip an independent coin. If Heads, move the $m$th card to the top of the deck. If Tails, move the $n$th card to the top of the deck.

Despite its simple construction, the overlapping cycles shuffle has interesting and surprising properties. Angel, Peres, and Wilson [1] determined the spectral gap of the chain which tracks a single card in the overlapping cycles shuffle. Their analysis determined that if $m = \lfloor \alpha n \rfloor$ for some $\alpha \in (0, 1)$ then the asymptotic relaxation time as $n$ grows depends on how well approximated $\alpha$ is by rational numbers. In particular, if $\alpha$ is rational the relaxation time is $O(n^2)$. However the relaxation time can be as short as $O(n^{\frac{3}{2}})$ which occurs when $\alpha = \frac{\sqrt{5}-1}{2}$, the inverse golden ratio.

Angeles, Peres, and Wilson ask if the mixing time of the entire deck is within a factor of $\log(n)$ of their result for individual cards. We prove something close: The mixing time is $O(n^2 \log^3(n))$ for rational $\alpha$ and $O(n^{\frac{3}{2}} \log^3(n))$ for"very" irrational $\alpha$ like the inverse golden ratio.

## 3.1. Description of the Shuffle

The overlapping cycles shuffle has a simple description as a random walk on the symmetric group $S_n$. In each round $g$ is equally likely to go to $(1, 2, \ldots, m)g$ or $(1, 2, \ldots, n)g$.

This explains where the name "overlapping cycles shuffle" comes from. Note that if $m$ and $n$ are both odd, then $(1, \ldots, m)$ and $(1, \ldots, n)$ will both be even permutations. Thus, the mixing time we seek to bound will be in respect to convergence to a distribution which is uniform across $A_n$, not $S_n$. If $m$ and $n$ are both even, then $(1, \ldots, m)$ and $(1, \ldots, n)$ will both be odd permutations, and the shuffle will be periodic. In this case we say the mixing time is the value $t$ such that if $r > t$

then the distribution after $r$ rounds of the shuffle is approximately uniform over $A_n$ if $r$ is even and $S_n \backslash A_n$ if $r$ is odd. If $(1, \ldots, m)$ and $(1, \ldots, n)$ have different parity then there are no issues and we consider mixing time in the typical sense.

The $m = n-1$ case of the overlapping cycles shuffle was shown to have a mixing time of $O(n^3 \log(n))$ by Hildebrand [3] in his dissertation. This case was studied before the overlapping cycles shuffle was defined in its general form and given its name. The $m = n - 1$ case is especially slow because the cycles $(1, \ldots, n-1)$ and $(1, \ldots, n)$ act identically on every element of $\{1, \ldots, n\}$ except for $n-1$ and $n$. We will only consider values of $m$ such that $\frac{m}{n}$ is "not too close" to 0 or 1, and this will allow our mixing time bound to be on a lower order.

It will be useful to analyze the distribution given by the inverse permutation of $t$ steps of the overlapping cycles shuffle. It turns out that this "inverse overlapping cycles shuffle" is just the overlapping cycles shuffle in disguise.

THEOREM 3.1.1. *Let $\pi_t$ be the random permutation that is $t$ steps of the overlapping cycles shuffle on $n$ cards with parameter $m$. Then,*

$$\pi_t^{-1} \stackrel{d}{=} \sigma \pi_t \sigma^{-1}$$

*where*

$$\sigma = \begin{pmatrix} 1 & 2 & \ldots & m & m+1 & m+2 & \ldots & n \\ m & m-1 & \ldots & 1 & n & n-1 & \ldots & m+1 \end{pmatrix}.$$

*In other words, the inverse overlapping cycles shuffle is also an overlapping cycles shuffle after reordering the cards.*

PROOF. Note that under $\pi_1^{-1}$ any $g$ is equally likely to go to $(m, m-1, \ldots, 1)g$ or $(n, n-1, \ldots, 1)g$. Also note that

$$\sigma(1, 2, \ldots, m)\sigma^{-1} = (m, m-1, m \ldots, 1)$$

and

$$\begin{aligned}\sigma(1,2,\ldots,n)\sigma^{-1} &= \sigma(1,2,\ldots,m,m+1,m+2,\ldots,n)\sigma^{-1}\\ &= (m,m-1,\ldots,1,n,n-1,\ldots,m+1)\\ &= (n,n-1,\ldots,1).\end{aligned}$$

Since each step of the overlapping cycles shuffle is the same as the inverse up to the given re-ordering of the cards, we know that the same is true for $t$ steps with the same reordering. □

### 3.2. Main Theorem

Our main goal of this chapter is to establish an upper bound on the mixing time of the overlapping cycles shuffle. The following theorem provides an upper bound on the mixing time of the overlapping cycles shuffle on $n$ cards with parameter $m$, which is tight up to a factor of $\log^3(n)$ provided that $\frac{m}{n}$ is bounded away from 0 and 1.

THEOREM 3.2.1. *Consider the overlapping cycles shuffle on $n$ cards with parameter $m$. Let $\ell_{\max}$ be defined by*

$$\ell_{\max} = \max_{\omega\in\{1,\ldots,2n-m\}} \Big\{ \min\big\{|a|+|b|\sqrt{n} \ : \ a,b\in\mathbb{Z},\ \omega \equiv a+bm \mod 2n-m+1\big\}\Big\}.$$

*Then the mixing time is at most*

$$\mathcal{A}\ell_{\max}^2\log^3(n)\mathcal{L}$$

*where $\mathcal{A}$ is a universal constant and*

$$\mathcal{L} = \exp\left(-864\left(\frac{192n\exp(\frac{10n}{m})}{(n-m)}\right)^2 - \frac{8n}{m}\right).$$

Note that if $\frac{m}{n}$ is bounded away from 0 and 1 (say $\frac{1}{100}\le\frac{m}{n}\le\frac{99}{100}$ for example) then $\mathcal{L}$ is a constant. Note that $\ell_{\max}$ is always bounded above by $2n$ because $(a,b)=(\omega,0)$ is always an element of the inner set, and $|\omega|+|0|\sqrt{n}=\omega\le 2n-m$. So for all values of $m$ bounded away from 0 and 1 the mixing time is at most a constant times $n^2\log^3(n)$.

Now fix some $\alpha\in(0,1)$ and for any deck size $n$ consider the shuffle where $m=\lfloor\alpha n\rfloor$. We are

interested in the asymptotic mixing time as $n$ approaches infinity. It turns out that if $\alpha$ is rational then we do no better than the universal upper bound of and get a mixing time of $O(n^2 \log^3(n))$. However, if $\alpha$ is an irrational number whose multiples form a low-discrepancy sequence, then the mixing time is $O(n^{\frac{3}{2}} \log^3(n))$. We show this in particular for $\alpha = \frac{\sqrt{5}-1}{2}$, the inverse golden ratio, at the very end of the chapter.

### 3.3. Movement of a Single Card: Intuition and Notation

In our analysis of the overlapping cycles shuffle we will make heavy use of sequences of coins. We defined the overlapping cycles shuffle in terms of a sequence of coins, and we will later imagine that in each step the shuffle draws from different pools of pre-flipped coins depending on its state. To start we define the following:

DEFINITION 3.3.1. *If $c = (c_1, c_2, \ldots, c_t)$ is a sequence of independent uniform {Heads, Tails}-valued random variables, we say it is a sequence of coins of length $t$.*

DEFINITION 3.3.2. *For a sequence of coins $S = (c_1, c_2, \ldots, c_t)$ and $r \leq t$, define the Heads-Tails differential $\text{Diff}_r(S)$ to be the number of Heads in $(c_1, \ldots, c_r)$ minus the number of Tails in $(c_1, \ldots, c_r)$.*

Note that $\text{Diff}_r(c)$ is a simple symmetric random walk. We will use this fact later to show with probability bounded away from 0 that $\text{Diff}_r(c)$ stays within constant standard deviations of 0.

To understand the overlapping cycles shuffle it is important to understand how each step affects cards in different parts of the deck. Note that if a card is in one of the top $m-1$ positions in the deck, then it will move down one position in the round regardless of if Heads or Tails is flipped. On the other hand, a card in a position between $m+1$ and $n$ is equally to stay put or move down one position, with the $n$th card "wrapping around" to the top of the deck if it moves "down". Lastly, the card in position $m$ is equally likely to either move to the top of the deck or move down one position.

Since the behavior of cards in positions 1 through $m-1$ varies from cards in positions $m+1$ through $n$, we should have language to quickly distinguish the two.

DEFINITION 3.3.3. *If a card is in a position between* 1 *and* $m-1$ *we say the card is in the* ***top part of the deck****. If a card is in a position between* $m+1$ *and* $n$ *we say the card is in the* ***bottom part of the deck****.*

Note that cards in the bottom part of the deck move, on average, half as quickly as cards in the top part of the deck. This is because cards in the top part of the deck move down one position each step, while cards in the bottom part of the deck move down only if Tails is flipped, which happens half the time. So if $x$ is a position in the top part of the deck, and $y$ is a position in the bottom part of the deck, we should think of positions $x$ and $x+1$ as being distance 1 from each other, and positions $y$ and $y+1$ as being distance 2 from each other. To quickly reference this notion, we define the following function:

DEFINITION 3.3.4. *If* $x \in \{1, \ldots, n\}$ *is a position in the deck, let*

$$p(x) = \begin{cases} x \text{ if } x \leq m \\ x + (x - m) \text{ if } x > m \end{cases}$$

We will be interested in finding the likelyhood of certain cards being in certain "nearby positions" after specific numbers of rounds. To do this we need to reevaluate our notion of distance away from the naive definition of physical distance in the deck. To see why, note that the card in position $m$ has a $\frac{1}{4}$ chance of being adjacent to the card in position $n$ after two steps. Thus, it makes sense to consider position $m$ and position $n$ as "close".

We will name cards after their initial position in the deck. So card 1 is the card initially on the top of the deck, card 2 is the card initially second to top, etc.

DEFINITION 3.3.5. *We use* $i_t$ *to denote the position of card* $i$ *(i.e. the card that was originally in position* $i$*) after* $t$ *steps of the shuffle.*

To determine where card $i$ is after $t$ steps it is enough to know the sequence of coin flips $c_1, c_2, \ldots c_t$ (where each $c_r \in \{\text{Heads}, \text{Tails}\}$). However this is much more information than we need. For example, if $i << m$, then $i$ will deterministicly move downwards for many steps regardless of the early values of $(c_r)$. The following proposition allows us to compute the position of $i$ only using the coins that actually influence the movement of $i$.

PROPOSITION 3.3.6. *Let $i$ be a card in a deck of $n$ cards (where as previously mentioned, $i$ begins in position $i$). Suppose $(c_r)$ is a sequence of coins. Let $i_t$ be the position of card $i$ after doing $t$ steps of the Overlapping Cycle Shuffle. Let $H_B$ be the number of times in the first $t$ steps a Heads is drawn while $i$ is in position $m$. Let $T_B$ be the number of times in the first $t$ steps a Heads is drawn while $i$ is in position $m$. Let $H_S$ be the number of "Heads" drawn in the first $t$ steps while $i$ is in the bottom part of the deck. Let $T_S$ be the number of "Tails" drawn in the first $t$ steps while $i$ is in the bottom part of the deck. Then,*

$$p(i_t) \equiv p(i) + t + (T_S - H_S) + (T_B - mH_B) \ mod\ (2n - m + 1).$$

PROOF. We proceed by induction. If $t = 0$ then the statement is trivially true. For the inductive step, it suffices to show that, if $i_{t-1}$ is the position of $i$ after $t-1$ steps, then $i$'s distance after step $t$ depends on the coin or lack of coin used in step $t$ exactly as the formula dictates. Specifically, we need to show that

$$\begin{aligned} p(i_t) \equiv & p(i_{t-1}) + 1 + \mathbb{1}(i_{t-1} \geq m \text{ and } c_t = T) - \mathbb{1}(i_{t-1} > m \text{ and } c_t = H) \\ & - m \cdot \mathbb{1}(i_{t-1} = m \text{ and } c_t = H). \end{aligned}$$

- If $i$ is positioned in the top part of the deck it must move down in the next step so $p(i_t) = p(i_{t-1}) + 1$. Similarly, if $i$ is in position $m$ and tails is flipped then $i_{t-1} = m$ and $i_t = m + 2$ and so $p(i_t) = m + 1 + 1 = p(i_{t-1}) + 1 + 1$.
- If $i$ is in the bottom part of the deck, but not in position $n$, and flips a Tails then $i_t = i_{t-1}+1$ and so $p(i_t) = p(i_{t-1}) + 2 = p(i_{t-1}) + 1 + 1$. If $i$ is in position $n$ and flips a Tails, then it moves to position 1, so $p(i_{t-1}) = 2n - m$ and $p(i_t) = 1$ and therefore $p(i_t) = p(i_{t-1}) + 1 + 1$ mod $(2n - m + 1)$.
- If $i$ is in the bottom part of the deck and flips a Heads then $i_t = i_{t-1}$ and so $p(i_t) = p(i_{t-1}) + 1 - 1$.
- If $i$ is in position $m$ and flips Heads, then $i_{t-1} = m$ and $i_t = 1$. So $p(i_t) = p(i_{t-1}) + 1 - m$.

□

We use $H_B$ and $T_B$ to denote movement in position $m$ because the choice of $i$ going to $m + 1$ or 1 is a "big" choice. We use $H_S$ and $T_S$ to denote movement in the bottom part of the deck, because

the choice of $i$ moving down one position or staying in place has a "small" impact on its movement. From now on we will use the letter B for "big coins" which are drawn when certain cards are in position $m$ and the letter S for "small coins" which are drawn when certain cards are in the bottom part of the deck.

As a consequence of Proposition 3.3.6, we have the following corollary:

COROLLARY 3.3.7. *Let $i$ and $j$ be two cards in a deck of $n$ cards. Suppose $(c_r)$ is a sequence of coins. Let $i_t$ be the position of card $i$ after doing $t$ steps of the Overlapping Cycle Shuffle drawing from $(c_r)$ and as described previously. Let $H_B(i)$ and $H_S(i)$ be the number of Heads drawn from $(c_r)$ when $i$ is in position $m$ and in the bottom part of the deck, respectively, in the first $t$ steps. Let $T_B(i)$ and $T_S(i)$ be the number of Tails drawn from $(c_r)$ when $i$ is in position $m$ and in the bottom part of the deck respectively in the first $t$ steps. Let $j_t, H_B(j), H_S(j), T_B(j), T_S(j)$ be defined similarly for card $j$ after $t$ steps drawing from $(c_r)$. Then,*

$$\begin{aligned} p(j_t) - p(i_t) \equiv p(j_0) - p(i_0) + (T_S(j) - H_S(j)) - (T_S(i) - H_S(i)) \\ + (T_B(j) - T_B(i)) - m(H_B(j) - H_B(i)) \ mod\ (2n - m + 1) \end{aligned}$$

This corollary helps inform us as to how we should consider the "closeness" of cards. Assume that $\frac{n}{10} < m < \frac{9n}{10}$. Note that after many steps, the magnitudes of $T_S(j) - H_S(j)$ and $T_S(i) - H_S(i)$ will each likely be much more than the magnitude of $H_B(i) - H_B(j)$. This is because whenever $i$ is in the bottom part of the deck the coins flipped add to either $T_S(i)$ or $H_S(i)$ and whenever $j$ is in the bottom part of the deck the coins flipped add to either $T_S(j)$ or $H_S(j)$. However coins only add to $H_B(i), T_B(i)$ or $H_B(j), T_B(j)$ when $i$ or $j$ is exactly in position $m$. We expect $i$ and $j$ to spend on the order of $n$ times more time in the bottom part of the deck than exactly in position $m$, so we should expect $|(T_S(i) - H_S(i))|$ and $|(T_S(j) - H_S(j))|$ to be about $\sqrt{n}|H_B(i) - H_B(j)|$ and $\sqrt{n}|T_B(i) - T_B(j)|$. With this intuition in mind, we define the following metric:

DEFINITION 3.3.8. *Let $\omega \in \mathbb{R}$. Then we define*

$$\|\omega\| = \min\{|a| + |b|\sqrt{n} \ : \ a \in \mathbb{R},\ b \in \mathbb{Z},\ \omega \equiv a + bm \ mod\ (2n - m + 1)\}.$$

*In particular if $i$ is a position in the deck then*

$$\|p(i)\| = \min\{|a| + |b|\sqrt{n} \ : \ a, b \in \mathbb{Z}, \ p(i) \equiv a + bm \ mod \ (2n - m + 1)\}$$

*and if $j$ and $k$ are positions in the deck then we have the distance*

$$\|p(k) - p(j)\| = \min\{|a| + |b|\sqrt{n} \ : \ p(k) \equiv p(j) + a + bm \ mod \ (2n - m + 1)\}.$$

It will be important to know how far apart cards can possibly be from each other. For this purpose we define the following constant.

DEFINITION 3.3.9. *Let $\ell_{\max}$ be defined by*

$$\ell_{\max} = \max_{\omega \in \{1,\ldots,2n-m\}} \|\omega\|.$$

We can show that $\ell_{\max} \geq \frac{1}{2} n^{\frac{3}{4}}$. To see this, note that if we choose $|a| < \ell$ and $|b|\sqrt{n} < \ell$ then we have $2\ell$ choices for $a$ and $\frac{2\ell}{\sqrt{n}}$ choices for $b$ and so there can be at most $\frac{4\ell^2}{\sqrt{n}}$ distinct cards with norms less than $\ell$. By the pigeonhole principle each of the elements of $\{1, \ldots, 2n-m\}$ are associated with exactly one duple $(a, b)$. In order to account for all $2n - m$ elements we need $\frac{4\ell^2}{\sqrt{n}} > 2n - m > n$ which translates to $\ell \geq \frac{1}{2} n^{\frac{3}{4}}$.

It will be useful to consider a more traditional "one dimensional" distance between positions in the deck, but allowing for "wrapping around" so that positions $n$ and 1 are considered close. We define this as follows.

DEFINITION 3.3.10. *Let $j$ and $k$ be positions in the deck. Then we define the distance*

$$|p(k) - p(j)|_M = \min\{|a| : p(k) \equiv p(j) + a \mod (2n - m + 1)\}.$$

One useful property of this notion of distance is that under the reordering of the cards used to simulate the inverse overlapping cycles shuffle, the distance between cards remains the same.

PROPOSITION 3.3.11. *Let $j$ and $k$ be positions in the deck. Let $\sigma$ be the permutation from Theorem 3.1.1. Then,*

$$p(k) - p(j) = p(\sigma(j)) - p(\sigma(k)) \mod 2n - m + 1.$$

PROOF. We consider three cases.

(1) $j, k \leq m$

In this case, $\sigma(j) = m + 1 - j$ and $\sigma(k) = m + 1 - k$. So,

$$\begin{aligned} p(\sigma(k)) - p(\sigma(j)) &= p(m+1-k) - p(m+1-j) \\ &= (m+1-k) - (m+1-j) \\ &= j - k \\ &= p(j) - p(k). \end{aligned}$$

(2) $j, k > m$

In this case, $\sigma(j) = n + m + 1 - j$ and $\sigma(k) = n + m + 1 - k$. So,

$$\begin{aligned} p(\sigma(k)) - p(\sigma(j)) &= p(n+m+1-k) - p(n+m+1-j) \\ &= (2n+m+2-2k) - (2n+m+2-2j) \\ &= 2j - 2k \\ &= (2j-m) - (2k-m) \\ &= p(j) - p(k). \end{aligned}$$

(3) $j \leq m < k$ In this case, $\sigma(j) = n + m + 1 - j$ and $\sigma(k) = n + m + 1 - k$. So,

$$\begin{aligned} p(\sigma(k)) - p(\sigma(j)) &= p(n+m+1-k) - p(m+1-j) \\ &= (2n+m+2-2k) - (m+1-j) \\ &= j - (2k-2n-1) \\ &\equiv (j) - (2k-m) \mod 2n-m+1 \\ &= p(j) - p(k). \end{aligned}$$

□

COROLLARY 3.3.12. *Let $j$ and $k$ be positions in the deck. Let $\sigma$ be the permutation from Theorem 3.1.1. Then,*

$$|p(k) - p(j)|_M = |p(\sigma(k)) - p(\sigma(j))|_M$$

*and*

$$\|p(k) - p(j)\| = \|p(\sigma(k)) - p(\sigma(j))\|.$$

PROOF. By Proposition 3.3.11 we know that

$$|p(k) - p(j)|_M = |p(\sigma(j)) - p(\sigma(k))|_M.$$

Since $|\cdot|$ is symmetric we get

$$|p(k) - p(j)|_M = |p(\sigma(k)) - p(\sigma(j))|_M.$$

Since $\|\cdot\|$ is a function of $|\cdot|_M$ it follows that

$$\|p(k) - p(j)\| = \|p(\sigma(k)) - p(\sigma(j))\|.$$

□

This next proposition is the main result of this section. It tells allows us to predict where a card $i$ will be after $t$ steps of the shuffle using the Heads-Tails differentials of $i$ from the times it is in position $m$ and in the bottom part of the deck.

