Università degli studi di Padova

DIPARTIMENTO DI MATEMATICA "TULLIO LEVI-CIVITA"
(Dipartimento di Ingegneria dell'Informazione)

Corso di Laurea Magistrale in Matematica

Approximate probability distributions for long hash chains in the random oracle model

Candidate:
Stefano Mattiello, 1237283

Thesis advisor:
Prof. Nicola Laurenti

21 July 2022

Contents

Introduction

The aim of the thesis is to improve the probabilistic model used for one-way hash chains in *Evaluating the Security of One-way key Chains in TESLA-based GNSS Navigation Message Authentication Schemes* by Caparra, Stuaro, Laurenti and Wullems [4].

In the first chapter we introduce some fundamental notions and useful results that will be used throughout the thesis. Then we give a description of what hash chains are and which are their applications. In particular, we focus on the TESLA protocol, in order to have a better understanding of the model we will be working on.

In the second chapter we define a probabilistic model for hash chain following the one presented in [4] and then we proceed to study its properties. Then we analyze the attack described in the above-mentioned article and derive general upper and lower bounds on the probability of success of the attack, by relaxing an assumption made in the paper which proves unnecessary.

In the final chapter we aim to find a probability distribution that well approximate our model, at least asymptotically. To do so we aim to give an upper bound on the Kullback–Leibler Divergence between the model and our target distribution to later derive an upper bound on the total variational distance through Pinsker inequality. Unfortunately we could not complete the task and we conclude the thesis with an intermediate result about these bounds.

Chapter 1

Preliminaries

1.1 Basic definitions

Definition 1.1. *A* random oracle *is a theoretical black box that responds to every query with a random response chosen uniformly from its output domain. In particular, a random oracle is a function* $f\colon X \to Y$*, where* $f(x) \sim \mathcal{U}(Y)$ *for all* $x \in X$ *with* $f(x_1), \ldots, f(x_n)$ *independent for distinct* $x_1, \ldots, x_n \in X$*.*

Definition 1.2. *A* discrete time martingale *is a discrete time sequence of random variables* $(X_i)_{i\in\mathbb{N}}$ *that satisfies*

$$E[|X_n|] < \infty \tag{1.1}$$
$$E[X_n|X_1, \ldots, X_{n-1}] = X_{n-1} \tag{1.2}$$

for all $n \in \mathbb{N}$*, i.e. the conditional expected value of the next observation, given all the past observations, is equal to the most recent observation.*

In particular if $(X_i)_{i\in\mathbb{N}}$ is a finite state Markov chain we have that condition (1.1) is satisfied and condition (1.2) become

$$E[X_n|X_1, \ldots, X_{n-1}] = E[X_n|X_{n-1}] = X_{n-1}$$

Definition 1.3. *The* Stirling numbers of the second kind $\left\{ {n \atop k} \right\}$ *are the number of ways to partition a set of* n *elements into* k *non-empty subsets. They can be computed as*

$$\left\{ {n \atop k} \right\} = \frac{1}{k!} \sum_{i=0}^{k} (-1)^i \binom{k}{i} (k-i)^n$$

They can also be characterized as the numbers that arise when expressinng powers of some $x \in \mathbb{N}$ *in terms of the falling factorials*

$$\sum_{k=0}^{n} \left\{ {n \atop k} \right\} x^{\underline{k}} = x^n \tag{1.3}$$

where $x^{\underline{k}} = \frac{x!}{(n-k)!}$.

Lemma 1.4. *Let* X *be a* (n, p) *binomial random variable. Then the* d^{th} *moment of* X *is*

$$E[X^d] = \sum_{k=1}^{d} \left\{ {d \atop k} \right\} n^{\underline{k}} p^k$$

A complete proof of the previous statement can be found in [8].

Lemma 1.5. *Let* $\{a_\ell\}_{\ell\in\mathbb{N}}, \{g_\ell\}_{\ell\in\mathbb{N}}$ *be two sequences in* $\mathbb{R}$ *and* $b \in \mathbb{R} \setminus \{0\}$. *If the recursive relation* $a_{\ell+1} = ba_\ell + g_\ell$ *holds, then*

$$a_\ell = b^\ell \Big(a_0 + \sum_{m=0}^{\ell-1} \frac{g_m}{b^{m+1}} \Big)$$

Proof. We have

$$\frac{a_{\ell+1}}{b^{\ell+1}} - \frac{ba_\ell}{b^{\ell+1}} = \frac{g_\ell}{b^{\ell+1}}$$

Let $A_\ell = \frac{a_\ell}{b^\ell}$. Then

$$\begin{aligned} A_{\ell+1} - A_\ell &= \frac{g_\ell}{b^{\ell+1}} \\ \sum_{m=0}^{\ell-1} (A_{m+1} - A_m) = A_\ell - A_0 &= \sum_{m=0}^{\ell-1} \frac{g_m}{b^{m+1}} \\ \frac{a_\ell}{b^\ell} &= A_0 + \sum_{m=0}^{\ell-1} \frac{g_m}{b^{m+1}} \\ a_\ell &= b^\ell \Big(a_0 + \sum_{m=0}^{\ell-1} \frac{g_m}{b^{m+1}} \Big) \end{aligned}$$

□

1.2 Problem description

The aim of the thesis is to improve the probabilistic model used for one-way hash chains in [4]. We will now give some contexual information about what a hash chain is and some of its applications.

One-way hash chains are build on one-way hash functions. A one-way hash function $H\colon \mathcal{X}^* \to \mathcal{Y}^\rho$ maps an input of any length to a fixed-length bit string. The function H should be simple to compute yet must be computationally infeasible to invert in general.