PROPOSITION 3.3.13. *Let $i$ be a card in the deck. Let $B$ be the record of flips when $i$ is in position $m$. Let the sequence $S$ be the record of flips when $i$ is in the bottom part of the deck. Let $x$ be the Heads-Tails differential of $S$ after $t$ steps of the shuffle. Let $y$ be the Heads-Tails differential of $B$ after $t$ steps of the shuffle. Then,*

$$\left\| p(i_t) - p(i) - \left( t - \left\lfloor \frac{t}{2n} \right\rfloor (m-1) - x - \left\lfloor y\left(1 - \frac{m}{2n}\right) \right\rfloor m \right) \right\| \leq \left|\frac{y}{2}\right| + \left|\frac{x}{2n}\right| (\sqrt{n}+1) + 4\sqrt{n}.$$

PROOF. By Theorem 3.3.6 we know

$$p(i_t) \equiv p(i) + t + (T_S - H_S) + (T_B - mH_B) \bmod (2n - m + 1). \tag{3.1}$$

Note that

$$H_B = \frac{1}{2}(\text{number of times } i \text{ is in position } m \text{ in the first } t \text{ steps}) + \frac{1}{2}y.$$

Imagine card $i$ has just reached position $m$ for the $T$th time. We are interested in how many steps it takes for $i$ return to position $m$ again (for the $(T+1)$th time). If the next flip is heads, then $i$ will move from position $m$ to position 1. It will then take $m-1$ more steps for $i$ to return to position $m$. This is a total of $m$ steps, and we call this a short return. If instead the next flip is tails, then $i$ will move to position $m+1$. It will then take $2(n-m)+\Delta_T$ steps for $i$ to cycle through the bottom of the deck back to position 1. $\Delta_T$ represents the deviation from the expected number of steps, and $\Delta_T$ contributes negatively to the Heads-Tails differential of $S$. Namely we have $x = -\Delta_1 - \Delta_2 - \dots$. After card $i$ reaches position 1 it will take $m-1$ more steps to reach position $m$. This is a total of $2n-m+\Delta_T$ steps and we call this a long return. Note that, ignoring $\Delta_T$, the average of the number of steps between the short return and long return is $n$ steps. Thus, if $i_0 = m$ and $y$ is nonnegative then the number of subsequent times $i$ hits position $m$ is

$$\left\lfloor y + \frac{t - my - x}{n} \right\rfloor.$$

This is because other than the excess $y$ short returns which take $m$ steps, the remaining visits are divided evenly between long and short returns, so $n$ steps before counting the increased/decreased speed through the bottom part of the deck as recorded by $x$. Similarly if $y$ is negative then the number of subsequent times $i$ hits position $m$ in $t$ steps is

$$\left\lfloor (-y) + \frac{t - (2n-m)(-y) - x}{n} \right\rfloor.$$

This means that

$$\begin{aligned} H_B &= \frac{y}{2} + \frac{t}{2n} - \frac{my}{2n} - \frac{x}{2n} + \zeta + \frac{y}{2} \\ &= \frac{t}{2n} + y\Big(1 - \frac{m}{2n}\Big) - \frac{x}{2n} + \zeta \end{aligned} \tag{3.2}$$

if $y$ is positive and

$$\begin{aligned} H_B &= \frac{-y}{2} + \frac{t}{2n} - \frac{(2n-m)(-y)}{2n} - \frac{x}{2n} + \zeta + \frac{y}{2} \\ &= \frac{t}{2n} + y\Big(1 - \frac{m}{2n}\Big) - \frac{x}{2n} + \zeta \end{aligned} \tag{3.3}$$

if $y$ is negative, where $\zeta \in [-\frac{1}{2}, 0]$ and rounds down so that $H_B$ is an integer. Note that both lines (3.2) and (3.3) are equal, so we have the same value regardless of if $y$ is positive or negative. These

values represent the case where $i_0 = m$ and we don't count the fact that $i$ starts at $m$ as an $m$ "hit". At the other extreme, where $i$ starts at $m$ and we do count that as a "hit" we would have

$$H_B = \frac{t}{2n} + y\Big(1 - \frac{m}{2n}\Big) - \frac{x}{2n} + \zeta + 1.$$

In the more general case where we have $i_0 \neq m$ and we count the number of times $i$ hits $m$ we get

$$H_B = \frac{t}{2n} + y\Big(1 - \frac{m}{2n}\Big) - \frac{x}{2n} + \upsilon$$

where $\upsilon \in [-\frac{1}{2}, 1]$. We can use the fact that

$$T_B = \frac{1}{2}(\text{number of times } i \text{ is in position } m \text{ in the first } t \text{ steps}) - \frac{1}{2}y$$

to similarly calculate

$$T_B = \frac{t}{2n} - y\Big(\frac{m}{2n}\Big) - \frac{x}{2n} + \upsilon'$$

where $\upsilon' \in [-\frac{1}{2}, 1]$. Plugging into (3.1) we get

$$p(i_t) \equiv p(i) + t - x + \left(\frac{t}{2n} - y\Big(\frac{m}{2n}\Big) - \frac{x}{2n} + \upsilon'\right) - m\left[\frac{t}{2n} + y\Big(1 - \frac{m}{2n}\Big) - \frac{x}{2n} + \upsilon\right]$$

where the terms in square brackets form an integer. Note that for any integer $z$ we have that $z = \lfloor z \rfloor + \delta$ where $\delta \in [0, 1)$. So,

$$p(i_t) \equiv p(i)+t-x+\left\lfloor \frac{t}{2n} \right\rfloor (1-m) - \left\lfloor y\Big(1 - \frac{m}{2n}\Big)\right\rfloor m + \Big(\delta_1 - y\Big(\frac{m}{2n}\Big) - \frac{x}{2n} + \upsilon'\Big) - m\left[\delta_2 + \delta_3 - \frac{x}{2n} + \upsilon\right]$$

where $\delta_1, \delta_2, \delta_3 \in [0, 1)$. This gives us

$$\begin{aligned} &\left\| p(i_t) - p(i) - t + \left\lfloor \frac{t}{2n} \right\rfloor (m-1) + x + \left\lfloor y\Big(1 - \frac{m}{2n}\Big)\right\rfloor m \right\| \\ \leq &\left\| \delta_1 - y\left(\frac{m}{2n}\right) - \frac{x}{2n} + \upsilon' - m\left[\delta_2 + \delta_3 - \frac{x}{2n} + \upsilon\right] \right\| \\ \leq &\left| \delta_1 - y\left(\frac{m}{2n}\right) - \frac{x}{2n} + \upsilon' \right| + \sqrt{n}\left| \delta_2 + \delta_3 - \frac{x}{2n} + \upsilon \right|. \end{aligned}$$

Utilize the fact that $|\delta_1|, |\delta_2|, |\delta_3|, |\upsilon|, |\upsilon'| \leq 1$ and $m < n$ we get

$$\left\| p(i_t) - p(i) - t + \left\lfloor \frac{t}{2n} \right\rfloor (m-1) + x + \left\lfloor y\Big(1 - \frac{m}{2n}\Big)\right\rfloor m \right\| \leq \left|\frac{y}{2}\right| + \left|\frac{x}{2n}\right| (\sqrt{n} + 1) + 4\sqrt{n}.$$

□

The norm on the left hand side of the inequality of Proposition 3.3.13 is quite complicated. To make our lives easier, it will be nice to consider values of $t$ such that $-t + \left\lfloor \frac{t}{2n} \right\rfloor (m-1) \equiv 0 \mod 2n - m + 1$ thus eliminating those two terms from the norm. To make sure this is possible we have the following lemma.

LEMMA 3.3.14. *Fix some $n, m \in \mathbb{N}$ such that $m < n$. Choose any $s \in \mathbb{N}$. Then there exists $t \in [s, s + 4n]$ such that*

$$-t + \left\lfloor \frac{t}{2n} \right\rfloor (m-1) \equiv 0 \mod 2n - m + 1$$

PROOF. Let $s^* \in [s, s + 2n)$ such that $s^*$ is a multiple of $2n$. Then for all $t^*$ in the interval $[s^*, s^* + 2n)$ we have

$$\left( -(t^* + 1) + \left\lfloor \frac{(t^* + 1)}{2n} \right\rfloor (m-1) \right) = \left( -t^* + \left\lfloor \frac{t^* + 1}{2n} \right\rfloor (m-1) \right) - 1$$

Since there are $2n$ terms of the interval and it will take at most $2n - m + 1$ unit steps to get to 0 mod $2n - m + 1$ we know there exists $t \in [s^*, s^* + 2n)$ such that

$$-t + \left\lfloor \frac{t}{2n} \right\rfloor (m-1) \equiv 0 \mod 2n - m + 1$$

□

## 3.4. Entropy and 3-Monte

We will use techniques involving entropy to bound the mixing time of the overlapping cycles shuffle. In this section, we provide the necessary background in entropy. We utilize a new technique involving entropy, the 3-Monte, first described by Senda [10], which is a generalization of a similar technique first described by Morris [9].

DEFINITION 3.4.1. *If $\pi$ is a random permutation in $S_n$, we define the relative entropy of $\pi$ with respect to the uniform distribution as*

$$\mathrm{ENT}(\pi) = \sum_{\varphi \in S_n} \mathbb{P}(\pi = \varphi) \log(n! \cdot \mathbb{P}(\pi = \varphi)).$$

Note that in the case that $\pi$ is uniform, we have $\mathrm{ENT}(\pi) = 0$. In the other extreme where $\pi$ is deterministic, we have $\mathrm{ENT}(\pi) = \log(n!)$.

This notion of relative entropy is useful, because we have by Pinsker's inequality [11] (Section 2.4, page 88) that

$$\sqrt{\frac{1}{2}\mathrm{ENT}(\pi)} \geq \|\pi - \xi\|_{\mathrm{TV}} \tag{3.4}$$

where $\xi$ is the uniform random permutation in $S_n$. We now define the notion of 3-Monte shuffles and collisions, which will be useful for finding bounds on relative entropy.

Definition 3.4.2. *We say a random permutation $\mu$ in $S_n$ is a 3-collision if for some distinct $x, y, z \in \{1, \ldots, n\}$ it is equally likely to be either the 3-cycle $(x, y, z)$ or the identity. So it has the distribution*

$$\mathbb{P}(\mu = (x, y, z)) = \mathbb{P}(\mu = \mathrm{id}) = \frac{1}{2}$$

*In particular we say $\mu = c(x, y, z)$ is the 3-collision which has a one half chance of being the $(x, y, z)$ 3-cycle.*

We will be interested in random permutations which can be written as products containing 3-collisions.

Definition 3.4.3. *We say a random permutation $\pi$ in $S_n$ is 3-Monte if it has the form*

$$\pi = \nu c(x_k, y_k, z_k) \ldots c(x_1, y_1, z_1)$$

*where $\nu$ is a random permutation and $x_1, x_2, x_3, \ldots, x_k, y_k, z_k$ and $k$ itself may be dependent on $\nu$, but conditional on $\nu$ the outcomes of $c(x_1, y_1, z_1), \ldots, c(x_k, y_k, z_k)$ must be independent.*

To be clear, any random permutation is technically 3-Monte, as $k$ could be trivially set to 0 conditioned on any $\nu$. Additionally, the same random permutation could be defined with different choices for the 3-collision. We will refer to random permutations as 3-Monte only after explicitly choosing 3-collisions and defining the permutation as a product involving those collisions. The following theorem regarding 3-Monte shuffles will allow us to bound entropy decay of the overlapping cycles shuffle.

THEOREM 3.4.4. *[10] (Chapter 4, page 16) Let $\pi$ be a 3-Monte shuffle on $n$ cards. Fix an integer $t > 0$ and suppose that $T$ is a random variable taking values in $\{1, \cdots, t\}$, which is independent of the shuffles $\{\pi_i : i \geq 0\}$. Consider a card $x$. If $x$ is involved in a 3-collision after time $T$ up to and including time $t$, then consider the first 3-collision it is involved in after time $T$; say that $x, y, z$ collide in that order. If that 3-collision is also the first 3-collision after time $T$ that $y$ is involved in and it is the first 3-collision that $z$ is involved in, then we say that* ***that collision matches*** *$x$ (with $y$ and $z$) and define $y$ to be the* ***front match*** *of $x$, written as $m_1(x) = y$, and $z$ to be the* ***back match*** *of $x$, written $m_2(x) = z$. If $x$ is in no such collision, define $m_1(x) = m_2(x) = x$. Suppose that for every card $i$ there is a constant $A_i \in [0, 1]$ such that $\mathbb{P}(m_2(i) = j, m_1(i) < i) \geq \frac{A_i}{i}$ for each $j \in \{1, \ldots, i-1\}$. (This also means that with probability at least $A_i$, $m_1(i) < i$ and $m_2(i) < i$. Note that it cannot be the case that exactly two of $i$, $m_1(i)$, and $m_2(i)$ are equal; the three are either all the same or all different.) Let $\mu$ be an arbitrary random permutation that is independent of $\{\pi_i : i \geq 0\}$. Then*

$$\mathbb{E}\big[\mathrm{ENT}(\pi_t\mu|\mathrm{sgn}(\pi_t\mu))\big] - \mathbb{E}\big[\mathrm{ENT}((\mu|\mathrm{sgn}(\mu)))\big] \leq \frac{-C}{\log(n)} \sum_{x=3}^{n} A_x E_x,$$

*where $E_k = \mathbb{E}\big[\mathrm{ENT}(\mu^{-1}(k) \mid \mu^{-1}(k+1), \mu^{-1}(k+2), \ldots, \mu^{-1}(n), \mathrm{sgn}(\mu))\big]$ and $C$ is a positive universal constant.*

The exact use of this theorem will be made apparent in Section 3.6. For now just know that our immanent goal is to bound the values $A_x$ from below.

We can write two steps of the overlapping cycles shuffle in 3-Monte form as follows: Let $\pi$ be the random permutation corresponding to two steps of the overlapping cycles shuffle. Then $\pi$ has the following distribution:

$$\mathbb{P}(\pi = (1, \ldots, m)(1, \ldots, m)) = \frac{1}{4}$$
$$\mathbb{P}(\pi = (1, \ldots, n)(1, \ldots, m)) = \frac{1}{4}$$
$$\mathbb{P}(\pi = (1, \ldots, m)(1, \ldots, n)) = \frac{1}{4}$$
$$\mathbb{P}(\pi = (1, \ldots, n)(1, \ldots, n)) = \frac{1}{4}$$

Note that $(1,\dots,n)(1,\dots,m) = (1,\dots,m)(1,\dots,n)(m-1,m,n)$. Thus, we can rewrite the distribution of $\pi$ in the following way:

$$\mathbb{P}(\pi = (1,\dots,m)^2) = \frac{1}{4}$$
$$\mathbb{P}(\pi = (1,\dots,n)^2) = \frac{1}{4}$$
$$\mathbb{P}(\pi = (1,\dots,m)(1,\dots,n)c(m-1,m,n)) = \frac{1}{2}$$

If we are being precise, this is not technically a definition of collisions for the overlapping cycles shuffle. We have defined collisions for the shuffle which is two steps at a time of the overlapping cycles shuffle. It will be inconvenient to from now on imagine that we do two steps at a time, so instead we will continue to consider the standard one-step-at-a-time overlapping cycles shuffle and say that the cards in positions $m-1, m, n$ are in collision at time $t$ if

- $t$ is even
- the coins flipped in steps $t, t+1$ land Heads, Tails or Tails, Heads.

To apply Theorem 3.4.4 to the overlapping cycles shuffle, we will need to examine the probabilities that cards $i, k, j$ end up in positions of $m-1, m, n$ respectively after an even number of steps. As a warm up we will first deal with a few special cases that elude the parameters of the general theorem in Section 3.6.

LEMMA 3.4.5. *Consider the overlapping cycles shuffle on $n$ cards where $m \in (10\sqrt{n}, n-10\sqrt{n})$. Let $i, j, k$ be cards in $(n-\sqrt{n}, n]$ such that $k = i+1$ and $j > i$. Let $T = 2n-5\sqrt{n}$ and $t = 2n+5\sqrt{n}+2$. Let $E$ be the event that the first time $i$ or $j$ or $k$ experiences a collision after time $T$, the collision is before time $t$ and with each other in the order $(i,k,j)$. Then*

$$\mathbb{P}(E) \geq \frac{D}{\sqrt{n}} \exp\left(-\frac{2n}{m}\right).$$

*for a universal constant $D$*

The idea for the proof is that we will show that with probability bounded away from 0 that $i$ and $k$ stay "glued together". Note that since $i$ and $k$ begin adjacent to each other in the part bottom of the deck, they will stay adjacent at least until one of them leaves the bottom part of the deck. Just before this, if $i_r = n-1$ and $k_r = n$, if the next two flips are tails, then we will have $i_{r+2} = 1$

and $k_{r+2} = 2$. So $i$ stays one position above $k$. If $i$ always copies the coin that $k$ uses when in positions $m$ and $n$ then $i$ will always stay in the position above $k$. It then suffices to show that when $j$ reaches position $n$ that $i$ reaches position $m-1$, because if $k$ stays "glued" then we will also have $k$ in position $m$. These are the correct positions for $i, k, j$ to collide in that order. Based on Corollary 3.3.7 we know that there on the order of $\sqrt{n}$ positions nearby position $n$ with respect to $\|\cdot\|$ which $i$ is likely to be in when $j$ reaches position $n$. Since $\|p(m-1) - p(n)\| = \sqrt{n}$ we see that $m-1$ is one of these positions.

PROOF. Let $H_1$ be the event that Tails is flipped the first time $i$ and $k$ are in position $m$, and Heads is flipped the first time $j$ is in position $m$. Then,

$$\mathbb{P}(H_1) = \frac{1}{8}.$$

Let $\tau_1$ be the random stopping time given by the minimum time $t$ such that $j_t = m$. Then

$$\tau_1 = m + 1 + (n - j) + \theta_1$$

where $\theta_1$ is the number of Heads flipped before $n - j$ Tails are flipped. Note that $\theta_1$ is a negative binomial random variable with a mean of $n-j$ and a standard deviation of approximately $\sqrt{n-j} \leq \sqrt[4]{n}$. Then, conditioned on $H_1$ we know at time $\tau_1$ that $j$ is in position 1 and $i, k$ are in the bottom part of the deck. Let $\tau_2$ be the minimum $t > \tau_1$ such that $j_t = m$. Then conditioned on $H_1$ we have

$$\tau_2 - \tau_1 = m.$$

In these $m$ steps $i$ and $k$ are adjacent to each other in the bottom part of the deck. Let $H_2$ be the event that Tails is flipped following $\tau_2$ (so that $j$ enters the bottom part of the deck), and that Tails is flipped immediately after $k$ reaches position 1 (so that $i$ follows and stays one position above $k$). Then,

$$\mathbb{P}(H_2 \mid H_1) \geq \frac{1}{4}.$$

Let $\tau_3$ be the random stopping time given by the minimum $t > \tau_2$ such that $j_t = n$. Then conditioned on $H_1, H_2$ we have

$$\tau_3 - \tau_2 = 1 + (n - m) + \theta_2$$

where $\theta_2$ is the number of Heads before $(n-m)$ Tails are flipped after $\tau_2$. Note that $\theta_2$ is a negative binomial random variable and has a mean of $(n-m)$ and a standard deviation of approximately $\sqrt{n-m} \leq \sqrt{n}$. Now let $Q_1$ be the event that $(i_{\tau_3}, k_{\tau_3}) = (m-1, m)$ and let $Q_2$ be the event that $\theta_1 + \theta_2 - (n-j) - (n-m) \in (-5\sqrt{n}, 5\sqrt{n})$. We claim that

$$\mathbb{P}\left(Q_1, Q_2 \mid H_1, H_2\right) \geq \frac{D_1}{\sqrt{n}} \exp\left(-\frac{2n}{m}\right)$$

for a universal constant $D_1$. To see why, let $S^{ik}$ be the record of coins flipped when $i$ and $k$ are in the bottom part of the deck and $j$ is not in the bottom part of the deck, which in particular happens for at least $m$ steps between times $\tau_2$ and $\tau_3$. Let $S^j$ be the record of coins flipped when $j$ is alone in the bottom part of the deck. Then $Q_1$ will occur as long as the Heads-Tails differentials of $S^{ik}$ and $S^j$ combine in such a way that the original gap of $p(j_0) - p(i_0)$ closes at $\tau_3$. Let $\Delta_i$ be the number of Tails minus the number of Heads drawn from $S^{ik}$ before $\tau_3$. Let $\Delta_j$ be defined similarly for $S^j$. Since at least $m$ coins will be used from $S^{ik}$ and no more than $n$ coins will be used from each of $S^{ik}, S^j$ we know that the distribution of $\Delta_j - \Delta_i$ is well approximated by a binomial random variable with probability $\frac{1}{2}$ chance of success and some number of trials between $m$ and $2n$. The standard deviation of such random a random variable (and its sum) will be at least $\sqrt{m}$. This is compared to $p(j_0) - p(i_0) \leq 2n$. So we need $\Delta_j - \Delta_i$ to equal a specific number no more than $\frac{\sqrt{2n}}{\sqrt{m}}$ standard deviations away. So we have

$$\mathbb{P}\left(\Delta_j - \Delta_i = p(j_0) - p(i_0) \mid H_1, H_2\right) \geq \frac{D_1}{\sqrt{n}} \exp\left(-\frac{2n}{m}\right).$$

By Corollary 3.3.6 we have $p(j_{\tau_3}) - p(i_{\tau_3}) \equiv p(j_0) - p(i_0) - \Delta_j + \Delta_i - m$ on $H_1, H_2$. So $Q_1$ follows from $\mathbb{P}\left(\Delta_j - \Delta_i = p(j_0) - p(i_0)\right)$. The event $Q_2$ also follows because $\theta_1$ and $\theta_2$ are derived from the surplus Tails flipped in $\Delta_i, \Delta_j$ and $|p(j_0) - p(i_0)| \leq 2\sqrt{n}$.

Note that on $H_1, H_2, Q_1, Q_2$ we have that $i, j, k$ are in the perfect position to collide in the order of $(i, k, j)$ after step $\tau_3$ with $\tau_3 \in (2n+1-5\sqrt{n}, 2n+1+5\sqrt{n})$. Also note that on these events, $i, j, k$ experience no collisions between time $2n - 5\sqrt{n}$ and $\tau_3$ because $i, k$ are in the top part of the deck throughout this interval and $j$ is in the bottom part of the deck throughout this interval. Now, as long as $\tau_3$ is even, there is a $\frac{1}{2}$ chance that we have a $(i, k, j)$ collision over the next two

steps. Let $G$ be the probability that $\tau_3$ is even. Then,

$$\mathbb{P}(G \mid H_1, H_2, Q_1, Q_2) \geq \frac{1}{3}.$$

This is because $i, j, k$ start out in the bottom part of the deck. With probability $\frac{1}{2}$ the first flip is tails and then $(i_1, j_1, k_1) = (i_0, j_0, k_0)$. So the chance that $\tau_3$ is odd cannot be more than twice the chance it is even. Together, this means that

$$\mathbb{P}(E) \geq \frac{1}{2} \cdot \frac{1}{3} \cdot \frac{D_1}{\sqrt{n}} \exp\left(-\frac{2n}{m}\right).$$

□

COROLLARY 3.4.6. *Consider the overlapping cycles shuffle on $n$ cards where $m \in (10\sqrt{n}, n-10\sqrt{n})$. Let $i, j, k$ be cards in $(n-\sqrt{n}, n]$ such that $k = i+2$ and $j = i+1$. Let $T = 2n - 5\sqrt{n}$ and $t = 2n + 5\sqrt{n} + 2$. Let $E$ be the event that the first time $i$ or $j$ or $k$ experiences a collision after time $T$, the collision is before time $t$ and with each other in the order $(i, k, j)$. Then*

$$\mathbb{P}(E) \geq \frac{D}{\sqrt{n}} \exp\left(-\frac{2n}{m}\right).$$

PROOF. The proof is nearly identical to the one in Lemma 3.4.5. In that proof we required $i, j, k$ to flip Tails, Heads, Tails respectively for each of their first visits to $m$. We require the same in this new case. Since $i, j, k$ start out adjacent to each other in that order, with probability $\frac{1}{8}$ they will all flip Tails when in position $n$ and so stay adjacent as they move to the top of the deck. Then, after $m-2$ more steps we will have $i, j, k$ in positions $m-2, m-1, m$ respectively. Then after the flips Tails, Heads, Tails we have $i, j, k$ in positions $m+1, 2, m+2$ respectively. This is exactly the same situation as in the proof of Lemma 3.4.5 after $\tau_1$. The rest of the proof is the same and the equivalent result holds. □

These two results tell us that if $i, j \in (n-\sqrt{n}, n]$ and $j > i$ then it is reasonably likely that $m_2(i) = j$ and $m_1(i) > i$ with regards to the notation in Theorem 3.4.4. We make this precise in the proposition below. Note that these inequalities are the opposite of those required in the Theorem, but this is okay. Our choice of labeling card 1 as the card in the top of the deck, and card 2 as second to top, etc was arbitrary. We can utilize Theorem 3.4.4 where the inequalities are with respect to any well-ordering of the deck, and we will in fact use a well-ordering later which starts its

count from the bottom of the deck upwards. As further justification that this is a valid application of Theorem 3.4.4, note that entropy decay will be the same if we first use some deterministic change of basis permutation to reorder the cards and then apply the overlapping cycles shuffle.

PROPOSITION 3.4.7. *Fix cards $i, j \in (n - \sqrt{n}, n]$ such that $i \geq n - 2$ and $i < j$. Let $T = 2n - 5\sqrt{n}$ and $t = 2n + 5\sqrt{n}$. Let $m_1(i)$ be front match of $i$ the and let $m_2(i)$ be the back match, as defined by Theorem 3.4.4. Then,*

$$\mathbb{P}\left(m_2(i) = j, m_1(i) > i\right) \geq \frac{D}{\sqrt{n}} \exp\left(-\frac{2n}{m}\right).$$

*for a universal constant $D$.*

PROOF. In the case that $j \neq i + 1$ we have by Lemma 3.4.5 that

$$\mathbb{P}\left(m_2(i) = j, m_1(i) > i\right) \geq \mathbb{P}\left(m_2(i) = j, m_1(i) = i + 1\right) \geq \frac{D_1}{\sqrt{n}} \exp\left(-\frac{2n}{m}\right).$$

In the case that $j = i + 1$ we have by Corollary 3.4.6 that

$$\mathbb{P}\left(m_2(i) = i + 1, m_1(i) > i\right) \geq \mathbb{P}\left(m_2(i) = i + 1, m_1(i) = i + 2\right) \geq \frac{D_2}{\sqrt{n}} \exp\left(-\frac{2n}{m}\right).$$

Taking $D = \min\{D_1, D_2\}$ completes the proof. □

## 3.5. Movement of 3 Cards

The goal of this section will be to show that after $t$ steps $i, j$ and $k$ have an approximately uniform distribution over all "nearby" positions, where our notion of "near" is related to the size of $t$. More specifically, we will show that after about $4\ell^2$ steps the cards $i, j, k$ are distributed approximately uniformly amongst positions $f_i, f_j, f_k$ such that $\|p(i) - p(f_i)\|, \|p(j) - p(f_j)\|, \|p(k) - p(f_k)\| < \ell$.

In order to prove things about the movement of cards $i, j$ and $k$ relative to each other, it will be useful to imagine that instead of using a single sequence of coins to determine if the card in position $m$ or $n$ is moved to the top in each step, we generate many sequence of coins, and choose which sequence to draw from in each step according to the state the deck is in. Here is one way of doing this:

We generate coin sequences $B^i, B^j, B^k, S^i, S^j, S^k, S^{ij}, S^{jk}, S^{ik}, S^{ijk}$ and simulate the movement of cards $i, j, k$ under the overlapping cycles shuffle according to the following rules:

- If card $i$ is in position $m$, then use the next coin from $B^i$ to do the shuffle. Similarly, if $j$ or $k$ is in position $m$, use $B^j$ or $B^k$.
- Otherwise, use $S^A$ where $A \subseteq \{i, j, k\}$ is the set of which $i, j, k$ are in the bottom part of the deck. If all of $i, j, k$ are in the top part of the deck then these three cards will move down deterministically in the next step and no coin is necessary.

We will be able to bound the movement of $i, j$ and $k$ by putting restrictions on $S^i, S^j, S^k, S^{ij}, S^{jk}, S^{ik}$, and $S^{ijk}$. However, it may be the case that some of these sequences are drawn from more than others. For example, if the entire sequence of $B^i$ is made up of Heads, then $i$ will spend all its time in the top part of the deck and coins from $S^i, S^{ij}, S^{ik}, S^{ijk}$ will not be used at all! We need to make sure something like this does not happen. In particular, it will be important to show that $S^i, S^j, S^k$ are each drawn from a constant proportion of the time. This is equivalent to saying that $i, j, k$ each spend a constant proportion of time alone in the bottom part of the deck. Fortunately, this happens with high probability.

LEMMA 3.5.1. *Consider the overlapping cycles shuffle on $n$ cards for sufficiently large $n$. Fix any cards $i, j, k$. Then with probability greater than $\frac{1}{8}\exp\left(-\frac{10n}{m}\right)$, at least at least $n-m$ out of the next $5n$ steps have $i$ in the bottom part of the deck and $j$ and $k$ in the top part of the deck.*

PROOF. Let $A$ be the event that every time $j$ and $k$ are in position $m$ in the next $5n$ steps, a Heads is flipped sending $j$ and $k$ respectively back to position 1. Since $j$ and $k$ each make at most $\frac{5n}{m}$ visits to position $m$ in $5n$ steps we have

$$\mathbb{P}(A) \geq \left(2^{-\frac{5n}{m}}\right)^2 \geq \exp\left(-\frac{10n}{m}\right).$$

Let $B$ be the event that the next two times $i$ is in position $m$, Tails is flipped sending $i$ to the bottom part of the deck. Then $\mathbb{P}(B) = \frac{1}{4}$. $A$ and $B$ are independent, so

$$\mathbb{P}(A, B) \geq \frac{1}{4}\exp\left(-\frac{10n}{m}\right).$$

Let $G$ be the event that out of the next $5n$ coin flips, no more than $3n$ are Heads. This happens with exponentially high probability, but we will just use that for large enough $n$, we have $\mathbb{P}(G) \geq$

$1 - \frac{1}{8}\exp\left(-\frac{10n}{m}\right)$. Then by the union bound we have,

$$\mathbb{P}(A, B, G) \geq \frac{1}{8}\exp\left(-\frac{10n}{m}\right).$$

On the events $A, B, G$ it must be that $i$ spends at least $n - m$ steps alone in the bottom part of the deck. This is because $G$ guarantees that, in the case that $j$ or $k$ begin in the bottom part of the deck, they will move swiftly to position $n$ and then wrap around to position 1. Then $A$ guarantees that $j$ and $k$ will continue to cycle through the top part of the deck, and $B$ guarantees that $i$ moves alone into the bottom part of the deck. Since there are $n - m$ positions in the bottoms part of the deck, and $C$ guarantees that $2n$ Tails are flipped to facilitate $i$ through at least 2 complete laps of the deck (at least one where it is alone in the bottom) we have proven the lemma. □

COROLLARY 3.5.2. *Consider the overlapping cycles shuffle on $n$ cards for sufficiently large $n$. For any $a \in \mathbb{N}$, after $t = 6an$ steps, the probability that each of $i, j, k$ spend at least $\frac{a(n-m)}{16}\exp\left(-\frac{10n}{m}\right)$ of these steps alone in the bottom part of the deck is at least*

$$1 - 3\exp\left(-\frac{a}{32\exp\left(\frac{10n}{m}\right)}\right).$$

PROOF. We can divide $t$ into $a$ steps of size $6n$. In each of these blocks of $6n$ steps, with probability greater than $\frac{1}{8}\exp\left(-\frac{10n}{m}\right)$ card $i$ spends at least $n - m$ steps in the bottom part of the deck. Let $Y$ be the number of steps spent by $i$ in the bottom part of the deck in $t$ steps. Then $\frac{Y}{n-m}$ stochastically dominates a Binomial$\left(a, \frac{1}{8}\exp\left(-\frac{10n}{m}\right)\right)$ random variable. We can use Theorem A.0.4, Hoeffing's inequality on the binomial random variable to get

$$\mathbb{P}\left(\frac{Y}{n-m} \leq \frac{a}{16\exp\left(\frac{10n}{m}\right)}\right) \leq \exp\left(-\frac{a}{32\exp\left(\frac{10n}{m}\right)}\right).$$

Thus with high probability $i$ spends a bounded away from 0 proportion of time alone in the bottom of the deck. In fact the probability that this does not happen is exponentially low. The same applies for $j$ and $k$ so using the union bound we get the result of the corollary. □

In the following Proposition we aim to simplify Corollary 3.5.2. We do this by defining a constant $L$ which depends on the ratios $\frac{n}{m}$ and $\frac{n}{n-m}$ which absorbs most of the complicated coefficents in the parameters and bound of the Corollary.

PROPOSITION 3.5.3. *Consider the overlapping cycles shuffle on $n$ cards for sufficiently large $n$. Let*

$$L = L(n,m) = \frac{192n\exp(\frac{10n}{m})}{(n-m)}$$

*Choose some $C \in \mathbb{N}$ such that $C \geq 1$. After $LCn$ steps, the probability that each of $i,j,k$ spend at least $Cn$ steps alone in the bottom part of the deck is at least*

$$1 - 3\exp\left(-\frac{1}{2}C\right).$$

PROOF. We will use the previous Corollary 3.5.2, setting $t = LCn$. Then, by that Corollary,

$$a = \left\lfloor \frac{LCn}{6n} \right\rfloor.$$

Since $L > 12$ and $C \geq 1$, we have $a \geq \frac{LC}{12}$. Note that

$$\frac{a(n-m)}{16}\exp\left(-\frac{10n}{m}\right) > Cn$$

and

$$\frac{a}{32\exp(\frac{10n}{m})} > \frac{Cn}{2(n-m)} \geq \frac{1}{2}C.$$

So the probability that $i,j,k$ each spend at least $t$ steps alone in the bottom part of the deck is at least

$$1 - 3\exp\left(-\frac{1}{2}C\right).$$

□

As previously mentioned, our goal in this section will be to control the movement of cards $i,j,k$. To do this, we will consider 3 "stages". In Stage 1, we will show that $i,j,k$ spread out from each other in terms of $\|\cdot\|$. In this first stage we will not need to make precise statements about exactly the positions $i,j,k$ travel to. It will be enough to guarantee that they have a gap between each other on the order of $\ell$ after $\ell^2$ steps.

In Stage 2 we will use a coupling argument to show that, provided that $i,j,k$ are spread out

from each other, their distribution after $\ell^2$ steps is approximately uniform over cards within distance $\ell$ (in the $\|\cdot\|$ sense) to $i, j, k$ respectively.

Stage 3 is the inverse of the first stage. We show that $i, k, k$ spread out under the inverse overlapping cycles shuffle as well. To put the three stages together, we use Stage 1 to show that $i, j, k$ spread out, Stage 2 to show that $i, j, k$ move precisely to locations likely to hit our desired targets, and Stage 3 to show $i, j, k$ move back together to hit their targets.

This next Proposition is useful for Stage 1 because it shows we can move $i, j, k$ nearby to targets which we will later choose to be spread out from each other.

PROPOSITION 3.5.4. *There exist universal constants $C, D$ such that the following holds: Fix $\ell$ such that $CL\sqrt{n} \leq \ell \leq n$. Fix any positions $i, j, k, f_i, f_j, f_k$ such that*

$$\|p(i) - p(f_i)\|, \|p(j) - p(f_j)\|, \|p(k) - p(f_k)\| \leq 3\ell.$$

*Choose $T \in [\ell^2, \ell^2 + 4n]$ such that $T + \left\lfloor \frac{T}{2n} \right\rfloor (m-1) \equiv 0 \mod 2n - m + 1$. Then,*

$$\mathbb{P}\left( \|p(i_T) - p(f_i)\|, \|p(j_T) - p(f_j)\|, \|p(k_T) - p(f_k)\| < \frac{\ell}{2000} \right) \geq D \exp(-432L^2).$$

PROOF. Since $\|p(i) - p(f_i)\| < 3\ell$ we know that $p(f_i) = p(i) + a_i + b_i m$ where $|a_i| < 3\ell$ and $|b_i| < \frac{3\ell}{\sqrt{n}}$. Suppose $B^i$ is the record of the coins card $i$ flips whenever it is in position $m$. Let $G_i$ be the event that the following holds:

- There exist $r \leq \frac{T}{2n}$ such that $\text{Diff}_r(B^i) = y_i$ where $\left\lfloor y_i\left(1 - \frac{m-1}{2n}\right) \right\rfloor = b_i$.
- For all $s$ with $r \leq s \leq \frac{2T}{n}$ we have $\text{Diff}_r(B^i) \in (y_i - \frac{\ell}{8000\sqrt{n}}, y_i + \frac{\ell}{8000\sqrt{n}})$.

Note that $|y_i| < (1 - \frac{m-1}{2n})^{-1}|b_i| < \frac{6\ell}{\sqrt{n}}$. By Theorems A.0.3 and A.0.6 the event $G_i$ occurs with probability $C_2$ for some constant $C_2$. Similarly define events $G_j, G_k$ for cards $j$ and $k$. Since the coins used by $i, j, k$ when in position $m$ are independent, we have that

$$\mathbb{P}(G_i, G_j, G_k) \geq C_2^3.$$

Now imagine that whenever $i, j, k$ are not in position $m$, we use the following sequences of $T$ small coins: $S^i, S^j, S^k, S^{ij}, S^{ik}, S^{jk}, S^{ijk}$ as follows: If the coins in set $A$ are in the bottom part of the deck, use the next coin in the sequence $S^A$. Let $H$ be the event that

- $\text{Diff}_r(S^{ij}), \text{Diff}_r(S^{ik}), \text{Diff}_r(S^{jk}), \text{Diff}_r(S^{ijk}) < \frac{\ell}{16000}$ for all $r \leq T$.
- There exists $r_i, r_j, r_k \leq \frac{\ell^2}{L}$ such that $\text{Diff}_{r_i}(S^i) = a_i,\ \text{Diff}_{r_j}(S^j) = a_j,\ \text{Diff}_{r_k}(S^k) = a_k$.
- For all $s_i, s_j, s_k$ with $r_i \leq s_i \leq T,\ r_j \leq s_j \leq T,\ r_k \leq s_k \leq T$ we have

$$\begin{aligned}
\text{Diff}_{s_i}(S^i) &\in (a_i - \frac{\ell}{16000}, a_i + \frac{\ell}{16000}), \\
\text{Diff}_{s_j}(S^j) &\in (a_j - \frac{\ell}{16000}, a_j + \frac{\ell}{16000}), \\
\text{and } \text{Diff}_{s_k}(S^k) &\in (a_k - \frac{\ell}{16000}, a_k + \frac{\ell}{16000}).
\end{aligned}$$

Again by Theorems A.0.3 and A.0.6 we get,

$$\mathbb{P}(H) \geq C_3^{10} \frac{1}{15} \exp(-16(3L)^2)^3.$$

This is because we require $S^i, S^j, S^k$ to move at up to $3L$ standard deviations, and then remain within a constant number of standard deviations, and we require $S^{ij}, S^{ik}, S^{jk} S^{ijk}$ to remain within a constant number of standard deviations. Since the "$B$ coin sequences" are generated independently of the "$S$ coin sequences" we get

$$\mathbb{P}(G_i, G_j, G_k, H) \geq C_2^3 C_3^{10} \frac{1}{15} \exp(-16(3L)^2)^3 = C_4 \exp(-432L^2).$$

Finally, let $Q$ be the event that $i, j, k$ each spend at least $\frac{\ell^2}{L}$ steps in the bottom of the deck. By Corollary 3.5.3 we have that

$$\begin{aligned}
\mathbb{P}(Q) &\geq 1 - C_5 \exp\left(-\left(\frac{\ell^2}{n}\right)^2\right) \\
&\geq 1 - C_5 \exp(-L^2 C^2).
\end{aligned}$$

Using the fact that $L > 192$ pick a universal $C$ large enough that $C_5 \exp(-L^2C^2) < \frac{1}{2} C_4 \exp(-432L^2)$. Then by the union bound,

$$\mathbb{P}(G_i, G_j, G_k, H, Q) \geq \frac{1}{2} C_4 \exp(-432L^2).$$

On the events $G_i, G_j, G_k, H, Q$ we have by Proposition 3.3.13 that

$$\|p(i_T) - p(f_i)\| < 4 \cdot \frac{\ell}{16000} + \frac{\ell}{8000} + \frac{1}{2}\left(|y_i| + \frac{\ell}{8000\sqrt{n}}\right) + \frac{|a_i| + \frac{\ell}{8000}}{2n}(\sqrt{n}+1) + 4\sqrt{n}$$

where the $4 \cdot \frac{\ell}{16000}$ comes from the fact that $S^i, S^{ij}, S^{ik}, S^{ijk}$ each differ at most $\frac{1}{16000}$ from their target Heads-Tails differential and the $\frac{\ell}{8000}$ comes from the fact that $B^i$ differs at most $\frac{\ell}{8000\sqrt{n}}$ from its target Heads-Tails differential. Using that fact that $\ell \leq n$ and $|a_i| \leq 3\ell$ and $|y_i| \leq \frac{6\ell}{\sqrt{n}}$ we get

$$\|p(i_T) - p(f_i)\| < \frac{3\ell}{8000} + \frac{3\ell}{\sqrt{n}} + \frac{3\ell}{2n} + \frac{\ell}{16000\sqrt{n}} + \frac{3\ell}{2\sqrt{n}} + \frac{3\ell}{2n} + \frac{\ell}{16000\sqrt{n}} + \frac{\ell}{16000n} + 4\sqrt{n}.$$

Recall that $\sqrt{n} \leq \frac{1}{CL}\ell \leq \frac{1}{C}\ell$. Choose $C$ large enough that for large $n$ we have

$$\|p(i_T) - p(f_i)\| < \frac{\ell}{2000}.$$

A similar argument shows the equivalent results for $j$ and $k$ follow from $G_i, G_j, H_k, H, Q$ and this completes the proof. □

Stage 3 is Stage 1 in reverse, so we will need the equivalent result for the inverse overlapping cycles shuffle.

Corollary 3.5.5. *For the same universal constants $C, D$ in Proposition 3.5.4 the following holds: Fix $\ell$ such that $CL\sqrt{n} \leq \ell \leq n$. Fix any positions $i, j, k, f_i, f_j, f_k$ such that*

$$\|p(i) - p(f_i)\|, \|p(j) - p(f_j)\|, \|p(k) - p(f_k)\| \leq 3\ell.$$

*Choose $T \in [\ell^2, \ell^2 + 4n]$ such that $T + \left\lfloor \frac{T}{2n} \right\rfloor \equiv 0 \mod 2n - m + 1$. If $i_{(-T)}, j_{(-T)}, k_{(-T)}$ are the locations of $i, j, k$ after $T$ steps of the inverse overlapping cycles shuffle we have*

$$\mathbb{P}\left(\|p(i_{(-T)}) - p(f_i)\|, \|p(j_{(-T)}) - p(f_j)\|, \|p(k_{(-T)}) - p(f_k)\| < \frac{\ell}{2000}\right) \geq D\exp(-432L^2).$$

Proof. We know that the inverse overlapping cycles shuffle is the same as the forward overlapping cycles shuffle after reordering the cards. As Corollary 3.3.12 tells us $\|p(\cdot) - p(\cdot)\|$ is invariant under $\sigma$. Since all parameters in Proposition 3.5.4 are in terms of $\|p(\cdot) - p(\cdot)\|$ the equivalent statement also holds for the inverse overlapping cycles shuffle. □

The next several lemmas will work to prove our desired result for Stage 2: that if $i, j, k$ are spread out then they will have an approximately uniform distribution after $T$ steps. The first lemma tells us that for a reasonable choice $b_i, b_j, b_k$ we can low bound the probability that $i, j, k$ flip exactly $b_i, b_j, b_k$ Heads when in position $m$.

LEMMA 3.5.6. *Choose cards $i, j, k$. Fix some $\ell$ such that $2\sqrt{n} \leq \ell \leq \frac{1}{2}n$. Choose any $T \in [\ell^2, \ell^2+4n]$ such that $T + \lfloor \frac{T}{2n} \rfloor \equiv 0 \mod 2n - m + 1$. Choose $b_i, b_j, b_k \in \left(\frac{T}{2n} - \frac{\ell}{\sqrt{n}}, \frac{T}{2n} + \frac{\ell}{\sqrt{n}}\right)$.*

- *Let $R_i$ be the number of Heads flipped by $i$ while in position $m$ in the first $T$ steps.*
- *Let $E_i$ be the event that $|p(i) + T + b_i m - p(i_T)|_M \leq 6\ell$.*

*Let $R_j, R_k, E_j, E_k$ be defined similarly for $j$ and $k$. Then,*

$$\mathbb{P}\left((R_i, R_j, R_k) = (b_i, b_j, b_k), E_i, E_j, E_k\right) \geq 7 \cdot 10^{-8} \cdot \frac{n^{\frac{3}{2}}}{\ell^3}$$

PROOF. Let $S^i, S^j, S^k$ be the record of all coins used by $i, j, k$ respectively when in the bottom part of the of deck. Let $A_i$ be the event that $\text{Diff}_r(S_i) \leq 3\sqrt{T}$ for all $r \leq T$. Let $A_j$,$A_k$ be similarly defined for $j$ and $k$. By Hoeffding's inequality A.0.6 and the union bound we get

$$\mathbb{P}\left(A_i, A_j, A_k\right) \geq 1 - 12\exp\left(-\frac{9}{2}\right) > \frac{17}{20}.$$

Let $B^i, B^j, B^k$ be the record of coins used by $i, j, k$ respectively when in position $m$. On the events $A_i, A_j, A_k$ we can make precise statements about the number of times $i, j, k$ pass through position $m$ based on $B^i, B^j, B^k$. Following the proof of Proposition 3.3.13 we recall that, on $A_i$, each time card $i$ is in position $m$ it will take exactly $m$ more steps to return to position $m$ for a short return and $2n - m$ steps to return to position $m$ for a long return, without including faster/slower movement in the bottom part of the deck dictated by the Heads-Tails differential of $S^i$. Thus if the first $c$ coins in $B^i$ have $h$ heads and $c - h$ tails, we know that it will take

$$m - i + hm + (c - h)(2n - m + 1) + \Delta \text{ steps}$$

if $i \leq m$ and

$$m - i + 2(n - m) + hm + (c - h)(2n - m + 1) + \Delta \text{ steps}$$

if $i > m$ to reach position $m$ for the $(c + 1)$th time, where $\Delta$ is the Heads-Tails differential of $S^i$ at that time. Let $\tau$ be the time when $i$ reaches position $m$ for the $(c + 1)$th time. On event $A_i$

we have that $|\Delta| \leq 3\sqrt{T} \leq 3n$. This along with the fact that $0 \leq m - i$ in the first equation and $0 \geq m - i$ in the second gives us,

$$hm + (c-h)(2n-m+1) - 3n \leq \tau \leq hm + (c-h)(2n-m+1) + 5n. \tag{3.5}$$

Fix some desired value of $h$ which we will call $h_0$. Setting the upper bound in (3.5) equal to $T$ and solving for $c$, we get

$$c = \frac{T-5n}{2n-m+1} - h_0 \cdot \frac{2m-2n-1}{2n-m+1}.$$

Thus, if we let

$$c^* = \left\lfloor \frac{T-5n}{2n-m+1} - h_0 \cdot \frac{2m-2n-1}{2n-m+1} \right\rfloor$$

then if $h_0$ heads appear in the first $c^*$ coins, we know that on event $A_i$ at time $\tau \leq T$ card $i$ will have drawn exactly $c^*$ coins from $B^i$ and therefore exactly $h_0$ Heads and $c^* - h_0$ Tails. Furthermore, we can substitute $c^*$ in to the lower bound for $\tau$ in (3.5) and get

$$T - 8n - (2n - m + 1) \leq \tau.$$

Let $G_i$ be the event that card $i$ flips exclusively tails when in position $m$ between time $\tau$ and time $T$. Then $i$ will take at least $n$ steps to cycle back to $m$ each time and therefore can make at most 10 of these cycles in the fewer than $8n + 2n - m + 1$ steps between $\tau$ and $T$. So

$$\mathbb{P}(G_i) \geq \left(\frac{1}{2}\right)^{10} = \frac{1}{1024}.$$

For any valid $h$ let $X(h)$ be a binomial random variable with $\left\lfloor \frac{T-3n}{2n-m+1} - h \cdot \frac{2m-2n-1}{2n-m+1} \right\rfloor$ trials and $\frac{1}{2}$ chance of success. Then,

$$\mathbb{P}(R_i = h) \geq \frac{1}{1024} \cdot \mathbb{P}(X(h) = h).$$

If $h = \frac{T}{2n} + \delta$ then the number of trials $X(h)$ has is

$$\begin{aligned}
&\left\lfloor \frac{T-5n}{2n-m+1} - \left(\frac{T}{2n}+\delta\right)\cdot\frac{2m-2n-1}{2n-m+1}\right\rfloor \\
&= \left\lfloor \frac{T-5n-\frac{m}{n}T+T+\frac{1}{2n}T}{2n-m+1} - \delta\cdot\frac{2m-2n-1}{2n-m+1}\right\rfloor \\
&= \left\lfloor \frac{(2-\frac{m}{n}+\frac{1}{n})T}{2n-m+1} - \frac{\frac{1}{2n}T}{2n-m+1} - \frac{5n}{2n-m+1} - \delta\cdot\frac{2m-2n-1}{2n-m+1}\right\rfloor \\
&= \left\lfloor \frac{T}{n} - \frac{\frac{1}{2n}T}{2n-m+1} - \frac{5n}{2n-m+1} + \delta\cdot\frac{2n-2m+1}{2n-m+1}\right\rfloor.
\end{aligned}$$

Recall that $T \le n^2$ so,

$$\frac{1}{2n}T \le \frac{1}{2}n < n$$

for large enough $n$. This means that

$$-\frac{\frac{1}{2n}T}{2n-m+1} - \frac{5n}{2n-m+1} \ge -\frac{6n}{2n-m+1} \ge -6.$$

Also note that

$$\left|\frac{2n-2m+1}{2n-m+1}\right| \le 1 \text{ for all } m \le n.$$

Thus we know that the number of trials $X(h)$ has is within

$$\left(\frac{T}{n} - 6 - \delta,\ \frac{T}{n} + 6 + \delta\right).$$

As long as $|\delta| \le \sqrt{\frac{T}{n}}$, the event that $X(h) = h$ boils down to asking if a binomial random variable hits a specific value within three standard deviations. This gives us

$$\mathbb{P}\left(X(h) = h\right) \ge \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{9}{2}\right)\sqrt{\frac{n}{T}} \qquad \text{for } h \in \left(\frac{T}{2} - \sqrt{\frac{T}{n}},\ \frac{T}{2} + \sqrt{\frac{T}{n}}\right).$$

Since we required $b_i \in \left(\frac{T}{2n} - \frac{\ell}{\sqrt{n}}, \frac{T}{2n} + \frac{\ell}{\sqrt{n}}\right)$ $\ell$ and $\ell \le \sqrt{T}$ we know $b_i$ meets these parameters.

The entire argument for card $i$ applies to $j$ and $k$ as well due to symmetry. Note that the bounds on $R_i, R_j, R_k$ are determined by considering the order of the coins used by $i, j, k$ when they are in

position $m$ respectively. Thus, the events considered to achieve these bounds are independent after conditioning on $A_i, A_j, A_k$. This means that

$$\begin{aligned}\mathbb{P}\big(A_i, A_j, A_k, (R_i, R_j, R_k) = (b_i, b_j, b_k)\big) &\geq \frac{17}{20}\left(\frac{1}{1024} \cdot \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{9}{2}\right)\sqrt{\frac{n}{T}}\right)^3 \\ &\geq 6 \cdot 10^{-17} \cdot \left(\frac{n}{T}\right)^{\frac{3}{2}}.\end{aligned}$$

Recall that $A_i, A_j$ and $A_k$ are the events that $|\mathrm{Diff}_r(S^i)|$, $|\mathrm{Diff}_r(S^j)|$, and $|\mathrm{Diff}_r(S^k)|$ respectively stay less than or equal to $3\sqrt{T}$. Since $T \leq \ell^2 + 4n$ and $\ell \geq 2\sqrt{n}$ we see that $3\sqrt{T} \leq \frac{9}{2}\ell$. On events $A_i, A_j, A_k$ and $(R_i, R_j, R_k) = (b_i, b_j, b_k)$ we have by Proposition 3.3.6 that,

$$\begin{aligned}\|p(i_t) - p(i) - t + mb_i\| &\leq |T_S - H_S| + T_B \\ &\leq \frac{9}{2}\ell + (\ell + 5) < 6\ell\end{aligned}$$

where we know $T_B \leq \ell + 5$ because at most $\frac{T}{n} + 1 \leq \frac{\ell^2+4n}{n} + 1$ Tails may be flipped by $T_B$ when in position $m$, and $\ell \leq n$. A similar argument shows the equivalent statement for $j$ and $k$ also follow for $A_i, A_j, A_k$ and $(R_i, R_j, R_k) = (b_i, b_j, b_k)$. This tells us that

$$\mathbb{P}\left((R_i, R_j, R_k) = (b_i, b_j, b_k), E_i, E_j, E_k\right) \geq 6 \cdot 10^{-17} \cdot \left(\frac{n}{T}\right)^{\frac{3}{2}}.$$

Since $T \geq \ell^2$ we get

$$\mathbb{P}\left((R_i, R_j, R_k) = (b_i, b_j, b_k), E_i, E_j, E_k\right) \geq 6 \cdot 10^{-17} \cdot \frac{n^{\frac{3}{2}}}{\ell^3}.$$

□

The following lemmas will collectively make the following argument: For cards $i, j, k$ fix some desired number of position $m$ Heads flips. Let $e_i, e_j, e_k$ be positions where you would expect $i, j, k$ to go with $T$ steps and your desired number of Heads flips while in position $m$. Then if we choose positions $f_i, f_j, f_k$ at uniformly at random nearby $e_i, e_j, e_k$ it is likely enough $i, j, k$ will go to $f_i, f_j, f_k$ in $T$ steps.

We then show that the same statement applies to the inverse overlapping cycles shuffle, and this gives us the following variation: Choose $a_i, a_j, a_k$ nearby $i, j, k$. Then after $T$ steps it is likely

enough $a_i, a_j, a_k$ will go to $e_i, e_j, e_k$. It will also be true that if we choose, for example, $a_i, a_j, a_k$ nearby $i+1, j+1, k+1$ then it is likely enough $a_i, a_j, a_k$ go to $e_i+1, e_j+1, e_k+1$. Using this logic, we show that if $a_i, a_j, a_k$ are chosen from an extra wide range of values around $i, j, k$ then $a_i, a_j, a_k$ have a distribution which is approximately uniform across values near to $e_i, e_j, e_k$.

LEMMA 3.5.7. *Fix positions $i, j, k$ and fix $\ell$ such that $2\sqrt{n} < \ell < \frac{n}{3}$. Set $T \in [\ell^2, \ell^2 + 4n]$ such that $T + \lfloor \frac{T}{2n} \rfloor m \equiv 0 \mod 2n - m + 1$. Fix any $\beta_i, \beta_j, \beta_k \in \mathbb{Z} \cap \left(-\frac{\ell}{\sqrt{n}}, \frac{\ell}{\sqrt{n}}\right)$. Let $b_i = \beta_i + \lfloor \frac{T}{2n} \rfloor$ and let $b_j, b_k$ be defined similarly. Now let $\Delta_i, \Delta_j, \Delta_k$ be iid uniformly chosen from $(-6\ell, 6\ell) \cap \mathbb{Z}$. Let $R_i, R_j, R_k$ be the number of Heads flipped by $i, j, k$ respectively when in position $m$ in the first $T$ steps. Let*

$$\omega_i \equiv p(i) + \beta_i m + \Delta_i \ (mod\ 2n - m + 1),$$
$$\omega_j \equiv p(j) + \beta_j m + \Delta_j \ (mod\ 2n - m + 1),$$
$$\omega_k \equiv p(k) + \beta_k m + \Delta_k \ (mod\ 2n - m + 1).$$

*Then,*

$$\mathbb{P}\left((R_i, R_j, R_k) = (b_i, b_j, b_k), (p(i_T), p(j_T), p(k_T)) = (\omega_i, \omega_j, \omega_j)\right) \geq 3 \cdot 10^{-20} \cdot \frac{n^{\frac{3}{2}}}{\ell^6}.$$

PROOF. By Proposition 3.5.6 we know that with probability at least

$$6 \cdot 10^{-17} \cdot \frac{n^{\frac{3}{2}}}{\ell^3}$$

we have

$$\begin{aligned} p(i_T) &\equiv p(i) + T + b_i m + \delta_i \\ &\equiv p(i) + T + \left\lfloor \frac{T}{2n} \right\rfloor m + \beta_i m + \delta_i \\ &\equiv p(i) + \beta_i m + \delta_i \end{aligned}$$

for some $\delta_i \in (-6\ell, 6\ell)$ and equivalent statements for $j, k$. Since there are $12\ell$ values in this range that $\Delta_i$ might take, the probability that $\Delta_i$ takes the "correct" one is $\frac{1}{12\ell}$. The same argument

applies for $j$ and $k$. So,

$$\mathbb{P}\left((R_i, R_j, R_k) = (b_i, b_j, b_k), (p(i_T), p(j_T), p(k_T)) = (\omega_i, \omega_j, \omega_j)\right) > 6 \cdot 10^{-17} \cdot \frac{n^{\frac{3}{2}}}{\ell^3} \cdot \frac{1}{12^3} \cdot \frac{1}{\ell^3}.$$

□

COROLLARY 3.5.8. *Fix $i, j, k, \ell, T, \beta_i, \beta_j, \beta_k, b_i, b_j, b_k$ as in Lemma 3.5.7. As before let $R_i, R_j, R_k$ be the number of Heads flipped by $i, j, k$ respectively when in position $m$ in $T$ steps. Now let $Z_i, Z_j, Z_k$ be uniformly chosen from all positions such that*

$$|p(Z_i) - p(i) - \beta_i m|_M < 6\ell,$$

$$|p(Z_j) - p(j) - \beta_j m|_M < 6\ell,$$

$$|p(Z_k) - p(k) - \beta_k m|_M < 6\ell.$$

*Then,*

$$\mathbb{P}\left((R_i, R_j, R_k) = (b_i, b_j, b_k), (i_T, j_T, k_T) = (Z_i, Z_j, Z_j)\right) \geq 3 \cdot 10^{-20} \cdot \frac{n^{\frac{3}{2}}}{\ell^6}.$$

PROOF. In the bound described by Lemma 3.5.7 we neglect to use the fact that only "valid" choices of $\delta_i, \delta_j, \delta_k$, which allow for $\omega_i, \omega_j, \omega_k$ to be in the image of $p$ have a nonzero chance of being the locations of $i, j, k$ after $T$ steps. For example, depending on the parameters, we might randomly choose $\omega_i = m + 1$. But $m + 1$ is not associated with any position in the deck, as $p(m) = m$ and $p(m + 1) = m + 2$. By conditioning on the probability 1 event that $i_T, j_T, k_T$ are associated with true positions (in the image of $p$) have this Corollary. □

We can reword the statement of Corollary 3.5.8 to immediately get the following Corollary:

COROLLARY 3.5.9. *Fix $i, j, k, \ell, T, \beta_i, \beta_j, \beta_k, b_i, b_j, b_k$ as in Lemma 3.5.7. Let $Z_i, Z_j, Z_k$ be chosen uniformly from all positions such that*

$$|p(Z_i) - p(i) - \beta_i m|_M < 6\ell,$$

$$|p(Z_j) - p(j) - \beta_j m|_M < 6\ell,$$

$$|p(Z_k) - p(k) - \beta_k m|_M < 6\ell.$$

*Let $c_i, c_j, c_k$ be the cards which, after $T$ steps end up in positions $Z_i, Z_j, Z_k$. Let $R_i, R_j, R_k$ be the number of Heads flipped in these $T$ steps while $c_i, c_j, c_k$ are in position $m$. Then,*

$$\mathbb{P}\left((R_i, R_j, R_k) = (b_i, b_j, b_k), (c_i, c_j, c_k) = (i, j, k)\right) \geq 3 \cdot 10^{-20} \cdot \frac{n^{\frac{3}{2}}}{\ell^6}.$$

This leads to the following lemma:

LEMMA 3.5.10. *Fix $i, j, k, \ell, T, \beta_i, \beta_j, \beta_k, b_i, b_j, b_k$ as in Lemma 3.5.7. Fix any cards $\alpha_i, \alpha_j, \alpha_k$ such that*

$$|p(\alpha_i) - p(i) + \beta_i m|_M < 2\ell,$$

$$|p(\alpha_j) - p(j) + \beta_j m|_M < 2\ell,$$

$$|p(\alpha_k) - p(k) + \beta_k m|_M < 2\ell.$$

*Now let $Z_i', Z_j', Z_k'$ be uniformly chosen from all positions such that*

$$|p(Z_i') - p(i)|_M < 8\ell,$$

$$|p(Z_j') - p(j)|_M < 8\ell,$$

$$|p(Z_k') - p(k)|_M < 8\ell.$$

*Let $c'_i, c'_j, c'_k$ be the cards which, after $T$ steps end up in positions $Z'_i, Z'_j, Z'_k$. Let $R_i, R_j, R_k$ be the number of heads which $c'_i, c'_j, c'_k$ flip in these $T$ steps while in position $m$. Then,*

$$\mathbb{P}\left((R_i, R_j, R_k) = (b_i, b_j, b_k), (c'_i, c'_j, c'_k) = (\alpha_i, \alpha_j, \alpha_k)\right) \geq 10^{-21} \cdot \frac{n^{\frac{3}{2}}}{\ell^6}.$$

Proof. Let $\Lambda_i$ be the set of all $\zeta$ such that that $|p(\zeta) - p(i)|_M < 8\ell$. Let $\Gamma_i$ be the set of all $\omega$ such that $|p(\omega) - p(\alpha_i) - \beta_i m|_M < 6\ell$. Note that $\Gamma_i$ is a subset of $\Lambda_i$ because $|p(\alpha_i) - p(i) + \beta_i m|_M < 2\ell$. Since $|p(z+1) - p(z)|_M \in \{1, 2\}$ for all $z$ we see that $|\Gamma_i| \geq 3\ell$ and $|\Lambda_i| \leq 8\ell$. Since $Z'_i$ is chosen uniformly from $\Lambda_i$ we get

$$\mathbb{P}\left(Z'_i \in \Gamma_i\right) \geq \frac{3}{8}.$$

Define $\Gamma_j, \Gamma_k$ similarly for $\alpha_j, \alpha_k$. Since $Z'_i, Z'_j, Z'_k$ are chosen independently we have

$$\mathbb{P}\left(Z'_i \in \Gamma_i, Z'_j \in \Gamma_j, Z'_k \in \Gamma_k\right) \geq \left(\frac{3}{8}\right)^3.$$

Conditioning on $Z'_i \in \Gamma_i, Z'_j \in \Gamma_j, Z'_k \in \Gamma_k$ we can apply Corollary 3.5.9 to get

$$\mathbb{P}\left((R_i, R_j, R_k) = (b_i, b_j, b_k), (c_i, c_j, c_k) = (\alpha_i, \alpha_j, \alpha_k) \mid Z'_i \in \Gamma_i, Z'_j \in \Gamma_j, Z'_k \in \Gamma_k\right) \geq 3 \cdot 10^{-20} \cdot \frac{n^{\frac{3}{2}}}{\ell^6}.$$

So,

$$\mathbb{P}\left((R_i, R_j, R_k) = (b_i, b_j, b_k), (c_i, c_j, c_k) = (\alpha_i, \alpha_j, \alpha_k)\right) \geq \left(\frac{3}{8}\right)^3 \cdot 3 \cdot 10^{-20} \cdot \frac{n^{\frac{3}{2}}}{\ell^6}.$$

□

Recall that our goal with these lemmas is to show that certain cards end up in an approximately uniform distribution. To this point we have shown that the triplets of cards we are interested in end up in "reasonable" positions with probability greater than or equal to

$$(\text{constant}) \cdot \left(\frac{\sqrt{n}}{\ell} \cdot \frac{1}{\ell}\right)^3. \tag{3.6}$$

This is good because it matches the $2\frac{\ell}{\sqrt{n}}$ choices for $\beta_i$ and the $2\ell$ to $4\ell$ outcomes for $\alpha_i$. However there is one problem. Consider the case of $m = \frac{n}{2}$. In this case $4m \equiv 1 \mod 2n - m + 1$. This means that, for example, if we for example choose $\beta_i = 3$ the positions nearby $p(i) - \beta_i m$ will be

almost the same set of positions as if we picked $\beta_i = 7$ or $\beta_i = -1$ etc. In the $m = \frac{n}{2}$ case we only have a constant times $\ell^3$ possible values for $\alpha_i, \alpha_j, \alpha_k$ across all choices of $\beta_i, \beta_j, \beta_k$. This does not match our probability in line (3.6).

To remedy this we introduce a constant $\gamma$ which depends on $n, m$ and $\ell$. It will give us the ability to increase the probability beyond that of (3.6) in cases like $m = \frac{n}{2}$.

DEFINITION 3.5.11. *Let $\gamma = \gamma(\ell, n, m)$ be defined as follows:*

$$\gamma = \left| \left\{ \kappa \in \mathbb{N} \ : \ |\kappa m|_M < \ell, \ \kappa < \frac{\ell}{\sqrt{n}} \right\} \right|.$$

Note that $\kappa = 0$ is always an element of the set and so $\gamma \geq 1$. It will be important to be able to use $\gamma$ to count how many cards there are with norms less than or equal to $\ell$, so we prove the following Lemma.

LEMMA 3.5.12. *Let $N^\ell$ be defined by*

$$N^\ell = \left\{ \textit{positions } \omega \textit{ such that } \omega \equiv a + bm \mod 2n - m + 1 \textit{ where } |a| < \ell, \ |b| \leq \frac{\ell}{\sqrt{n}} \right\}.$$

*Then,*

$$|N^\ell| \geq \frac{2\ell^2}{\gamma\sqrt{n}}$$

PROOF. First consider some integer $z \in (1, \ldots, 2n - m + 1)$ with $\|z\| \leq \ell$, ignoring for a moment if $z$ actually represents a position in the deck. Then we know $z = a + b\sqrt{n}$ with $|a| < \ell$ and $|b|\sqrt{n} < \ell$. Let $\mathcal{K}^\ell$ be the set of all such $z$. First consider the extreme case where there exist $\gamma$ natural numbers $\kappa$ where $\kappa m \equiv 0 \mod 2n - m + 1$. Then

$$|\mathcal{K}^\ell| \geq 2\ell \cdot \frac{2\ell}{\gamma\sqrt{n}}$$

because there are $2\ell$ choices for $a$ and $2\frac{\ell}{\sqrt{n}}$ choices for $b$ but each of the choices for $b$ are over-counted by at most $\gamma$. Since this is the extreme case the bound holds in general.

Recall that $N^\ell$ is the subset of $\mathcal{K}^\ell$ that only includes $\omega \in (1, \ldots, 2n - m + 1)$ such that $p(a) = \omega$

for some $a \in (1, \ldots, n)$. Since all $\omega \leq m$ have this property and every other $\omega > m$ has it we get

$$|N^\ell| \geq \frac{1}{2}|\mathcal{K}^\ell|$$

This completes the proof. □

Now that we have defined $\gamma$ we will incorporate it into our probability bound.

LEMMA 3.5.13. *Fix positions $i, j, k$ and fix $\ell$ such that $2\sqrt{n} < \ell < \frac{n}{3}$. Set $T \in [\ell^2, \ell^2 + 4n]$ such that $T + \lfloor \frac{T}{2n} \rfloor m \equiv 0 \mod 2n - m + 1$. Fix any $\beta_i', \beta_j', \beta_k' \in \mathbb{Z} \cap \left(-\frac{\ell}{\sqrt{n}}, \frac{\ell}{\sqrt{n}}\right)$. Fix any cards $\alpha_i', \alpha_j', \alpha_k'$ such that*

$$|p(\alpha_i') - p(i) + \beta_i' m|_M < \ell,$$
$$|p(\alpha_j') - p(j) + \beta_j' m|_M < \ell,$$
$$|p(\alpha_k') - p(k) + \beta_k' m|_M < \ell.$$

*Now let $\mathcal{Z}_i, \mathcal{Z}_j, \mathcal{Z}_k$ be uniformly chosen from all positions such that*

$$|p(\mathcal{Z}_i) - p(i)|_M < 9\ell,$$
$$|p(\mathcal{Z}_j) - p(j)|_M < 9\ell,$$
$$|p(\mathcal{Z}_k) - p(k)|_M < 9\ell.$$

*Let $\mathcal{C}_i, \mathcal{C}_j, \mathcal{C}_k$ be the cards which, after $T$ steps end up in positions $\mathcal{Z}_i, \mathcal{Z}_j, \mathcal{Z}_k$. Then,*

$$\mathbb{P}\left((\mathcal{C}_i, \mathcal{C}_j, \mathcal{C}_k) = (\alpha_i', \alpha_j', \alpha_k')\right) \geq 10^{-23} \cdot \frac{\gamma^3 n^{\frac{3}{2}}}{\ell^6}.$$

PROOF. Fix any $\kappa_i \in \mathbb{N}$ such that $|\kappa m|_M < \ell$ and $\kappa_i < \frac{\ell}{\sqrt{n}}$. If $\beta_i' \geq 0$ then let $\beta_i = \beta_i' + \kappa_i$ and if $\beta_i' < 0$ then let $\beta_i = \beta_i' - \kappa_i$. Note that this insures $\beta_i \in \left(-\frac{\ell}{\sqrt{n}}, \frac{\ell}{\sqrt{n}}\right)$. Now because $|\kappa_i m|_M < \ell$ there exists a position $\eta_i$ (where $p(\eta_i)$ approximately equals $p(i) + \beta_i$) such that

$$|p(\eta_i) - p(i)|_M < \ell,$$

which implies

$$|p(\alpha_i') - p(\eta_i) + \beta_i' m|_M < 2\ell.$$

Let $\Lambda_i$ be the set of all $\zeta$ such that $|p(\zeta) - p(i)|_M < 9\ell$. Let $\Gamma_i$ be the set of all $\omega$ such that $|p(\omega) - p(\eta_i)|_M < 8\ell$. Note that $\Gamma_i$ is a subset of $\Lambda_i$ because $|p(\eta_i) - p(i)|_M < \ell$. Since $\mathcal{Z}_i$ is chosen uniformly from $\Lambda$ we can show similarly to the proof of Lemma 3.5.10 that

$$\mathbb{P}\left(\mathcal{Z}_i \in \Gamma_i\right) \geq \frac{4}{9}.$$

Now fix some $\kappa_j, \kappa_k$ from the same subset of $\mathbb{N}$ as $\kappa_i$ was chosen from. Use these to similarly define $\eta_j, \eta_k, \Gamma_j, \Gamma_k$. Then we can use Lemma 3.5.10 to show that the probability of

$$(R_i, R_j, R_k) = \left(\beta_i' + \left\lfloor \frac{T}{2n} \right\rfloor + \kappa_i, \beta_j' + \left\lfloor \frac{T}{2n} \right\rfloor + \kappa_j, \beta_k' + \left\lfloor \frac{T}{2n} \right\rfloor + \kappa_k\right) \text{ and } (\mathcal{C}_i, \mathcal{C}_j, \mathcal{C}_k) = (\alpha_i', \alpha_j', \alpha_k')$$

conditioned on $\mathcal{Z}_i \in \Gamma_i, \mathcal{Z}_j \in \Gamma_j, \mathcal{Z}_k \in \Gamma_k$ is at least

$$10^{-21} \cdot \frac{n^{\frac{3}{2}}}{\ell^6}.$$

So

$$\mathbb{P}\left((R_i, R_j, R_k) = \left(\beta_i' + \left\lfloor \frac{T}{2n} \right\rfloor + \kappa_i, \beta_j' + \left\lfloor \frac{T}{2n} \right\rfloor + \kappa_j, \beta_k' + \left\lfloor \frac{T}{2n} \right\rfloor + \kappa_k\right),\ (\mathcal{C}_i, \mathcal{C}_j, \mathcal{C}_k) = (\alpha_i', \alpha_j', \alpha_k')\right)$$

is at least

$$\left(\frac{4}{9}\right)^3 \cdot 10^{-21} \cdot \frac{n^{\frac{3}{2}}}{\ell^6}.$$

Summing over all $\gamma^3$ ways to choose $\kappa_i, \kappa_j, \kappa_k$ we get

$$\mathbb{P}\left((\mathcal{C}_i, \mathcal{C}_j, \mathcal{C}_k) = (\alpha_i', \alpha_j', \alpha_k')\right) \geq \gamma^3 \cdot \left(\frac{4}{9}\right)^3 \cdot 10^{-21} \cdot \frac{n^{\frac{3}{2}}}{\ell^6}.$$

□

Be rewording Lemma 3.5.13 the following Corollary is immediate:

COROLLARY 3.5.14. *Fix positions $i, j, k$ and fix $\ell$ such that $2\sqrt{n} < \ell < \frac{n}{3}$. Set $T \in [\ell^2, \ell^2 + 4n]$ such that $T + \lfloor \frac{T}{2n} \rfloor m \equiv 0 \mod 2n - m + 1$. Fix any $\beta_i', \beta_j', \beta_k' \in \mathbb{Z} \cap \left(-\frac{\ell}{\sqrt{n}}, \frac{\ell}{\sqrt{n}}\right)$. Fix any positions*

*$f_i, f_j, f_k$ such that*

$$|p(f_i) - p(i) + \beta_i' m|_M < \ell,$$

$$|p(f_j) - p(j) + \beta_j' m|_M < \ell,$$

$$|p(f_k) - p(k) + \beta_k' m|_M < \ell.$$

*Now let $i', j', k'$ be uniformly chosen from all cards such that*

$$|p(i') - p(i)|_M < 9\ell,$$

$$|p(j') - p(j)|_M < 9\ell,$$

$$|p(k') - p(k)|_M < 9\ell.$$

*Let $i'_{(-T)}, j'_{(-T)}, k'_{(-T)}$ be the locations of $i', j', k'$ respectively after $T$ steps of the inverse overlapping cycles shuffle. Then,*

$$\mathbb{P}\left((i'_{(-T)}, j'_{(-T)}, k'_{(-T)}) = (f_i, f_j, f_k)\right) \geq 10^{-23} \cdot \frac{\gamma^3 n^{\frac{3}{2}}}{\ell^6}.$$

We now exploit the similarity between the inverse overlapping cycles shuffle and the normal forwards overlapping cycles shuffle to get the equivalent statement for $T$ positive steps.

PROPOSITION 3.5.15. *Fix positions $i, j, k$ and fix $\ell$ such that $2\sqrt{n} < \ell < \frac{n}{3}$. Set $T \in [\ell^2, \ell^2 + 4n]$ such that $T + \lfloor \frac{T}{2n} \rfloor m \equiv 0 \mod 2n - m + 1$. Fix any $\beta_i, \beta_j, \beta_k \in \mathbb{Z} \cap \left(-\frac{\ell}{\sqrt{n}}, \frac{\ell}{\sqrt{n}}\right)$. Fix any positions $f_i, f_j, f_k$ such that*

$$|p(f_i) - p(i) + \beta_i m|_M < \ell,$$

$$|p(f_j) - p(j) + \beta_j m|_M < \ell,$$

$$|p(f_k) - p(k) + \beta_k m|_M < \ell.$$

*Now let $i', j', k'$ be uniformly chosen from all cards such that*

$$|p(i') - p(i)|_M < 9\ell,$$
$$|p(j') - p(j)|_M < 9\ell,$$
$$|p(k') - p(k)|_M < 9\ell.$$

*Then,*

$$\mathbb{P}\left((i'_T, j'_T, k'_T) = (f_i, f_j, f_k)\right) \geq 10^{-23} \cdot \frac{\gamma^3 n^{\frac{3}{2}}}{\ell^6}.$$

PROOF. This follows from Corollary 3.5.14 and Corollary 3.3.12. Corollary 3.5.14 describes the distribution of randomly chosen $i', j', k'$ after $T$ steps of the inverse overlapping cycles shuffle. In particular it says that if $\beta_i, \beta_j, \beta_k$ are appropriately fixed and $i', j', k'$ are uniformly sampled so that

$$|p(i') - p(i)|_M < 9\ell$$
$$|p(j') - p(j)|_M < 9\ell$$
$$|p(k') - p(k)|_M < 9\ell$$

then for any positions $f_i, f_j, f_k$ such that

$$|p(f_i) - p(i)|_M \mod 2n - m + 1 \in (-\beta_i - \ell,\ -\beta_k m + \ell),$$
$$|p(f_j) - p(j)|_M \mod 2n - m + 1 \in (-\beta_j - \ell,\ -\beta_k m + \ell),$$
$$|p(f_k) - p(k)|_M \mod 2n - m + 1 \in (-\beta_k - \ell,\ -\beta_k m + \ell),$$

we have that $i', j', k'$ are sufficiently likely to travel to $f_i, f_j, f_k$ in $T$ steps of the inverse overlapping cycles shuffle. By Theorem 3.1.1 we know that the inverse overlapping cycles shuffle is the same as the forward overlapping cycles shuffle after a reordering of the cards, and by Corollary 3.3.12 we know that $|p(\cdot) - p(\cdot)|_M$ is fixed under this reordering. Thus the equivalent result to Corollary 3.5.14 applies to the forwards overlapping cycles shuffle. □

We have shown that if we choose random cards nearby $i, j, k$ then the distribution of these random cards will be close to uniform over nearby positions. But we really care about the distribution of $i, j, k$ themselves, not some randomly chosen neighbors. We now use a coupling argument to show that the distribution of cards $i, j, k$ themselves will also be approximately uniform. The basic idea

is that we choose random neighbors $i', j', k'$ and then show that under a certain coupling that $i, j, k$ will couple with $i', j', k'$ with probability bounded away from 0. The argument will require $i, j, k$ to be spread out from each other, which is why we require Stage 1 and Stage 3.

To start with, we describe the coupling.

Fix any cards $i, j, k$ such that $\|i - j\|, \|i - k\|, \|j - k\| > 199\ell$. Let $T = \ell^2$. Let $i', j', k'$ be chosen uniformly from cards such that $|p(i) - p(i')|_M, |p(j) - p(j')|_M, |p(k) - p(k')|_M < 9\ell$. We now run two overlapping cycles shuffles $(\pi_t)$ and $(\pi'_t)$ and we track $i, j, k$ in $(\pi_t)$ and $i', j', k'$ in $(\pi'_t)$. We couple the two shuffle as follows:

- Generate coin sequences $B^i, B^j, B^k$. Have $(\pi'_t)$ draw from $B^i$ or $B^j$ or $B^k$ whenever $i'$ or $j'$ or $k'$ is in position $m$. Similarly have $(\pi_t)$ draw from $B^i$ or $B^j$ or $B^k$ whenever $i$ or $j$ or $k$ is in position $m$ with the following exceptions:
  - If $i' \leq m < i$ then have $i$ skip $B^i_1$ and draw from $B^i_2$ on its first visit to $m$ and $B^i_3$ on its second visit etc.
  - If $i \leq m < i'$ then have $i'$ skip $B^i_1$ and draw from $B^i_2$ on its first visit to $m$ and $B^i_3$ on its second visit etc.
  - Have the equivalent exceptions for $j$ and $k$.

  This will ensure that $i, j, k$ follow the same choice of big coins as their counterparts. For example, imagine that $i_0 = m - 3$ and $i'_0 = m - 1$. Then after one step we have $i_1 = m_2$ and $i'_1 = m$. Now if $B^i_1$ is a Heads, then $\pi'$ will flip Heads on its second step. So $i_2 = m - 1$ and $i'_2 = 1$. One more step and we get $i'_3 = 2$ and $i_3 = m$. Note that this is the first time $i$ reaches position $m$, so now $\pi$ must use $B^i_1$ which is Heads. So $i'_4 = 3$ and $i_4 = 1$. In this way, $i$ and $i'$ will always follow the same trajectory of "big" coins and $i$ will "follow behind" $i'$. The reason we have the aforementioned exceptions is because if, for example, $j_0 = m - 1$ and $j'_0 = m + 1$, then we want to wait to synchronize the draws from $B^j$ until after $j$ and $j'$ are in the same part of the deck. That way, as long as $B^j_1$ is a Tails (which happens with probability $\frac{1}{2}$) we will get that $j$ follows behind $j'$ as they will both start drawing from $B^j_2$ once they cycle back to position $m$.

- To determine the movement of $i', j', k'$ when none of these cards are in position $m$, generate sequences of coins $S^i, S^j, S^k$. Whenever $i'$ is in the bottom part of the deck (and neither $j'$ nor $k'$ is in position $m$) use a coin from $S^i$ for $(\pi'_t)$. Whenever $i'$ is in the top part of the deck and $j'$ is in the bottom part of the deck (and $k'$ is not in position $m$) use a coin from $S^j$. Whenever $i'$ and $j'$ are in the top part of the deck and $k'$ is in the bottom part of the deck, use a coin from $S^k$. You can think of $S^j$ having "priority" over $S^k$ and $S^i$ having priority over both $S^j$ and $S^k$. Think of $B^i, B^j, B^k$ as all having priority over $S^i, S^j, S^k$.
- The movement of $i, j, k$ when none of these cards is in position $m$ is broken into four phases. Generate sequences of coins $X^i, X^j, X^k, X^{ij}, X^{ik}, X^{jk}, X^{ijk}, Y^j, Y^k, Y^{jk}, Z^k$.
  - At the start of the process, we say we are in "Phase 1". In this phase, if none of $i, j, k$ are in position $m$, use the next coin from $X^A$ to determine the movement of $i, j, k$ where $A$ is the set of exactly which of $i, j, k$ are in the bottom part of the deck. Let $\tau_1$ be random stopping time which is the minimum $t$ such that $i_t = i'_t + \Delta(i, t)$ where $\Delta(i, t)$ is a small random variable (very likely in $\{-1, 0, 1\}$) which we will define later. You can imagine $\tau_1$ is more or less when $i_t = i'_t$, and the reason that we add $\Delta(i, t)$ is a technicality which we will explain later. At time $\tau_1$ we move to Phase 2, and we change the rules for how $i, j, k$ move in order to couple $i$ to $i'$.
  - At time $\tau_1$, suppose $r_i$ is the number of coins $i'$ has drawn from $S^i$. Let $\kappa^i$ to be such that $\kappa^i_s = S^i_{r_i+s}$. Now whenever $i$ is in the bottom part of the deck, draw the next coin from $\kappa^i$. In other words, have $i$ start following the same sequence of coins that $i'$ uses. Since $i$ and $i'$ will both draw from the same sequence of coins when they are in the bottom part of the deck, we know $i$ and $i'$ will stay coupled together, except for a small technicality which we will explain later. When $i$ is in the top part of the deck, use the next coin from $Y^j$ or $Y^k$ or $Y^{jk}$ depending on if $j$ is in the bottom of the deck or $k$ is or both. Let $\tau_2$ be the random stopping time which is the minimum $t \geq \tau_1$ such that $j_t = j'_t + \Delta(j, t)$ where $\Delta(j, t)$ is small random variable we will define later. At time $\tau_2$ we move to Phase 3, and change how $i, j, k$ move in order to couple $(i, j)$ to $(i', j')$.
  - At time $\tau_2$, suppose $r_j$ is the number of coins $j'$ has drawn from $S^j$. Define the sequence $\kappa^j$ to be such that $\kappa^j_s = S^j_{r_j+s}$. Now whenever $i$ is in the top part of the

deck and $j$ is in the bottom part of the deck, draw the next coin from $\kappa^j$. Continue to draw from $\kappa_i$ whenever $i$ is in the bottom part of the deck, regardless of if $j$ is in the bottom part of the deck or not. In other words, give $\kappa^i$ priority over $\kappa^j$ so $j$ follows the same sequence of coins that $j'$ uses. Now we know (other than a small technicality which will be explained later) that both $i, i'$ and $j, j'$ will stay coupled together because the joint movement of $(i, j)$ and $(i', j')$ follow the same rules. If $k$ is alone in the bottom part of the deck, draw from $Z^k$. Let $\tau_3$ be the random stopping time which is the minimum $t \geq \tau_2$ such that $k_t = k'_t + \Delta(k, t)$ where once again $\Delta(k, t)$ will be defined later. At time $\tau_3$ we move the Phase 4, where $(i, j, k)$ is permanently coupled with $(i', j', k')$.

- At time $\tau_3$, suppose $r_k$ is the number of coins $k'$ has drawn from $S^k$. Define the sequence $\kappa^k$ to be such that $\kappa_s = S^k_{r_k+s}$. Now whenever $i$ and $j$ are in the top part of the deck and $k$ is in the bottom part of the deck, draw the next coin from $\kappa^k$. Now $(i, j, k)$ obey the same rules as $(i', j', k')$, so they will stay coupled forever.

The technicality that could cause $i$ to "decouple" from $i'$ after time $\tau_1$ is the fact that $i$ is obligated to use coins from $B^j$ or $B^k$ whenever $j$ or $k$ are in position $m$, whereas $i'$ uses coins from $B^j$ or $B^k$ whenever $j'$ or $k'$ are in position $m$. Since $j$ will be in position $m$ at different times than $j'$ during Phase 2, there will probably be times during Phase 2 where $i_t \neq i'_t$. However, as long as we make sure that $j_t$ and $j'_t$ stay closer to each other than to $i_t$ and $i'_t$ this will not be an issue. To see why, consider the situation in Phase 2 where

$$\begin{pmatrix} i_t \\ j_t \\ i'_t \\ j'_t \end{pmatrix} = \begin{pmatrix} m + 100 \\ m \\ m + 100 \\ m - 3 \end{pmatrix}$$

Assume that $k_t$ and $k'_t$ are far away from position $m$. Also assume that the upcoming coin in $B^j$ is a Tails and $i_t, i'_t$ are at coin number $r$ in sequence $S^i$. Finally assume

$$(S^i_r, S^i_{r+1}, S^i_{r+2}, S^i_{r+3}) = (H, T, H, T).$$

(1) In the first step $\pi$ draws from $B^j$ to get a Tails, and $\pi'$ draws $S^i_r$ from $S^i$ to get a Heads. So

$$\begin{pmatrix} i_{t+1} \\ j_{t+1} \\ i'_{t+1} \\ j'_{t+1} \end{pmatrix} = \begin{pmatrix} m+101 \\ m+1 \\ m+100 \\ m-2 \end{pmatrix}$$

Now $i$ and $i'$ have decoupled! This looks bad, but lets see what happens over the course of the next few moves.

(2) Now $\pi$ draws from $S^i$ but it is a step behind where $\pi'$ is in $S^i$. So $\pi$ draws $S^i_r$ from $S^i$ to get a Heads. On the other hand $\pi'$ draws $S^i_{r+1}$ from $S^i$ to get a Tails. So now

$$\begin{pmatrix} i_{t+2} \\ j_{t+2} \\ i'_{t+2} \\ j'_{t+2} \end{pmatrix} = \begin{pmatrix} m+101 \\ m+1 \\ m+101 \\ m-1 \end{pmatrix}$$

It may look like everything is fixed now, but remember $\pi$ and $\pi'$ are at different points in the sequence $S^i$. This is going to cause trouble the next step.

(3) Now $\pi$ draws $S^i_{r+1}$ from $S^i$ to get a Tails and $\pi'$ draws $S^i_{r+2}$ from $S^i$ to get a Heads. So

$$\begin{pmatrix} i_{t+3} \\ j_{t+3} \\ i'_{t+3} \\ j'_{t+3} \end{pmatrix} = \begin{pmatrix} m+102 \\ m+2 \\ m+101 \\ m \end{pmatrix}$$

Now that $j'$ is in position $m$ everything will be fixed in the next step.

(4) We know $\pi$ will draw $S^i_{r+2}$ from $S^i$ to get a Heads. But now $\pi'$ will draw a Tails from $B^j$, the "same" Tails that $\pi$ drew three steps earlier. So now

$$\begin{pmatrix} i_{t+4} \\ j_{t+4} \\ i'_{t+4} \\ j'_{t+4} \end{pmatrix} = \begin{pmatrix} m+102 \\ m+2 \\ m+102 \\ m+1 \end{pmatrix}$$

We see that $i$ and $i'$ have re-synchronized. Not only are they in the same position again, but $\pi$ and $\pi'$ are in the same place in $S^i$. This means that they will stick together until another discrepancy from $B^j$ or $B^k$ causes them to split. For example, lets look at one more step.

(5) Now $\pi$ and $\pi'$ BOTH draw $S^i_{r+3}$ from $S^i$ to get a Heads. So,

$$\begin{pmatrix} i_{t+5} \\ j_{t+5} \\ i'_{t+5} \\ j'_{t+5} \end{pmatrix} = \begin{pmatrix} m+103 \\ m+3 \\ m+103 \\ m+2 \end{pmatrix}$$

In short, in any situation where $j$ hits position $m$ before $j'$ hits position $m$ (or vice versa), the decoupling of $i$ from $i'$ will only last until $j'$ hits position $m$, as long as this re-synchronization happens before $i$ or $i'$ hit either position $m$ or $n$ themselves (moving into a different "part" of the deck with different rules and potentially causing more trouble). The reason for this, in short, is because if we look at all the steps in between when $j$ and when $j'$ hit position $m$, as long as $i$ and $i'$ both stay in the bottom part of the deck then they will both use the same number of coins from $S^i$ plus one of the same coin from $B^j$. If instead $i$ and $i'$ stay in the top of the deck, it is even easier: $i$ and $i'$ deterministically move down one position each step. If we can guarantee that

$$\|p(i_t) - p(j_t)\|, \|p(i'_t) - p(j'_t)\| \geq |j'_t - j_t|_M + \sqrt{n}$$

for all $t$ then we can ensure that re-synchronization happens before $i$ or $i'$ hit position $m$ or $n$. This is because if, for example, $j'_t < m < j_t$ then we know $i_t, i'_t \neq m$ because otherwise we would have

$$\|p(i_t) - p(j_t)\| < |j'_t - j_t|_M$$
$$\text{or } \|p(i'_t) - p(j_t)'\| < |j'_t - j_t|_M.$$

We also know that $i_t, i'_t \neq n$ because otherwise we would have

$$\|p(i_t) - p(j_t)\| \leq \|p(n) - p(m)\| + |m - j_t| < \sqrt{n} + |j'_t - j_t|_M$$
$$= \text{ or } \|p(i'_t) - p(j'_t)\| \leq \|p(n) - p(m)\| + |m - j_t| < \sqrt{n} + |j'_t - j_t|_M.$$

The equivalent logic works comparing the distances between $i, i'$ to $i, k$ and $i', k'$ as well as comparing the distances between $j, j'$ to $j, k$ and $k, k'$. In short, as long as we keep each card closer to its counterpart than to the other tracked cards, we don't need to worry about desynchronizaiton.

This same reasoning is how we define $\Delta(i,t), \Delta(j,t)$ and $\Delta(k,t)$. It would be a mistake to say $\tau_1$ is always when $i_t = i'_t$. This is because, for example, if $i_t = i'_t$ during some time $t$ when $j_t < m < j'_t$, if we start Phase 2 and switch $i$ to start using coins from $S^i$ then $i$ and $i'$ will become decoupled after $j$ reaches position $m$. So, we define $\Delta(i,t)$ to be such that if $j, j'$ or $k, k'$ straddle position $m$ at time $t$ then if we start Phase 2 when $i_t = i'_t + \Delta(i,t)$ we will have $i_r = i'_r$ AND $i_r, i'_r$ at the same point in the sequence $S^i$ at the soonest time $r > t$ when $j, j'$ or $k, k'$ stop straddling $m$. Define $\Delta(j,t)$ and $\Delta(k,t)$ similarly for $j$ and $k$.

By our previous discussion, we know that $|\Delta(i,t)|, |\Delta(j,t)|, |\Delta(k,t)| \leq 1$ for all $t$ as long as $i, j, k$ and $i', j', k'$ stay closer to their counterparts than to each other. This will help us ensure that $i, j, k$ couple to $i', j', k'$. If for example $p(i_0) = p(i'_0) - \ell$ then as previously explained $i$ will "follow behind" $i'$. If we can show that at some point, due to a surplus of Tails flipped by $i$ in the bottom part of the deck, that $i$ "passes" $i'$ then $i$ must hit $i' - 1, i', i' + 1$ on the way by and is therefore guaranteed to couple with $i'$. The same idea is used to show $j$ couples with $j'$ and $k$ couples with $k'$.

Now that we have described the coupling, it is time to use it in the main proposition of this section.

PROPOSITION 3.5.16. *There exist universal constants $C, D$ such that the following holds: Fix $\ell > CL\sqrt{n}$. Suppose $i, j, k$ are cards such that $\|p(i) - p(j)\|, \|p(i) - p(k)\|, \|p(j) - p(k)\| > 199\ell$. Let $T \in [\ell^2, \ell^2 + 4n]$ such that $T + \left\lfloor \frac{T}{2n} \right\rfloor m \equiv 0 \mod 2n - m + 1$. Define the set*

$$N_i = \left\{ \omega \text{ such that } p(i) + a + bm = p(\omega) \mod 2n - m + 1 \text{ with } |a| \leq \ell,\ |b| \leq \frac{\ell}{\sqrt{n}} \right\}.$$

*Define $N_j$ and $N_k$ similarly for $j$ and $k$. Then,*

$$\mathbb{P}\left(i_T = f_i, j_T = f_j, k_T = f_k\right) \geq \frac{D\gamma^3 n^{\frac{3}{2}}}{\ell^6}$$

*for at least* $\frac{3}{4}$ *of all* $(f_i, f_j, f_k)$ *in* $N_i \times N_j \times N_k$.

PROOF. Let $i', j', k'$ be chosen uniformly from positions such that $|p(i') - p(i)|_M, |p(j') - p(j)|_M, |p(k') - p(k)|_M < 9\ell$ as described by the previous coupling. Run two Overlapping Cycle shuffles $\pi$ and $\pi'$ tracking $i, j, k$ and $i', j', k'$ respectively and let $\pi$ be coupled to $\pi'$ as previously described. Now let $H$ be the event that the following is true:

- There exists $r_i, r_j, r_k, s_i, s_j, s_k \leq \frac{\ell^2}{3L}$ such that $\text{Diff}_{r_i}(X^i), \text{Diff}_{r_j}(Y^j), \text{Diff}_{r_k}(Z^k) = 34\ell$ and $\text{Diff}_{s_i}(X^i), \text{Diff}_{s_j}(Y^j), \text{Diff}_{s_k}(Z^k) = -34\ell$.
- $|\text{Diff}_t(X^i)|, |\text{Diff}_t(Y^j)|, |\text{Diff}_t(Z^k)| \leq 35\ell$ for all $t \leq T$.

In other words, $H$ is the event that when $i, j, k$ are alone in the bottom part of the deck in Phase 1, Phase 2, and Phase 3 respectively, they oscillate greatly so that they will be likely to pass by (and couple with) $i', j', k'$. Since we are controlling movement within at most $3 \cdot 35 \cdot L$ standard deviations, we see that

$$\mathbb{P}(H) \geq D_1 \exp(-C_1 L^2)$$

for some universal constants $C_1, D_1$. Note that the event $H$ is independent of the distribution of $i', j', k'$ since $H$ concerns coins which are not used to generate the paths of $i', j', k'$. We now show, conditioning on $H$, that $(i, j, k)$ is likely to couple to $(i', j', k')$. To see this, let $W^j$ be the record of coins used by $j$ before time $\tau_1$ while in the bottom part of the deck. Let $W^k$ be the record of coins used by $k$ before time $\tau_2$ while in the bottom part of the deck. Let $V^i, V^j, V^k$ be the records of coins used by $i', j', k'$ before time $T$ while in the bottom part of the deck. Let $G$ be the event that the following is true:

- $|\text{Diff}_t(W^j)|, |\text{Diff}_t(W^k)|, |\text{Diff}_t(V^i)|, |\text{Diff}_t(V^j)|, |\text{Diff}_t(V^k)| \leq 12\ell$ for all $t \leq T$.

Then by Theorem A.0.6 we have that

$$\mathbb{P}(G) \geq 1 - 5 \cdot 4 \exp\left(-\frac{12^2}{2}\right) \geq 1 - 10^{-30}.$$

Note that $G$ is independent of $H$ because $G$ is a record of coins separate from the coins that determine $H$. Let $Q$ be the event that $i$ spends at least $\frac{\ell^2}{3L}$ steps in the bottom part of the deck in the first $\frac{\ell^2}{3}$ steps of Phase 1, and $j$ spends at least $\frac{\ell^2}{3L}$ steps in the bottom part of the deck in the first $\frac{\ell^2}{3}$ steps of Phase 2, and $k$ spends at least $\frac{\ell^2}{3L}$ steps in the bottom part of the deck in the first

$\frac{\ell^2}{3}$ steps of Phase 3. By Proposition 3.5.3 we know that

$$\mathbb{P}(Q) \geq 1 - 3\exp\left(-\frac{1}{2}L^2C^2\right).$$

Note that,

$$\begin{aligned}
\mathbb{P}(Q \mid H) &= 1 - \mathbb{P}\left(Q^C \mid H\right) \\
&= 1 - \frac{\mathbb{P}\left(Q^C, H\right)}{\mathbb{P}(H)} \\
&\geq 1 - \frac{\mathbb{P}\left(Q^C\right)}{\mathbb{P}(H)} \\
&\geq 1 - \frac{D_1\exp(C_1L^2)}{3\exp(\frac{1}{2}L^2C^2)}.
\end{aligned}$$

Using the fact that $L > 192$ we universally choose $C$ large enough that $\mathbb{P}(Q \mid H) > 1 - 10^{-30}$. Note that on the events $H, G, Q$ we can show that $\tau_3 < T$. This is because $G$ forces $i', j', k'$ to stay within $12\ell$ steps of their expected position. On the other hand, $i$ starts out within distance $9\ell$ distance of $i'$ in phase 1, and for $j$ and $k$ the event $G$ means that $j$ and $k$ don't drift more than $12\ell$ steps further than this initial $9\ell$ due to $W^j, W^k$. Thus, the total distance $i, j, k$ are stretched from $i', j', k'$ respectively before we count $X^i, Y^j, Z^k$ is at most $9\ell + 12\ell + 12\ell = 33\ell$. Conditioning on $H$ and $Q$ we know that $X^i$ overcomes this gap in the first $\frac{\ell^2}{3L}$ coins, and that $X^i$ uses these coins in $\frac{\ell^3}{3}$ steps. So at some point in the first $\frac{\ell^2}{3}$ steps $i$ passes by $i'$ and couples to end Phase 1. By a similar argument Phase 2 and Phase 3 last no more than $\frac{\ell^2}{3}$ steps each.

Finally, it we will need to ensure $i, j, k$ stay separated from each other so that $i, j, k$ don't decouple from $i', j', k'$. Let $E_i$ be the event that in between the $x$th time $i$ hits position $m$ and the $x$th time $i'$ hits position $m$, none of $j, j', k, k'$ hit position $m$ (where we don't count the first time $i'$ hits position $m$ if $i' \leq m < i$ and we don't count the first time $i$ hits position $m$ if $i \leq m < i'$). Let $E_j$ and $E_k$ be the equivalent events for $j, j'$ and $k, k'$. Now we want to show that $E_i, E_j, E_k$ are all likely, even when conditioning on $H$. To see why, let $U^i, U^j, U^k$ be the records of all coins used by $i, j, k$ while in the bottom part of the deck except those drawn from $X^i, Y^j, Z^k$. Let $V^i, V^j, V^k$ be the records of all coins used by $i', j', k'$ while in the bottom part of the deck. Then as long as

$$|\text{Diff}_t(U^i)|, |\text{Diff}_t(U^j)|, |\text{Diff}_t(U^k)|, |\text{Diff}_t(V^i)|, |\text{Diff}_t(V^j)|, |\text{Diff}_t(V^k)| \leq 12\ell \text{ for all } t \leq T$$

and

$$|\text{Diff}_t(B^i)|, |\text{Diff}_t(B^k)|, |\text{Diff}_t(B^k)| \leq \frac{12\ell}{\sqrt{n}} \text{ for all } t \leq \frac{T}{n},$$

then the gaps in norm between $i, j, k$ and the gaps in norm between $i', j', k'$ will never smaller than the gaps between $i, i'$ and $j, j'$ and $k, k'$. This is because the gaps between $i, j, k$ start out at more than $199\ell$ and the gaps between $i', j', k'$ start out more than $199\ell - 2 \cdot 9\ell$. Our restrictions on the $U^i, U^j, B^i, B^j, X^i, Y^j$ sequences mean that these gap between $i$ and $j$ can shrink at most to $199\ell - (12 + 12 + 12 + 12 + 35 + 35)\ell = 81\ell$. Our restrictions on $V^i, V^j, B^i, B^j$ mean that the gap between $i'$ and $j$ can shrink at most to $199\ell - 2 \cdot 9\ell - (12 + 12 + 12 + 12)\ell = 133\ell$. On the other hand, due to our restrictions on $U^i, V^i, X^i$ the gap between $i$ and $i'$ which start out at most $9\ell$ can grow to at most $9\ell + (12 + 12 + 35)\ell = 68\ell$. Similarly the gap between $j$ and $j'$ can grow to at most $68\ell$. Since $68\ell < 81\ell, 141\ell$ we don't have to worry about $i, j$ or $i', j'$ interfering with each other and causing a decoupling.

All constants used are symmetric so same reasoning applies to the pairs $i, k$ with $i', k'$ and $j, k$ with $j', k'$. There are 9 coin sequences we need to bound ($U^i, U^j, U^k, V^i, V^j, V^k, B^i, B^j, B^k$ excluding $X^i, X^j, X^k$ because we already have those bounds from event $H$) within 10 standard deviations. So we calculate

$$\mathbb{P}\left(E_i, E_j, E_k \mid H\right) \geq 1 - 9 \cdot 4 \exp\left(-\frac{12^2}{2}\right) \geq 1 - 10^{-30}.$$

All together this means

$$\mathbb{P}\left(G, Q, E_i, E_j, E_k \mid H\right) \geq 1 - 3 \cdot 10^{-20}.$$

Recall that the distribution of $i', j', k'$ is independent of $H$. By Lemma 3.5.12 we get that

$$|N_i \times N_j \times N_k| \geq \frac{2^3 \ell^6}{\gamma^3 n^{\frac{3}{2}}}.$$

By Corollary 3.5.15 we have that

$$\mathbb{P}\left((i'_T, j'_T, k'_T) = (f_i, f_j, f_k) \mid H\right) \geq 10^{-23} \cdot \frac{\gamma^3 n^{\frac{3}{2}}}{\ell^6}$$

for all $(f_i, f_j, f_k) \in N_i \times N_j \times N_k$. Let $\mathcal{D} = 10^{23}$. Then,

$$\begin{aligned}\frac{1}{\mathcal{D}|N_i \times N_j \times N_k|} &\leq \left(10^{23} \cdot \frac{2^3 \ell^6}{\gamma^3 n^{\frac{3}{2}}}\right)^{-1} \\ &\leq 10^{-23} \cdot \frac{\gamma^3 n^{\frac{3}{2}}}{\ell^6} \\ &\leq \mathbb{P}\left((i'_T, j'_T, k'_T) = (f_i, f_j, f_k) \mid H\right)\end{aligned}$$

and

$$1 - \frac{1}{8\mathcal{D}} \leq 1 - 3 \cdot 10^{-30} \leq \mathbb{P}\left(G, Q, E_i, E_j, E_k \mid H\right).$$

So by Theorem A.0.1 in the appendix we have

$$\mathbb{P}\left((i'_T, j'_T, k'_T) = (f_i, f_j, f_k) \mid H, G, Q, E_i, E_j, E_k\right) \geq \frac{1}{2} \cdot 10^{-23} \cdot \frac{\gamma^3 n^{\frac{3}{2}}}{\ell^6}$$

for at least $\frac{3}{4}$ of all $(f_i, f_j, f_k) \in N_i \times N_j \times N_k$. Since $H, G, Q, E_i, E_j, E_k$ ensure that $(i, j, k)$ couple to $(i', j', k')$, we arrive at the statement of the theorem. □

Now we one again exploit the symmetry between the overlapping cycles shuffle and the inverse overlapping cycles shuffle to show the equivalent statement holds in reverse.

Corollary 3.5.17. *For the same universal constants in Theorem 3.5.16 the following holds: Fix $\ell > CL\sqrt{n}$. Suppose $i, j, k$ are cards such that $\|p(i) - p(j)\|, \|p(i) - p(k)\|, \|p(i) - p(k)\| > 199\ell$. Let $T \in [\ell^2, \ell^2 + 4n]$ such that $T + \left\lfloor \frac{T}{2n} \right\rfloor m \equiv 0 \mod 2n - m + 1$. Define the set*

$$N_i = \left\{\omega \text{ such that } p(i) + a + bm = p(\omega) \mod 2n - m + 1 \text{ with } |a| \leq \ell,\ |b| \leq \frac{\ell}{\sqrt{n}}\right\}$$

*Define $N_j$ and $N_k$ similarly for $j$ and $k$. Now let $i_{(-T)}$, $j_{(-T)}$, $k_{(-T)}$ be the locations of $i, j, k$ after doing $T$ steps of the inverse overlapping cycles shuffle. Then,*

$$\mathbb{P}\left(i_{(-T)} = f_i, j_{(-T)} = f_j, k_{(-T)} = f_k\right) \geq \frac{D\gamma^3 n^{\frac{3}{2}}}{\ell^6}$$

*for at least $\frac{3}{4}$ of all $(f_i, f_j, f_k)$ in $N_i \times N_j \times N_k$*

Proof. This corollary holds because the inverse overlapping cycles shuffle is itself an overlapping cycles shuffle up to the reordering of the cards in Proposition 3.1.1. According to Corollary

3.3.12 the distances $|p(\cdot) - p(\cdot)|_M$ are $\|p(\cdot) - p(\cdot)\|$ are fixed under this reordering. Since the parameters for $i, j, k$ and $N_i, N_j, N_k$ are defined using these distances the proposition holds for the inverse overlapping cycles shuffle. □

We are now very close to what we want. We have shown that $i, j, k$ reach an approximately uniform distribution, but it's a distribution that excludes up to $\frac{1}{4}$ of our desired terms. What we really want is a distribution that includes *every* term at a probability of a constant times the number of terms. To accomplish this, we make the following argument:

Pick some cards $i, j, k$ and some reasonable targets $f_i, f_j, f_k$. Now run the overlapping cycles shuffle forward from $i, j, k$ and backwards from $f_i, f_j, f_k$. In the middle, $i, j, k$ and $f_i, f_j, f_k$ will be distributed over $\frac{3}{4}$ of all nearby terms. So a constant fraction of those terms will overlap, and we can show that $i, j, k$ goes to $f_i, f_j, f_k$ by summing over those overlapping middle terms. We formalize this in the following Theorem.

THEOREM 3.5.18. *There exist universal constants $C, D$ such that the following holds: Fix $\ell > CL\sqrt{n}$. Fix any cards $i, j, k$ such that $\|i-k\|, \|i-k\|, \|j-k\| > 199\ell$. Fix any positions $f_i, f_j, f_k$ such that $\|i-f_i\|, \|j-f_j\|, \|k-f_k\| < \frac{\ell}{10}$ and $\|f_i-f_k\|, \|f_i-f_k\|, \|f_j-f_k\| > 199\ell$. Let $T \in (\ell^2, \ell^2+2n)$ such that $T + \lfloor \frac{T}{2n} \rfloor m \equiv 0 \mod 2n-m+1$. Then,*

$$\mathbb{P}\left((i_{2T}, j_{2T}, k_{2T}) = (f_i, f_j, f_k)\right) \geq \frac{D\gamma^3 n^{\frac{3}{2}}}{\ell^6}$$

PROOF. Define the set

$$N_i = \left\{ \text{positions } \omega \text{ such that } p(i) + a + bm = p(\omega) \mod 2n-m+1 \text{ with } |a| \leq \ell,\ |b| \leq \frac{\ell}{\sqrt{n}} \right\}.$$

Define the set

$$N_i' = \left\{ \text{positions } \omega \text{ such that } p(f_i) + a + bm = p(\omega) \mod 2n-m+1 \text{ with } |a| \leq \ell,\ |b| \leq \frac{\ell}{\sqrt{n}} \right\}.$$

Note that because $\|i - f_i\| < \frac{\ell}{10}$ we know there exists $a_i, b_i$ such that $p(i) + a_i + b_i m$ with $|a_i| < \frac{\ell}{10}, |b_i| < \frac{\ell}{10\sqrt{n}}$. For this reason $N_i$ and $N_i'$ overlap at least $(\frac{9}{10})^2 > \frac{4}{5}$ of their elements. Define $N_j, N_k, N_j', N_k'$ similarly for $j, k, f_j, f_k$. Then $N_i \times N_j \times N_k$ and $N_i' \times N_j' \times N_k'$ overlap at least $(\frac{4}{5})^3 \geq \frac{51}{100}$ of their elements. By Proposition 3.5.16 and Corollary 3.5.17 we can bound the distribution

of $(i_T, j_T, k_T)$ from below over $\frac{3}{4}$ of the elements of $N_j, N_k, N_k$ and bound the distribution of $(f_{i(-T)}, f_{j(-T)}, f_{k(-T)})$ from below over $\frac{3}{4}$ of the elements of $N'_i, N'_j, N'_k$. Let $\mathcal{S}$ be the subset of $(N_i \times N_j \times N_k) \cap (N'_i \times N'_j \times N'_k)$ where the bounds from Proposition 3.5.16 and Corollary 3.5.17 hold. Then,

$$|\mathcal{S}| \geq \left(1 - \frac{1}{4} - \frac{1}{4} - \frac{49}{100}\right)|N_i \times N_j \times N_k| = \frac{1}{100}|N_i \times N_j \times N_k|.$$

If we now run the shuffle $T$ steps forward from $i, j, k$ and $T$ steps backwards from $f_i, f_j, f_k$ we can compute the probability that they meet in the middle.

$$\begin{aligned}
&\mathbb{P}\left((i_{2T}, j_{2T}, k_{2T}) = (f_i, f_j, f_k)\right) \\
&\geq \sum_{(z_i,z_j,z_k)\in\mathcal{S}} \mathbb{P}\left((i_{2T}, j_{2T}, k_{2T}) = (z_i, z_j, z_k)\right) \cdot \mathbb{P}\left((f_{i(-T)}, f_{j(-T)}, f_{k(-T)}) = (z_i, z_j, z_k)\right) \\
&\geq \sum_{(z_i,z_j,z_k)\in\mathcal{S}} \frac{D_1^2\gamma^6 n^3}{\ell^{12}} \\
&= \frac{D_1^2\gamma^6 n^3}{100\ell^{12}}|N_i \times N_j \times N_k|.
\end{aligned}$$

As we showed in Lemma 3.5.12,

$$|N_i| \geq \frac{2\ell^2}{\gamma\sqrt{n}}.$$

By symmetry the same bound applies to $N_j, N_k$ so

$$|N_i \times N_j \times N_k| \geq \frac{8\ell^6}{\gamma^3 n^{\frac{3}{2}}}.$$

This gives us,

$$\mathbb{P}\left((i_{2T}, j_{2T}, k_{2T}) = (f_i, f_j, f_k)\right) \geq \frac{8D_1^2\gamma^3 n^{\frac{3}{2}}}{100\ell^6}$$

□

### 3.6. Entropy Decay

In this section we will find our mixing time bound for the overlapping cycles shuffle. We will do this by applying Theorem 3.5.18 to Theorem 3.4.4 to bound each $A_x$ in the sum. As previously explained, we need $i, j, k$ to be spread out to apply Theorem 3.5.18, and so we use Theorem 3.5.4

and Corollary 3.5.5 as part of Stage 1 and Stage 3 to spread out $i, j, k$ and our targets $g_i, g_j, g_k$ to meet the requirements of $i, j, k$ and $g_i, g_j, g_k$ being distance $199\ell$ from each other. However there is one small technicality we have not addressed: How do we even know there exist positions $f_i, f_j, f_k$ that are spread out from each other and also a reachable distance from $i, j, k, g_i, g_j, g_k$? In other words, do there exist $f_i, f_j, f_k$ such that $\|p(f_i)\|, \|p(f_j)\|, \|p(f_k)\| < \ell$ but $\|p(f_i) - p(f_j)\|, \|p(f_i) - p(f_k)\|, \|p(f_j) - p(f_k)\| > c\ell$ for some not-to-small constant $c$? It seem clear that such positions should exist but a priori it is not obvious why. We deal with this in the following Theorem.

LEMMA 3.6.1. *Fix some $\ell$ such that $100\sqrt{n} < \ell < \ell_{\max}$ where $\ell_{\max}$ is the maximum value of $\|\cdot\|$ for the shuffle. There exist positions $f_i, f_j, f_k$ such that*

- $\|p(f_i)\|, \|p(f_j)\|, \|p(f_k)\| < \ell$
- $\|p(f_i) - p(f_j)\|, \|(f_i) - p(f_k)\|, \|p(f_j) - p(f_k)\| > \frac{\ell}{5}$

PROOF. Let $z$ be an element of $\{1, \ldots, 2n - m\}$ that maximizes $\|z\|$. Then $z = a + bm$ where $\|z\| = |a| + |b|\sqrt{n}$. For any constant $k \in (0, 1)$, let $\left\langle \frac{z}{2} \right\rangle$ be defined by

$$\left\langle \frac{z}{2} \right\rangle = \left\lceil \frac{a}{2} \right\rceil + \left\lceil \frac{b}{2} \right\rceil m.$$

. Then,

$$\left\| \left\langle \frac{z}{2} \right\rangle \right\| \leq \frac{1}{2}\|z\| + 1 + \sqrt{n},$$
$$\left\| \left\langle \frac{z}{2} \right\rangle \right\| \geq \frac{1}{2}\|z\| - 1 - \sqrt{n}.$$

The second item is true because

$$\|z\| \leq \left\| \left\langle \frac{z}{2} \right\rangle + \left\langle \frac{z}{2} \right\rangle \right\| + 2 + 2\sqrt{n} \leq 2\left\| \left\langle \frac{z}{2} \right\rangle \right\| + 2 + 2\sqrt{n}.$$

Also note that

$$\begin{aligned} \left\| z - \left\langle \frac{z}{2} \right\rangle \right\| &= \left\| \left( a - \left\lceil \frac{a}{2} \right\rceil \right) + \left( b - \left\lceil \frac{b}{2} \right\rceil \right) m \right\| \\ &= \left\| \left\lfloor \frac{a}{2} \right\rfloor + \left\lfloor \frac{b}{2} \right\rfloor m \right\| \\ &\geq \left\| \left\langle \frac{z}{2} \right\rangle \right\| - 2 - 2\sqrt{n}. \end{aligned}$$

Define $\left\langle \frac{z}{4} \right\rangle$ as $\left\langle \frac{\left\langle \frac{z}{2} \right\rangle}{2} \right\rangle$ and $\left\langle \frac{z}{8} \right\rangle$ as $\left\langle \frac{\left\langle \frac{z}{4} \right\rangle}{2} \right\rangle$, etc. Then it follows inductively that

$$\left\| \left\langle \frac{z}{2^x} \right\rangle \right\| \leq \frac{1}{2^x}\|z\| + 2 + 2\sqrt{n}$$
$$\left\| \left\langle \frac{z}{2^x} \right\rangle \right\| \geq \frac{1}{2^x}\|z\| - 2 - 2\sqrt{n}.$$

Now choose $x$ such that

$$\frac{\ell}{5} < \frac{1}{2^{x+1}}\|z\| - 6 - 4\sqrt{n} < \frac{1}{2^x}\|z\| + 2 + 2\sqrt{n} < \ell. \tag{3.7}$$

Using the inequality on the right of (3.7) gives us

$$\left\| \left\langle \frac{z}{2^{x+1}} \right\rangle \right\| < \left\| \left\langle \frac{z}{2^x} \right\rangle \right\| < \ell.$$

Using the inequality on the left of (3.7) gives us

$$\begin{aligned} \left\| \left\langle \frac{z}{2^x} \right\rangle - \left\langle \frac{z}{2^{x+1}} \right\rangle \right\| &\geq \left\| \left\langle \frac{z}{2^{x+1}} \right\rangle \right\| - 2 - 2\sqrt{n} \\ &\geq \frac{1}{2^{x+1}}\|z\| - 4 - 4\sqrt{n} \\ &\geq \frac{\ell}{5} + 2. \end{aligned}$$

Let $f_i$ be position 1 in the deck and let $f_j, f_k$ be positions in the deck such that

$$\left| p(f_j) - \left\langle \frac{z}{2^{x+1}} \right\rangle \right| \leq 1$$
$$\left| p(f_k) - \left\langle \frac{z}{2^x} \right\rangle \right| \leq 1.$$

Then,

$$\|p(f_i) - p(f_j)\| \geq \left\| \left\langle \frac{z}{2^{x+1}} \right\rangle \right\| - 2 \geq \frac{1}{2^{x+1}}\|z\| - 4 - 2\sqrt{n} \geq \frac{\ell}{5}$$
$$\|p(f_i) - p(f_k)\| \geq \left\| \left\langle \frac{z}{2^x} \right\rangle \right\| - 2 \geq \frac{1}{2^x}\|z\| - 4 - 2\sqrt{n} \geq \frac{\ell}{5}$$
$$\|p(f_j) - p(f_k)\| \geq \left\| \left\langle \frac{z}{2^x} \right\rangle - \left\langle \frac{z}{2^{x+1}} \right\rangle \right\| - 2 \geq \frac{\ell}{5} + 2 - 2 \geq \frac{\ell}{5}.$$

□

We now have finished all the work necessary to justify Stage 1, Stage 2, and Stage 3. There is one final proposition we need to show before we apply Theorem 3.4.4. We need to show that after $i, j, k$ get close together, they are likely to collide with each other.

PROPOSITION 3.6.2. *Consider the overlapping cycles shuffle where* $10\sqrt{n} < m < n - 10\sqrt{n}$. *Consider any cards* $i, j, k$ *and suppose that* $|p(i_T) - p(j_T)|_M, |p(i_T) - p(k_T)|_M, |p(j_T) - p(k_T)|_M < \sqrt{n}$. *Let* $E$ *be the event that the next time* $i$ *or* $j$ *or* $k$ *collides after time* $T$ *it is with each other and in the order* $(i, k, j)$. *Let* $G$ *be the event that the next time* $i$ *collides after time* $T$ *it is before time* $T + 10n$. *Then,*

$$\mathbb{P}(E, G) \geq \frac{D}{n} \exp\left(-\frac{8n}{m}\right)$$

*for a universal constant* $D$.

Intuitively, this makes sense. If $i, j, k$ begin distance $\sqrt{n}$ from each other with respect to $\|\cdot\|$ then by the analysis we have done in previous sections we should believe that $i, j, k$ can travel to positions $m - 1, n, m$ respectively in $C(\sqrt{n})^2 = Cn$ steps, as positions $m - 1, n, m$ are also distance $\sqrt{n}$ from each other with respect to $\|\cdot\|$. The only wrinkle is making sure $i, j, k$ do not collide with any other cards along the way.

PROOF. Let $\tau_1$ be the random stopping time given by the minimal $t \geq T$ such that $i, j, k \in (m, n)$. Let $H_1$ be the event that

- $i, j, k$ each flip Tails on their first visit to position $m$ before time $\tau_1$ (if it exists),
- $i, j, k$ do not experience any collisions between time $T$ and $\tau_1$.

Note that we can guarantee that $i, j, k$ do not experience any collisions between time $T$ and $\tau_1$ if, in addition to flipping Tails when $i, j, k$ are in position $m$, we also have Tails flips in the previous or subsequent flips (depending on if $i, j, k$ reach $m$ in an even or odd number of steps). So $\mathbb{P}(H_1) \geq \left(\frac{1}{2}\right)^6$ accounting for at most 6 Tails flips. On $H_1$ we expect $\tau_1 - T$ to not be much larger than $m$ as $i, j, k$ start out close together. In the longest case we have something like $i_T = 1, j_T = 2$, and $k_T = n - \sqrt{n}$. It will likely take about $m + 2\sqrt{n}$ steps for $k_T$ to reach position $m + 1$, and at that point cards $i$ and $j$ will be in $(m, n)$ as well. Let $H_2$ be the event that $\tau_1 - T < 2n$. Then we know

$H_2$ occurs with high probability. In particular we have

$$\mathbb{P}(H_1, H_2) \geq \frac{1}{2^7}.$$

Let $\tau_2$ be the random stopping time given by the minimal $t > \tau_1$ such that $i_t, j_t, k_t \in (1, 10\sqrt{n})$. In other words $\tau_2$ is the time after $\tau_1$ when $i, j, k$ reach the top part of the deck. Let $\Delta_i, \Delta_j$ be defined by

$$\Delta_i = i_{\tau_2} - k_{\tau_2},$$

$$\Delta_j = j_{\tau_2} - k_{\tau_2}.$$

Since $i, j, k$ are at most $n - m < n - 10\sqrt{n}$ steps from the bottom of the deck at time $\tau_1$ we know the event $\tau_2 - \tau_1 < 2n$ has high probability. In this case we have $\tau_1 + \tau_2 - T < 4n$ and we expect $i, j, k$ to drift at most $4\sqrt{n}$ steps further from each other in these steps. Let $H_3$ be the event that

- $|\Delta_i|, |\Delta_j| < 4\sqrt{n}$,
- $\tau_2 - \tau_1 < 2n$,
- $i, j, k$ experience no collisions between time $\tau_1$ and $\tau_2$.

Then $\mathbb{P}(H_3 \mid H_1, H_2) > D_1$ for a universal constant $D_1$ because the first two items in $H_3$ follow from modest constraints on the Heads-Tails differentials of $i, j, k$ while in the bottom part of the deck before $\tau_2$. The last item follows from $i, j, k$ flipping tails twice in a row when in position $n - 1$ or $n$ depending on if they reach such a position at an even or odd time) which occurs with probability bounded away from 0. Let $\tau_3$ be the random stopping time given by the minimal $t > \tau_2$ such that $j_t = n$. Let $H_4$ be the event that after $\tau_2$,

- the next two times $i$ is in position $m$ it flips Tails, Heads,
- the next two times $k$ is in position $m$ it flips Heads, Tails,
- the next three times $j$ is in position $m$ it flips Heads, Heads, Tails,
- $i, j, k$ experience no collisions between time $\tau_2$ and $\tau_3$,
- $\tau_3 - \tau_2 < 3m + 3n$.

We claim that

$$\mathbb{P}(H_4 \mid H_1, H_2, H_3) \geq D_2$$

This is because the first four items in $H_4$ follow from flips $i, j, k$ make at the finitely many instances when they are position $m$ or $n$. For the last item, note that we expect $\tau_3 - \tau_2$ to last around $3m + 2(n-m) \pm \sqrt{n}$ flips because $j$ will make three loops between 1 and $m$ and then need to travel $n - m$ positions beyond $m$ to position $n$, moving down on average once every other flip. So $\tau_3 - \tau_2 < 3m + 3n$ happens with high probability.

Now let $H_5$ be the event that

- $i_{\tau_3} = m - 1$,
- $k_{\tau_3} = m$.

We claim that

$$\mathbb{P}\left(H_5 \mid H_1, H_2, H_3, H_4\right) \geq \frac{D_3}{n} \exp\left(-\frac{8n}{m}\right)$$

for a universal constant $D_3$. To see why, note that on $H_4$ we know $i$ will spend exactly $m + \Delta_i$ steps alone in the bottom part of the deck before $k$ enters the bottom part of the deck, and $k$ will spend exactly $m - \Delta_j$ steps in the bottom part of the deck (with or without $i$) before $j$ enters the bottom part of the deck. Finally, $j$ will spend some time alone in the bottom part of the deck after $i, k$ exit to the top. Let $S^i, S^{ik}, S^k, S^{ijk}, S^{jk}, S^j$ be the records of flips during these times, where $S^A$ corresponds to when cards in the set $A$ are in the bottom part of the deck. Note that some of these sequences may be empty, but we know that $S^i$ has at least $m + \Delta_i$ coins and $S^{ik}, S^k$ combined have at least $m - \Delta_j$ coins. Thus, $i$ spends enough time in the bottom of the deck alone and $k$ spends enough time in the bottom of the deck apart from $j$ that there is ample time for $i_t - k_t \mod m$ and $j_t - k_t \mod m$ to vary according to a binomial random variable with at least $m + \Delta_i$ and $m - \Delta_k$ trials respectively. We want the outcomes of these binomial random variables to be $\Delta_i$ and $\Delta_k$ off their mean, which represents at most $\frac{2\sqrt{n}}{\sqrt{m}}$ standard deviations. Thus,

$$\mathbb{P}\left(H_5 \mid H_1, H_2, H_3, H_4\right) \geq \left(\frac{D_4}{\sqrt{n}} \exp\left(-\frac{4n}{m}\right)\right)^2$$

Now on $H_1, H_2, H_3, H_4, H_5$ we see that $i, j, k$ are in the correct position to collide in the next two steps immediately after $\tau_3$, provided that $\tau_3$ is even. Let $Q$ be the event that $\tau_3$ is even. Then

$$\mathbb{P}\left(Q \mid H_1, H_2, H_3, H_4\right) \geq \frac{1}{3}$$

This is because $i_{\tau_1}, j_{\tau_1}, k_{\tau_1}$ are all in the bottom part of the deck. There is a $\frac{1}{2}$ the first flip after $\tau_1$ is Heads and in that case $(i_{\tau_1+1}, j_{\tau_1+1}, k_{\tau_1+1}) = (i_{\tau_1}, j_{\tau_1}, k_{\tau_1})$. So an odd time for $\tau_3$ cannot be more than twice as likely as an even time.

Finally, conditioning on $H_1, H_2, H_3, H_4, H_5, Q$ there is a $\frac{1}{2}$ chance $i, j, k$ do in fact collide in the steps immediately after $\tau_3$ (as a collision occurs if Heads, Tails or Tails, Heads are flipped). In this case the collision occurs in fewer than $2n + 2n + 3m + 3n + 2 < 10n$ steps. So,

$$\mathbb{P}(E, G) \geq \frac{1}{2^7} \cdot D_1 \cdot D_2 \cdot \frac{D_3}{n} \cdot \exp\left(-\frac{8n}{m}\right).$$

□

Now we are ready to apply Theorem 3.4.4. As a reminder, our strategy is broken into 3 stages.

- In Stage 1 we use Theorem 3.5.4 to show that $i, j, k$ spread out sufficiently.
- In Stage 2 we use Theorem 3.5.18 to show that $i, j, k$ move to a precise position.
- In Stage 3 we use Theorem 3.5.5 to show that $i, j, k$ collapse back together.

Then finally we use Theorem 3.6.2 to show that $i, j, k$ collide with each other.

THEOREM 3.6.3. *There exist universal constants $C, D$ such that the following is true: Fix any $\ell$ such that $CL\sqrt{n} \leq \ell \leq \ell_{\max}$. Let $i, j, k$ be cards such that $\|p(i)\|, \|p(j)\|, \|p(k)\| < \ell$. Let $T_1 \in [\ell^2, \ell^2 + 4n]$ such that $T_1 + \left\lfloor \frac{T_1}{2n} \right\rfloor \equiv 0 \mod 2n - m + 1$. Let $T_2 \in [10^{-6}\ell^2, 10^{-6}\ell^2 + 4n]$ such that $T_2 + \left\lfloor \frac{T_2}{2n} \right\rfloor \equiv 0$. Let $t = 2T_1 + 2T_2 + 10n$. Let $E$ be the event that the first time $i$ collides after time $T$, it is with $j$ and $k$ on the front and back respectively, and it happens before time $t$. Then,*

$$\mathbb{P}(E) \geq \frac{D\gamma^2 n}{\ell^4} \exp\left(-864L^2 - \frac{8n}{m}\right).$$

PROOF. Let $N^\ell$ be the set of positions

$$N^\ell = \left\{ \text{positions } \omega \text{ such that } p(\omega) = a + bm \text{ with } |a| \leq \ell,\ |b| \leq \frac{\ell}{\sqrt{n}} \right\}$$

Choose any positions $g_i, g_j, g_k$ such that

- $p(g_i) \in N^\ell$
- $|p(g_i) - p(g_j)|_M, |p(g_i) - p(g_k)|_M, |p(g_j) - p(g_k)|_M < \sqrt{n}$

Now according to Lemma 3.6.1 there exist positions $f_i, f_j, f_k$ such that

$$\|p(f_i)\|, \|p(f_j)\|, \|p(f_k)\| < \ell,$$
$$\|p(f_i) - p(f_j)\|, \|p(f_i) - p(f_j)\|, \|p(f_i) - p(f_k)\| > \frac{\ell}{5}.$$

Then as long as $C$ is sufficiently large, we have according to Proposition 3.5.4 that

$$\mathbb{P}(E_1) := \mathbb{P}\left(\|p(i_{T_1}) - p(f_i)\|, \|p(j_{T_1}) - p(f_j)\|, \|p(k_{T_1}) - p(f_k)\| < \frac{\ell}{2000}\right) \geq D_1 \exp(-432L^2).$$

Similarly by Proposition 3.5.5 we have that

$$\mathbb{P}(E_2) := \mathbb{P}\left(\|p(g_{i(-T_1)}) - p(f_i)\|, \|p(g_{j(-T_1)}) - p(f_j)\|, \|p(g_{k(-T_1)}) - p(f_k)\| < \frac{\ell}{2000}\right) \geq D_1 \exp(-432L^2).$$

On $E_1, E_2$ we get that

- $\|p(i_{T_1}) - p(g_{i(-T_1)})\|, \|p(j_{T_1}) - p(g_{j(-T_1)})\|, \|p(k_{T_1}) - p(g_{k(-T_1)})\| < \frac{\ell}{1000}$,
- $\|p(i_{T_1}) - p(j_{T_1})\|, \|p(i_{T_1}) - p(k_{T_1})\|, \|p(j_{T_1}) - p(k_{T_1})\| > \frac{\ell}{5} - \frac{2\ell}{2000} > \frac{199\ell}{1000}$,
- $\|p(g_{i(-T_1)}) - p(g_{j(-T_2)})\|, \|p(g_{i(-T_1)}) - p(g_{k(-T_1)})\|, \|p(g_{j(-T_1)}) - p(g_{k(-T_2)})\| > \frac{\ell}{5} - \frac{2\ell}{1000} > \frac{199\ell}{1000}$.

So by Theorem 3.5.18 using $10^{-3}\ell$ in place of $\ell$ we have

$$\mathbb{P}\left((i_{T_1+2T_2}, j_{T_1+2T_2}, k_{T_1+2T_2}) = (g_{i(-T_1)}, g_{j(-T_1)}, g_{k(-T_1)}) \mid E_1, E_2\right) \geq \frac{D_2\gamma^3 n^{\frac{3}{2}}}{\ell^6}.$$

Together we get that

$$\mathbb{P}\left((i_{T_1+2T_2+T_1}, i_{T_1+2T_2+T_1}, i_{T_1+2T_2+T_1}) = (g_i, g_j, g_k)\right) \geq \frac{D_1^2 D_2\gamma^3 n^{\frac{3}{2}}}{\ell^6} \exp(-864L^2).$$

Let $H(g_i, g_j, g_k)$ be the event that the next time $g_i$ collides it is with $g_j$ as its front match and $g_k$ as its back match, and that this collision happens within $8n$ steps. As shown in Theorem 3.6.2,

$$\mathbb{P}(H(g_i, g_k, g_k)) \geq \frac{D_3}{n} \exp\left(-\frac{8n}{m}\right).$$

Let $\mathcal{R}_\ell$ be the set of choices for the triplet $(g_i, g_j, g_k)$ according to our parameters at the start of the proof. To be specific, let

$$\mathcal{R}_\ell = \{(g_i, g_j, g_k) \text{ such that } g_i, g_j, g_k \in N^\ell \text{ and } |p(g_i)-p(g_j)|_M, |p(g_i)-p(g_k)|, |p(g_j)-p(g_k)|_M < \sqrt{n}\}.$$

Then,

$$\begin{aligned}
\mathbb{P}(E) &\geq \sum_{(g_i,g_j,g_k)\in\mathcal{R}_\ell} \mathbb{P}\left(E, (i_{T_1+2T_2+T_1}, j_{T_1+2T_2+T_1}, k_{T_1+2T_2+T_1}) = (g_i, g_j, g_k)\right) \\
&\geq \sum_{(g_i,g_j,g_k)\in\mathcal{R}_\ell} \mathbb{P}\left((i_{T_1+2T_2+T_1}, j_{T_1+2T_2+T_1}, k_{T_1+2T_2+T_1}) = (g_i, g_j, g_k)\right) \cdot \mathbb{P}\left(H(g_i, g_j, g_k)\right) \\
&\geq \sum_{(g_i,g_j,g_k)\in\mathcal{R}_\ell} \frac{D\gamma^3 n^{\frac{3}{2}}}{n\ell^6} \exp\left(-864L^2 - \frac{8n}{m}\right).
\end{aligned}$$

We now need to bound $|\mathcal{R}_\ell|$. Note that $g_i$ is chosen from $N^\ell$. For each choice of $g_i \in N^\ell$ we can choose $g_j$ such that $p(g_j) = p(g_i) + \delta_j$ for $\delta_j \in (\sqrt{n}, \sqrt{n})$. At least half of these values of $\delta_j$ are associated with valid positions in the deck, so there are at least $\sqrt{n}$ choices for $g_j$. If $\delta_j$ is negative then we can choose $g_k$ such that $p(g_k) = p(g_j) + \delta_k$ with $\delta_k \in \{1, \ldots, \sqrt{n}\}$ and if $\delta_j$ is positive we can do the same with $\delta_k \in \{-1, \ldots, -\sqrt{n}\}$. This will meet the requirement that $|p(g_i) - p(g_j)|_M$ and $|p(g_i) - p(g_k)|_M$ and $|p(g_j) - p(g_k)|_M$ are all less than or equal to $\ell$. This gives us at least $\sqrt{n} \cdot \frac{\sqrt{n}}{2} = \frac{n}{2}$ choices for $(g_j, g_k)$. Thus,

$$|\mathcal{R}_\ell| \geq \frac{n}{2}|N^\ell|.$$

As we showed in Lemma 3.5.12 we have

$$|N^\ell| \geq \frac{2\ell^2}{\gamma\sqrt{n}}. \tag{3.8}$$

So,

$$|\mathcal{R}_\ell| \geq \frac{\sqrt{n}\ell^2}{\gamma}$$

and

$$\mathbb{P}(E) \geq \frac{D\gamma^2 n}{\ell^4} \exp\left(-864L^2 - \frac{8n}{m}\right).$$

□

Now that we have a bound on the $A_x$ in the sum of Theorem 3.4.4 it is time to find a bound on the mixing time. To do this we first show that the Entropy of the overlapping cycles shuffle decays at an exponential rate. As part of the proof, we will use a slight variation of Theorem 3.4.4. The only variation will be that, instead of denoting $y > x$ for cards $y, x$ where $x$ is above $y$ in the deck, we will use a different well-ordering of the deck. We will define a permutation $\nu$ which translates from the top-to-bottom ordering to our new well-ordering. Using this well-ordering Theorem 3.4.4 will still hold, because the ordering assumed in the Theorem is arbitrary. Since the Theorem is defined for any random permutation, there is no reason any particular ordering is preferred.

LEMMA 3.6.4. *Consider the overlapping cycles shuffle with $n$ cards and parameter $m$. There exist universal constants $C, D$ such that the following is true: If $\pi_t$ is the overlapping cycles shuffle with $t$ steps then there is a value $t \in \{1, \dots, C\ell_{\max}^2\}$ such that*

$$\mathbb{E}\big[\mathrm{ENT}(\mu\pi_t|\mathrm{sgn}(\mu\pi_t))\big] \leq \left(1 - \frac{Dt}{\ell_{\max}^2 \log^2(n)} \exp\left(-864L^2 - \frac{8n}{m}\right)\right)\mathbb{E}[\mathrm{ENT}(\mu \mid \mathrm{sgn}(\mu))]$$

PROOF. Let $a = \lceil \log_2(\ell_{\max}) - \frac{1}{2}\log_2(n) \rceil$. For $k \in \{1, \dots, a\}$, let $\ell_k = 2^k\sqrt{n}$. Now we partition the deck of $n$ cards as follows:

- Let $J_0 := \{n, n-1, \dots, n - \lfloor\sqrt{n}\rfloor\}$.
- For $k \geq 1$ let $J_k = \{i \ : \ \|p(i)\| \leq \ell_k\}$.
- Let $I_0 = J_0 \backslash \{n, n-1\}$ and for $k \geq 1$ let $I_k = J_k \backslash J_{k-1}$.

Let $\nu$ be a permutation which reorders the deck with the following properties:

- $\nu(n) = 1, \nu(n-1) = 2, \dots, \nu\left(n - \lfloor\sqrt{n}\rfloor\right) = \lfloor\sqrt{n}\rfloor$
- $\nu$ respects the natural ordering of $I_k$. Specifically, if $x \in I_k$ and $y \in I_{k+1}$ then $\nu(x) < \nu(y)$.

Note that the second item does not contradict the first because $p(n-2), p(n-3), \dots, p(n - \lfloor\sqrt{n}\rfloor)$ all have norms less than or equal to $2\sqrt{n}$. Note that under this reordering, if $x \in I_k$ with $k \geq 1$ then $\nu(x) \geq |J_{k-1}|$.

For each $k$, let $\gamma_k$ be defined by

$$\gamma_k = \left|\left\{\kappa \in \mathbb{N} \ : \ |\kappa m|_M < \ell_k, \ \kappa < \frac{\ell_k}{\sqrt{n}}\right\}\right|.$$

Let $E_j = \mathbb{E}[\text{ENT}(\mu^{-1}(\nu(j)) \mid \text{sgn}(\mu), \mu^{-1}(\nu(j)+1), \mu^{-1}(\nu(j)+2), \ldots, \mu^{-1}(\nu(n))]$. Then as shown by Senda [10] in Appendix B we can decompose,

$$\mathbb{E}[\text{ENT}(\mu \mid \text{sgn}(\mu))] = \sum_{\nu(j)=3}^{n} E_j = \sum_{k=1}^{a} \sum_{j \in I_k} E_j.$$

Let $k^*$ be such that $\sum_{j \in I_{k^*}} E_j$ is maximal. Then,

$$\begin{aligned} \mathbb{E}[\text{ENT}(\mu \mid \text{sgn}(\mu))] &\leq a \sum_{j \in I_{k^*}} E_j, \\ \frac{1}{a}\mathbb{E}[\text{ENT}(\mu \mid \text{sgn}(\mu))] &\leq \sum_{j \in I_{k^*}} E_j. \end{aligned} \tag{3.9}$$

For any card $x$ in the deck let $A_x$ be the maximal value such that

$$\mathbb{P}\left(m_2(x) = y, m_1(x) = z\right) \geq \frac{A_x}{\nu(x)^2}$$

for all distinct cards $y, z$ such that $\nu(y), \nu(z) < \nu(x)$. Note that this value $A_x$ also has the property that

$$\mathbb{P}\left(m_2(x) = y, \nu(m_1(x)) < \nu(x)\right) \geq \frac{A_x}{\nu(x)}$$

for all cards $y$ such that $\nu(y) < \nu(x)$. Now, by Theorem 3.4.4, if we examine the shuffle after any number $t$ steps, we have

$$\begin{aligned} \mathbb{E}\big[\text{ENT}(\mu\pi_t|\text{sgn}(\mu\pi_t))\big] - \mathbb{E}\big[\text{ENT}((\mu|\text{sgn}(\mu)))\big] &\leq \frac{-C_1}{\log(n)} \sum_{\nu(x)=3}^{n} A_x E_x \\ &\leq \frac{-C_1}{\log(n)} \sum_{x \in I_{k^*}} A_x E_x \end{aligned} \tag{3.10}$$

where the second inequality comes from the fact that we are summing over fewer negative terms. We now consider three cases for $k^*$.

(1) $k^* = 0$

Fix any $x \in I_0$. Then $x \in [n-2, n-3, \ldots, n-\lfloor\sqrt{n}\rfloor]$. Fix $t = 2n+5\sqrt{n}$ and $T = 2n-5\sqrt{n}$.

Then by Proposition 3.4.7 we have

$$A_x \geq \frac{D_1}{\sqrt{n}} \exp\left(-\frac{2n}{m}\right).$$

Plugging into the bound from (3.10) we get

$$\mathbb{E}\big[\mathrm{ENT}(\mu\pi_t|\mathrm{sgn}(\mu\pi_t))\big] - \mathbb{E}\big[\mathrm{ENT}((\mu|\mathrm{sgn}(\mu)))\big] \leq \frac{-C_1 D_1}{\sqrt{n}\log(n)} \exp\left(-\frac{2n}{m}\right) \sum_{x \in I_{k^*}} E_x.$$

Since $t$ is less than a constant times $n$ we have,

$$\mathbb{E}\big[\mathrm{ENT}(\mu\pi_t|\mathrm{sgn}(\mu\pi_t))\big] - \mathbb{E}\big[\mathrm{ENT}((\mu|\mathrm{sgn}(\mu)))\big] \leq \frac{-D_2 t}{n^{\frac{3}{2}}\log(n)} \exp\left(-\frac{2n}{m}\right) \sum_{x \in I_{k^*}} E_x.$$

Recall that $\ell_{\max} \geq \frac{1}{2} n^{\frac{3}{4}}$, so $n^{\frac{3}{2}} \leq 4\ell_{\max}^2$. This gives us

$$\mathbb{E}\big[\mathrm{ENT}(\mu\pi_t|\mathrm{sgn}(\mu\pi_t))\big] - \mathbb{E}\big[\mathrm{ENT}((\mu|\mathrm{sgn}(\mu)))\big] \leq \frac{-D_3 t}{\ell_{\max}^2\log(n)} \exp\left(-\frac{2n}{m}\right) \sum_{x \in I_{k^*}} E_x.$$

(2) $1 \leq k^* \leq \log_2(C_2 L)$ where $C_2$ is the universal constant from Theorem 3.6.3.

In this case all cards $x$ in $I_k^*$ have $\|x\| < C_1 L\sqrt{n}$. Set $t = (C_2^2 L^2)(2+2\cdot 10^{-6})n + 4\cdot 4n + 10n$. Then by Theorem 3.6.3 for each $A_x$ in the sum $\sum_{x \in I_{k^*}} A_x E_x$ we have

$$A_x = \mathbb{P}(E) \cdot \nu(x)^2 \geq \frac{D_4 \gamma_{k^*}^2 n}{\ell^4} \exp\left(-864L^2 - \frac{8n}{m}\right) \cdot \nu(x)^2$$

where $E$ is the event in the statement of that Theorem and $\ell = C_1 L\sqrt{n}$ and $\nu(x) \geq \sqrt{n}$ and $\gamma_{k^*} \geq 1$. Plugging in this information we get

$$A_x \geq \frac{D_4}{C_2^2 L^2} \exp\left(-864L^2 - \frac{8n}{m}\right).$$

Now plugging this into the bound from (3.10) gives us

$$\mathbb{E}\big[\mathrm{ENT}(\mu\pi_t|\mathrm{sgn}(\mu\pi_t))\big] - \mathbb{E}\big[\mathrm{ENT}((\mu|\mathrm{sgn}(\mu)))\big] \leq \frac{-C_1 D_4}{C_2^2 L^2 \log(n)} \exp\left(-864L^2 - \frac{8n}{m}\right) \sum_{x \in I_{k^*}} E_x.$$

Note that $L^2 t < n^{\frac{3}{2}}$ for sufficiently large $n$. So for large $n$ we have

$$\mathbb{E}\big[\mathrm{ENT}(\mu\pi_t|\mathrm{sgn}(\mu\pi_t))\big] - \mathbb{E}\big[\mathrm{ENT}((\mu|\mathrm{sgn}(\mu)))\big] \leq \frac{-D_5 t}{n^{\frac{3}{2}}\log(n)\sqrt{n}} \exp\left(-864L^2 - \frac{8n}{m}\right) \sum_{x\in I_{k^*}} E_x.$$

Using again that $n^{\frac{3}{2}} \leq 4\ell_{\max}$ we get

$$\mathbb{E}\big[\mathrm{ENT}(\mu\pi_t|\mathrm{sgn}(\mu\pi_t))\big] - \mathbb{E}\big[\mathrm{ENT}((\mu|\mathrm{sgn}(\mu)))\big] \leq \frac{-D_6 t}{\ell_{\max}^2\log(n)\sqrt{n}} \exp\left(-864L^2 - \frac{8n}{m}\right) \sum_{x\in I_{k^*}} E_x.$$

(3) $k^* \geq \log_2(C_2 L)$

Set $t = \ell_{k^*}^2(2 + 2\cdot 10^{-6}) + 4\cdot 4n + 10n$. Then by Theorem 3.6.3 for each $A_x$ in the sum $\sum_{x\in I_{k^*}} A_x E_x$ we have

$$A_x = \mathbb{P}(E)\cdot\nu(x)^2 \geq \frac{D_4\gamma_{k^*}^2 n}{\ell_k^{*4}} \exp\left(-864L^2 - \frac{8n}{m}\right)\cdot\nu(x)^2.$$

Recall that for all $x \in I_{k^*}$ we have $\nu(x) > |J_{k^*-1}|$. As we did in (3.8) we can compute

$$|J_{k^*-1}| \geq \frac{2\ell_{k^*-1}^2}{\gamma_{k^*-1}\sqrt{n}}.$$

This gives us

$$A_x \geq D_4\left(\frac{\ell_{k^*-1}}{\ell_{k^*}}\right)^4\left(\frac{\gamma_{k^*}}{\gamma_{k^*-1}}\right)^2 \exp\left(-864L^2 - \frac{8n}{m}\right).$$

Note that $\gamma_{k^*} > \gamma_{k^*-1}$ and recall that $\ell_{k^*-1} = \frac{1}{2}\ell_{k^*}$. So we have

$$A_x \geq D_7 \exp\left(-864L^2 - \frac{8n}{m}\right).$$

Plugging this into (3.10) we get

$$\mathbb{E}\big[\mathrm{ENT}(\mu\pi_t|\mathrm{sgn}(\mu\pi_t))\big] - \mathbb{E}\big[\mathrm{ENT}((\mu|\mathrm{sgn}(\mu)))\big] \leq \frac{-C_1 D_7}{\log(n)} \exp\left(-864L^2 - \frac{8n}{m}\right) \sum_{x\in I_{k^*}} E_x.$$

Since $t$ is less than a constant times $\ell_{k^*}^2$ we have that

$$\mathbb{E}\big[\mathrm{ENT}(\mu\pi_t|\mathrm{sgn}(\mu\pi_t))\big] - \mathbb{E}\big[\mathrm{ENT}((\mu|\mathrm{sgn}(\mu)))\big] \leq \frac{-D_8 t}{\ell_{k^*}^2 \log(n)} \exp\left(-864L^2 - \frac{8n}{m}\right) \sum_{x \in I_{k^*}} E_x$$
$$\leq \frac{-D_9 t}{\ell_{\max}^2 \log(n)} \exp\left(-864L^2 - \frac{8n}{m}\right) \sum_{x \in I_{k^*}} E_x.$$

Set $D_{10}$ to be the minimum of $D_3, D_6, D_9$. Then we have the bound

$$\mathbb{E}\big[\mathrm{ENT}(\mu\pi_t|\mathrm{sgn}(\mu\pi_t))\big] - \mathbb{E}\big[\mathrm{ENT}((\mu|\mathrm{sgn}(\mu)))\big] \leq \frac{-D_{10} t}{\ell_{\max}^2 \log(n)} \exp\left(-864L^2 - \frac{8n}{m}\right) \sum_{x \in I_{k^*}} E_x$$

independent of the value of $k^*$. Recall from line (3.9) that

$$\frac{1}{a}\mathbb{E}[\mathrm{ENT}(\mu \mid \mathrm{sgn}(\mu))] \leq \sum_{x \in I_{k^*}} E_j.$$

So we have

$$\mathbb{E}\big[\mathrm{ENT}(\mu\pi_t|\mathrm{sgn}(\mu\pi_t))\big] \leq \left(1 - \frac{D_{10} t}{a\ell_{\max}^2 \log(n)} \exp\left(-864L^2 - \frac{8n}{m}\right)\right) \mathbb{E}[\mathrm{ENT}(\mu \mid \mathrm{sgn}(\mu))].$$

Since $a \leq \log_2(\ell_{\max}) + 1 \leq \log_2(2n) + 1 \leq C_3 \log(n)$ we have

$$\mathbb{E}\big[\mathrm{ENT}(\mu\pi_t|\mathrm{sgn}(\mu\pi_t))\big] \leq \left(1 - \frac{Dt}{\ell_{\max}^2 \log^2(n)} \exp\left(-864L^2 - \frac{8n}{m}\right)\right) \mathbb{E}[\mathrm{ENT}(\mu \mid \mathrm{sgn}(\mu))].$$

□

Now we are ready to find a bound on the mixing time for the overlapping cycles shuffle.

THEOREM 3.2.1. *The overlapping cycles shuffle has a mixing time which is at most*

$$\mathcal{A}\ell_{\max}^2 \log^3(n)\mathcal{L}$$

*where $\mathcal{A}$ is a universal constant and*

$$\mathcal{L} = \mathcal{L}\left(\frac{n}{m}\right) = \exp\left(-864L^2 - \frac{8n}{m}\right) = \exp\left(-864\left(\frac{192n\exp(\frac{10n}{m})}{(n-m)}\right)^2 - \frac{8n}{m}\right).$$

PROOF. The previous Lemma 3.6.4 implies that there exists $t_1 \in \{1, \ldots, C\ell_{\max}^2\}$ such that

$$\mathbb{E}\big[\mathrm{ENT}(\pi_{(t_1)}|\mathrm{sgn}(\pi_{(t_1)}))\big] \leq \left(1 - \frac{D\mathcal{L}t_1}{\log^2(n)\ell_{\max}^2}\right) \mathbb{E}\big[\mathrm{ENT}((\mathrm{id}|\mathrm{sgn}(\mathrm{id}))\big].$$

Choose such a $t_1$, and then, by the same theorem, there exists $t_2 \in \{1, \dots, C\ell_{\max}^2\}$ such that

$$\mathbb{E}\big[\mathrm{ENT}(\pi_{(t_2)}\pi_{(t_1)}|\mathrm{sgn}(\pi_{(t_2)}\pi_{(t_1)}))\big] \leq \left(1 - \frac{D\mathcal{L}t_2}{\log^2(n)\ell_{\max}^2}\right)\mathbb{E}\big[\mathrm{ENT}((\pi_{(t_1)}|\mathrm{sgn}(\pi_{(t_1)})\big].$$

Repeat this inductively, so choosing $t_k \in \{1, \dots, C\ell_{\max}^2\}$ such that

$$\mathbb{E}(\mathrm{ENT}(\pi_{(t_k)} \dots \pi_{(t_1)}|\mathrm{sgn}(\pi_{(t_k)} \dots \pi_{(t_1)}))) \leq \left(1 - \frac{D\mathcal{L}t_k}{\log^2(n)\ell_{\max}^2}\right)\mathbb{E}(\mathrm{ENT}((\pi_{(t_{k-1})} \dots \pi_{(t_1)}|\mathrm{sgn}(\pi_{(t_{k-1})} \dots \pi_{(t_1)})),$$

and therefore

$$\mathbb{E}(\mathrm{ENT}(\pi_{(t_k)} \dots \pi_{(t_1)}|\mathrm{sgn}(\pi_{(t_k)} \dots \pi_{(t_1)}))) \leq \prod_{i=1}^{k}\left(1 - \frac{D\mathcal{L}t_i}{\log^2(n)\ell_{\max}^2}\right)\mathbb{E}\big[\mathrm{ENT}((\mathrm{id}|\mathrm{sgn}(\mathrm{id}))\big].$$

Note that

$$\begin{aligned}\prod_{i=1}^{k}\left(1 - \frac{D\mathcal{L}t_i}{\log^2(n)\ell_{\max}^2}\right) &\leq \exp\left(-\sum_{i=1}^{j}\frac{D\mathcal{L}t_i}{\log^2(n)\ell_{\max}^2}\right)\mathbb{E}\big[\mathrm{ENT}(\mathrm{id} \mid \mathrm{sgn}(\mathrm{id}))\big] \\ &= \exp\left(\frac{-D\mathcal{L}}{\log^2(n)\ell_{\max}^2}\sum_{i=1}^{j}t_j\right)\mathbb{E}\big[\mathrm{ENT}(\mathrm{id} \mid \mathrm{sgn}(\mathrm{id}))\big].\end{aligned}$$

and

$$\mathbb{E}\big[\mathrm{ENT}(\mathrm{id} \mid \mathrm{sgn}(\mathrm{id}))\big] = \mathrm{ENT}(\mathrm{id} \mid \mathrm{sgn}(\mathrm{id})) = \log\left(\frac{n!}{2}\right) \leq n\log(n).$$

With this in mind, let

$$t = \frac{2}{D\mathcal{L}}\log^2(n)\ell_{\max}^2\big(\log(n) + \log(\log(n))\big)\log(-\epsilon) + C\ell_{\max}^2.$$

Since each $t_k$ is less than $C\ell_{\max}^2$, then there exists some $\kappa$ such that

$$\frac{2}{D\mathcal{L}}\log^2(n)\ell_{\max}^2\big(\log(n) + \log(\log(n))\big)\log(-\epsilon) < t_1 + \dots + t_\kappa < t.$$

So,

$$\begin{aligned}\mathbb{E}\big[\mathrm{ENT}(\pi_t|\mathrm{sgn}(\pi_t))\big] &\leq \mathbb{E}\big[\mathrm{ENT}(\pi_{(t_\kappa)}\dots\pi_{(t_1)}|\mathrm{sgn}(\pi_{(t_\kappa)}\dots\pi_{(t_1)}))\big] \\ &\leq \exp\left(\frac{-D\mathcal{L}}{\log^2(n)\ell_{\max}^2}\frac{2}{D\mathcal{L}}\log^2(n)\ell_{\max}^2\big(\log(n)+\log(\log(n))\big)\log(-\epsilon)\right)n\log(n) \\ &= \epsilon^2\end{aligned}$$

This is a bound on the conditional entropy given the sign of $\pi_t$. If $(1,\dots,m)$ and $(1,\dots,n)$ have the same sign, then this is the best we can hope for because we will always know if $\pi_t$ is even or odd by looking at if $t$ is even or odd. If $(1,\dots,m)$ and $(1,\dots,n)$ have different signs, then we can get a bound on the total entropy by doing a single additional step. Since the group element we multiply by in this additional step is equally likely to have an even or odd sign, we get

$$\mathbb{E}\big[\mathrm{ENT}(\pi_{(t+1)})\big] \leq \epsilon^2$$

Plugging this into (3.4) tells us that

$$\|\pi_{(t+1)} - \xi\|_{\mathrm{TV}} \leq \epsilon$$

and this gives us the mixing time. □

It should be noted that for any constant $\delta > 0$ the function $\mathcal{L}$ can be bounded from below on choices of $m$ where $\delta < \frac{m}{n} < 1-\delta$. So if we let $\alpha \in (\frac{1}{100}, \frac{99}{100})$ and consider the shuffles where $m = \lfloor \alpha n \rfloor$ we get that the mixing time is $O(\ell_{\max}^2 \log^3(n))$. This matches the mixing time shown by Angel, Peres, and Wilson for a single card after multiplying by the factor of $C\log^3(n)$. In the longest case, since we trivially have $\ell_{\max} \leq 2n$, we get that the mixing time is $O(n^2\log^3(n))$. This longest case is admitted if $\alpha$ is any rational, although the constant in front of $n^2\log^3(n)$ is smaller for rationals that have larger denominators in their reduced form.

In the shortest case, since $\ell_{\max} \geq \frac{1}{2}n^{\frac{3}{4}}$, we have a mixing time of $O(n^{\frac{3}{2}}\log^3(n))$. This shortest case is admitted when $\alpha = \phi$ where $\phi = \frac{\sqrt{5}-1}{2}$ is the inverse golden ratio. This is because by Corollary A.0.8 we have that $\ell_{\max}$ for $m = \lfloor \phi n \rfloor$ is a constant times $n^{\frac{3}{4}}$. This follows from multiples of $\phi$ being equally distributed across $(0,1) \mod 1$ which also means they are equally distributed across $(0, 2n-\phi n+1) \mod 2n-\phi n+1$. If a more thorough justification is required, consider the following:

Let $x \in \{1, \ldots, 2n - \lfloor \phi n \rfloor + 1\}$. As per Corollary A.0.8 choose some $\beta \in \{1, \ldots, \sqrt[4]{n}\}$ such that

$$\left| \frac{x}{2n} - (\beta\phi \mod 1) \right| \leq \frac{1}{2\phi^2} \cdot \frac{1}{\sqrt[4]{n}}.$$

where $(z \mod \mathcal{M})$ refers to the number $\zeta \in (0, \mathcal{M}]$ such that $z \equiv \zeta \mod \mathcal{M}$. Then,

$$|x - (2\beta\phi n \mod 2n)| \leq 2n \cdot \frac{1}{2\phi^2} \cdot \frac{1}{\sqrt[4]{n}} = \frac{1}{\phi^2} \cdot n^{\frac{3}{4}}.$$

Note that $\beta\phi n \leq \beta(2n)$. So,

$$(2\beta\phi n \mod 2n) = \kappa(2n) + 2\beta\phi n$$

where $|\kappa| \leq \beta$. This means that

$$(2\beta\phi n \mod 2n) = \kappa(2n - \phi n) + (2\beta + \kappa)\phi n.$$

So,

$$|x - \kappa(2n - \phi n) - (2\beta + \kappa)\phi n| \leq \frac{1}{\phi^2} \cdot n^{\frac{3}{4}}$$

$$|x - \kappa(2n - \lfloor \phi n \rfloor + 1) - (2\beta + \kappa)\phi n| \leq \frac{1}{\phi^2} \cdot n^{\frac{3}{4}} + 2|\kappa|$$

$$|x - \kappa(2n - \lfloor \phi n \rfloor + 1) - (2\beta + \kappa)\lfloor \phi n \rfloor| \leq \frac{1}{\phi^2} \cdot n^{\frac{3}{4}} + 2|\beta| + 3|\kappa|.$$

Let $b = 2\beta + \kappa$. Then,

$$|(x - b\lfloor \phi n \rfloor) \mod 2n - \lfloor \phi n \rfloor| \leq \frac{1}{\phi^2} \cdot n^{\frac{3}{4}} + 2|\beta| + 3|\kappa|$$

so

$$x = b\lfloor \phi n \rfloor + a \text{ where } |b| \leq 3\sqrt[4]{n} \text{ and } |a| \leq \frac{1}{\phi^2} \cdot n^{\frac{3}{4}} + 5\sqrt[4]{n}.$$

This means,

$$\begin{aligned}\|x\| &\leq 3\sqrt[4]{n} \cdot \sqrt{n} + \frac{1}{\phi^2} \cdot n^{\frac{3}{4}} + 5\sqrt[4]{n} \\ &\leq \left(3 + \frac{1}{\phi^2}\right) n^{\frac{3}{2}} + 5\sqrt[4]{n} \\ &\leq 6n^{\frac{3}{2}} \text{ for large enough } n.\end{aligned}$$

# APPENDIX A

Here we have included some theorems cited throughout the thesis whose uses are more generally applicable across probability.

THEOREM A.0.1. *Suppose $\mu$ is a probability measure on a finite probability space $\Omega$ such that for each $\omega \in \Omega$, we have $\mu(\omega) \geq \frac{1}{\mathcal{D}|\Omega|}$. Let $E$ be an event such that $\mu(E) \geq 1 - \frac{1}{8\mathcal{D}}$. Then there exists at least $\frac{3}{4}|\Omega|$ values $\alpha \in \Omega$ such that $\mu(\alpha|E) > \frac{1}{2\mathcal{D}|\Omega|}$.*

PROOF. Let $S \subset \Omega$ be the set of values $\beta \in \Omega$ such that $\mu(\beta|E) \leq \frac{1}{2\mathcal{D}|\Omega|}$. Then,

$$\mu(S, E) = \mu(E) \sum_{\beta \in S} \leq \frac{|S|}{2\mathcal{D}|\Omega|}.$$

On the other hand,

$$\mu(S, E) = \mu(S) - \mu(S, E^C) \geq \mu(S) - \mu(E^C) \geq \frac{|S|}{\mathcal{D}|\Omega|} - \frac{1}{8\mathcal{D}}.$$

So,

$$\begin{aligned} \frac{|S|}{\mathcal{D}|\Omega|} - \frac{1}{8\mathcal{D}} &\leq \frac{|S|}{2\mathcal{D}|\Omega|}, \\ \frac{|S|}{2\mathcal{D}|\Omega|} &\leq \frac{1}{8\mathcal{D}}, \\ |S| &\leq \frac{1}{4}|\Omega|. \end{aligned}$$

Since at most $\frac{1}{4}|\Omega|$ of $\beta \in \Omega$ have $\mu(\beta|E) \leq \frac{1}{2\mathcal{D}|\Omega|}$ we know that at least $\frac{3}{4}|\Omega|$ of $\alpha \in \Omega$ have $\mu(\alpha|E) > \frac{1}{2\mathcal{D}|\Omega|}$. □

We now provide a more general version of Theorem [A.0.1]. While we do not use the general version for any of our results, we provide it for the potential interest of the reader.

THEOREM A.0.2. *Let $\mu, \nu$ be probability measures on $\Omega$. Assume that there exist constants $a, b, \epsilon, \delta \in [0, 1]$ such that*

- $\mu(x) \geq \frac{a}{|\Omega|}$ *at least* $(1-\epsilon)|\Omega|$ *of* $x \in \Omega$
- $\nu(x) \leq \frac{b}{|\Omega|}$ *at least* $(1-\delta)|\Omega|$ *of* $x \in \Omega$

*Then,*

$$\|\mu - \nu\|_{\text{TV}} \geq (1-\epsilon)(1-\delta)(a-b)$$

PROOF. Let $S$ be the set of all $x \in \Omega$ such that $\mu(x) \geq \frac{a}{|\Omega|}$. Let $T$ be the set of all $x \in \Omega$ such that $\nu(x) \leq \frac{b}{|\Omega|}$. Then using one of the definitions of total variation distance we get,

$$\begin{aligned}
\|\mu - \nu\|_{\text{TV}} &= \sum_{x \in |\Omega|} (\mu(x) - \nu(x))^+ \\
&\geq \sum_{x \in S \cap T} (\mu(x) - \nu(x))^+ \\
&\geq \sum_{x \in S \cap T} \frac{a-b}{|\Omega|} \\
&= \frac{|S \cap T|}{|\Omega|}(a-b) \geq (1-\epsilon)(1-\delta)(a-b).
\end{aligned}$$

□

We can get Theorem [A.0.1] from Theorem [A.0.2] if we let $a = \frac{1}{\mathcal{D}}$ and $\epsilon = 0$. So $\mu$ is a probability measure where $\mu(x) \geq \frac{1}{\mathcal{D}|\Omega|}$ for all $x \in \Omega$. Then pick an event $E$ where $\mu(E) \geq 1 - \frac{1}{8\mathcal{D}}$ and let $\nu$ be the measure $\mu$ conditioned on $E$. Then $\|\mu - \nu\|_{\text{TV}} \leq \frac{1}{8\mathcal{D}}$. Then if we set $b = \frac{1}{2\mathcal{D}}$ and solve for $\delta$ we will find that $\delta \geq \frac{3}{4}$ which means that no more than $\frac{1}{4}$ of all $x \in \Omega$ can have $\nu(x) = \mu(x|E) \leq \frac{1}{2\mathcal{D}}$.

THEOREM A.0.3. *[7] (Section 7.3, page 46) Let $X$ be a binomial random variable with $n$ trials and probability $\frac{1}{2}$ chance of success. Let $k \geq 0$. Then,*

$$\mathbb{P}\left(X - \frac{n}{2} \geq k\right) \geq \frac{1}{15} \exp\left(\frac{-16k^2}{n}\right)$$

THEOREM A.0.4 (Hoeffding's inequality). *[5] (Section 2, page 15) Let $X$ be a binomial random variable with $n$ trials and probability $p$ of success.*

$$\mathbb{P}\left(X - np \geq k\right) \leq \exp\left(-\frac{2k^2}{n}\right)$$

COROLLARY A.0.5. *Let $X_t$ be the simple random walk on the integers. Then*

$$\mathbb{P}(|X_t| \geq a\sqrt{n}) \leq 2\exp\left(-\frac{a^2}{2}\right)$$

PROOF. After $t$ steps of the simple symmetric random walk, let $R$ be the amount of right steps. Then $X_t = 2R - t$, and $R$ is a Binomial random variable with $t$ trials and probability $\frac{1}{2}$ of success. So,

$$\begin{aligned}\mathbb{P}(X_t \geq a\sqrt{n}) &= \mathbb{P}\left(R \geq \frac{t}{2} + \frac{a}{2}\sqrt{n}\right) \\ &\leq \exp\left(-\frac{a^2}{2}\right).\end{aligned}$$

By symmetry, $-X_t$ has the same distribution. So by the union bound we get

$$\mathbb{P}(|X_t| > a\sqrt{n}) \leq 2\exp\left(-\frac{a^2}{2}\right).$$

□

THEOREM A.0.6. *Let $X_t$ be the simple random walk on the integers. Let*

$$A_t = \max\{|X_s| \ : \ s \leq t\}.$$

*Then*

$$\mathbb{P}(A_t > a\sqrt{n}) \leq 4\exp\left(-\frac{a^2}{2}\right).$$

PROOF. Let

$$M_t = \max\{X_s \ : \ s \leq t\}$$

Note that

$$\begin{aligned}\mathbb{P}(M_t \geq k) &= \mathbb{P}(\text{ there exists } s \leq t : X_s = k) \\ &= \mathbb{P}(\text{ there exists } s \leq t : X_s = k,\ X_t > k) \\ &\quad + \mathbb{P}(\text{ there exists } s \leq t : X_s = k,\ X_t < k) \\ &\quad + \mathbb{P}(\text{ there exists } s \leq t : X_s = k,\ X_t = k).\end{aligned}$$

By the Markov Property we see that

$$\mathbb{P}(\text{ there exists } s \leq t : X_s = k,\ X_t > k) = \mathbb{P}(\text{ there exists } s \leq t : X_s = k,\ X_t < k).$$

Also note that

$$\mathbb{P}(\text{ there exists } s \leq t : X_s = k,\ X_t > k) = \mathbb{P}(X_t > k)$$

$$\mathbb{P}(\text{ there exists } s \leq t : X_s = k,\ X_t = k) = \mathbb{P}(X_t = k)$$

so,

$$\begin{aligned}\mathbb{P}(M_t \geq k) &= 2\mathbb{P}(X_t > k) + \mathbb{P}(X_t = k) \\ &\leq 2\mathbb{P}(X_t \geq k).\end{aligned}$$

Setting $k = a\sqrt{n}$ and applying Hoeffding's inequality gives

$$\mathbb{P}(M_t \geq a\sqrt{n}) \leq 2\exp\left(-\frac{a^2}{2}\right).$$

Due to symmetry the minimum value of $X_s$ over the first $t$ steps has the same distribution as $-M_t$. So by the union bound

$$\mathbb{P}(A_t \geq a\sqrt{n}) \leq 4\exp\left(-\frac{a^2}{2}\right).$$

□

THEOREM A.0.7. *[12] Let $\phi = \frac{\sqrt{5}-1}{2}$ be the inverse golden ratio. Fix any $N \in \mathbb{N}$. Now we define $0 = a_0 < a_1 < \cdots < a_N < 1$ as the numbers where each $a_i = k\phi \mod 1$ for some natural number $k \leq N$. In other words, $a_0, \ldots, a_N$ is a reordering of $0, \phi, 2\phi, \ldots, N\phi \mod 1$ from least to greatest. Then for any $i \in \{1, \ldots, N\}$ we have*

$$a_i - a_{i-1} \in \{\phi^z, \phi^{z+1}, \phi^{z+2}\}$$

$$1 - a_N \in \{\phi^z, \phi^{z+1}, \phi^{z+2}\}$$

*where $z$ is defined as follows: $F_z$, the $z$th Fibbonacci number, is the largest Fibbonacci number less than or equal to $N$ (using the convention that $F_1 = F_2 = 1$). For a more numerical definition we*

*can also write*

$$z = \max\left\{x \in \mathbb{N} \text{ such that } \frac{\phi^{-x} - (-\phi)^x}{\sqrt{5}} \leq N\right\}.$$

COROLLARY A.0.8. *Fix any $N \in \mathbb{N}$. Then for any $x \in [0,1]$ there exists $k \in \{0, 1, \ldots, N\}$ such that*

$$|x - b_k| \leq \frac{1}{2\phi^2} \cdot \frac{1}{N+1}$$

*for some $b_k \in [0,1]$ with $b_k \equiv k\phi \mod 1$.*

PROOF. Let $a_0, \ldots, a_N$ be a reordering of $b_0, \ldots, b_N$ in increasing order. I.e let $\{a_0, \ldots, a_N\} = \{b_0, \ldots, b_N\}$ where $a_0 < a_1 < \cdots < a_N$. Let $g$ be the smallest gap between adjacent elements of $(a_0, \ldots, a_N, 1)$. Note that $g \leq \frac{1}{N+1}$ by the pigeon hold principle. By Theorem [A.0.7] we know that all gaps take the form $\phi^z, \phi^{z+1}, \phi^{z+2}$ for a particular $z$. Using the fact that $\frac{1}{N+1} \geq g \geq \phi^{z+2}$ we see that

$$z \leq -\log_\phi(N+1) - 2$$

Let $G$ be the largest gap between adjacent elements of $\{a_0, \ldots, a_N, 1\}$. Then $G \leq \phi^z$ so

$$G \leq \frac{1}{(N+1)\phi^2}$$

In the furthest case $x \in [0,1]$ is in the middle of a gap, in which case $x$ is at most distance $\frac{1}{2\phi^2} \cdot \frac{1}{N+1}$ from an element of $\{b_0, \ldots, b_N\}$. □

# Bibliography


[1] O. Angel, Y. Peres, and D. B. Wilson, *Card shuffling and diophantine approximation*, The Annals of Applied Probability, 18 (2008), pp. 1215–1231.

[2] W. Dai, V. T. Hoang, and S. Tessaro, *Information-theoretic indistinguishability via the chi-squared method*, in Advances in Cryptology – CRYPTO 2017, J. Katz and H. Shacham, eds., Cham, 2017, Springer International Publishing, pp. 497–523.

[3] M. V. Hildebrand, *Rates of convergence of some random processes on finite groups*, PhD thesis, Harvard University, 1990.

[4] V. T. Hoang, B. Morris, and P. Rogaway, *An enciphering scheme based on a card shuffle*, in Advances in Cryptology – CRYPTO 2012, R. Safavi-Naini and R. Canetti, eds., Berlin, Heidelberg, 2012, Springer Berlin Heidelberg, pp. 1–13.

[5] W. Hoeffding, *Probability inequalities for sums of bounded random variables*, Journal of the American Statistical Association, 58 (1963), pp. 13–30.

[6] J. Jonasson, *Biased random-to-top shuffling*, The Annals of Applied Probability, 16 (2006), pp. 1034–1058.

[7] J. Matoušek and J. Vondrák, *The Probabilistic Method*, Lecture Notes, Department of Applied Mathematics, Charles University, Prague, (2001).

[8] U. Maurer, K. Pietrzak, and R. Renner, *Indistinguishability amplification*, in Advances in Cryptology - CRYPTO 2007, A. Menezes, ed., Berlin, Heidelberg, 2007, Springer Berlin Heidelberg, pp. 130–149.

[9] B. Morris, *Improved mixing time bounds for the Thorp shuffle and L-reversal chain*, The Annals of Probability, 37 (2009), pp. 453–477.

[10] A. E. Senda, *A Mixing Time Bound for the Diaconis Shuffle*, PhD thesis, University of California Davis, 2022.

[11] A. Tsybakov, *Introduction to Nonparametric Estimation*, Springer Series in Statistics, Springer New York, 2008.

[12] S. Świerczkowski, *On successive settings of an arc on the circumference of a circle*, Fundamenta Mathematicae, 46 (1958), pp. 187–189.