Definition 1.6. *Let $H\colon \mathcal{X}^* \to \mathcal{X}^\rho$ be a hash function with the same input and output alphabet and $h_0 \in \mathcal{X}^\rho$. A* time invariant one-way hash chain *is the list of values*

$$h_1, h_2, h_3, \ldots, h_n$$

where $h_i = H(h_{i-1})$ for $1 \leq i \leq n$.

Given an element h_i of a one-way hash chain, it is possible to verify that a previous element h_{i-j} belong to the sequence simply by computing $H^j(h_{i-j})$. On the other hand it should be difficult to get previous elements of the chain, due to the properties of the hash function H.

An application of time invariant one-way hash chains is given by Hu, Johnson and Perrig in [6], where they design the SEAD (Secure Efficient Ad hoc Distance vector) routing protocol. In this protocol the hash chains are used in routing updates to authenticate the distance to the destination node for each entry of the routing table and to prevent attackers from generating false values to mislead others nodes.

In long time invariant one-way hash chains loops could arise when exixts some $x \in \mathcal{X}$ such that $H(x) = H^j(x)$ for some $j \in \mathbb{N}$. This is a problem since when a loop is detected it is possible to get previous elements of the chain simply by repeating the elements of the loop in the reverse order over and over. To avoid the presence of loops and make the chain more robust it is possible to define time varying hash chains.

Definition 1.7. *Let $H\colon \mathcal{X}^* \times \mathbb{N} \to \mathcal{X}^\rho$ be a hash function and $h_0 \in \mathcal{X}^\rho$. A* time varying one-way hash chain *is the list of values*

$$h_1, h_2, h_3, \ldots, h_n$$

where $h_i = H(h_{i-1}, i)$ for $1 \leq i \leq n$.

When dealing with time varying hash chain we will denote for simplicity the hash function by H_i, with $H_i(k) \coloneqq H(k, i)$.

The application we will consider in this thesis is the TESLA (Timed Efficient Stream Loss-Tolerant Authentication) protocol. The TESLA protocol provides data authentication for messages sequences without relying on computing the digital signature of every message. A complete description of the TESLA protocol can be found in [10]; we will give here a simplified overview of the protocol that will be useful in our discussion.

First the sender creates a generator key k_0 used to develop a sequence of keys $k_0, k_1, \ldots, k_L$ using a time varying one-way hash function H_i, where $k_i = H_i(k_{i-1})$. Then the final key k_L can be securely delivered to the receiver. We assume that in this passage the receiver can verify that the key comes from the true source. For the following messages the sender can use the previous keys of the chain to authenticate the source of the messages. Indeed, due to the properties of the hash function an attacker cannot generate previous

value of the chain and, on the other hand, it is easy for the receiver to verify that $H_i(k_{i-1}) = k_i$, which prove that the sender is the original source.

Given this context we have to make sure that the attacker cannot find easily some key k such that $(H_{L-1} \circ \cdots \circ H_{i+1} \circ H_i)(k) = k_L$.

Since in the TESLA protocol the function H_i is a pseudo-random function and therefore in general not injective, for long sequences we could have that the image of large portions of the domain are sets of only few elements. This constitues a problem because it means that, given the key k_L, there is a significant probability that random element b in the domain satisfies $(H_{L-1} \circ \cdots \circ H_{i+1} \circ H_i)(k) = k_L$.

This unwanted behaviour make the hash chain vulnerable to various attack, since one of the assumed properties of the hash chain is that it is difficult to get previous elements of the chain. However, numerical experimentation show that this phenomenom is negligible if the length of the hash chain is small compared to the cardinality of the domain of the hash function.

Chapter 2

Probabilistic model

2.1 Definition

In this section we will define the probabilistic model for the hash chain that we will use throughout the thesis.

The time varying hash function will be represented as a random oracle $f_i \colon \mathcal{N} \to \mathcal{N}$, where $|\mathcal{N}| = N \gg 1$, i.e. $f_i(a)$, $a \in \mathcal{N}$, $i = 1, \ldots, L-1$ are independent random variables uniformly distributed in $\mathcal{N}$. For each $b \in \mathcal{N}$, $i = 0, \ldots, L-1$, let us denote by $J_{i,j}(b) = |[f_i^j]^{-1}(\{b\})|$ the cardinality of the preimage of b under $f_i^j = f_i \circ f_{i+1} \circ \cdots \circ f_{j-1}$. By extension we consider f_i^i as the identity function on $\mathcal{N}$ and $J_{i,i}(b) = 1$ for all b.

We will denote by $\overline{\boldsymbol{J}}_{i,\ell}$ the vector $[J_{i,\ell}(b_1), \ldots, J_{i,\ell}(b_N)]$, where $b_1, \ldots, b_N$ is the complete list of the elements of $\mathcal{N}$.

Then, by denoting $\underline{1} = [1, \ldots, 1]$ the random vector

$$\overline{\boldsymbol{J}}_{i,i+1} = [J_{i,i+1}(b_1), \ldots, J_{i,i+1}(b_N)]$$

has the $(N, \frac{1}{N})$ multinomial distribution

$$p_{\boldsymbol{J}_{i,i+1}}(\boldsymbol{c}) = \begin{cases} 0 & \text{for } \sum_{b \in \mathcal{N}} c_b \neq N \\ \frac{N!}{\prod_{b \in \mathcal{N}} c_b!} \frac{1}{N^N} & \text{for } \sum_{b \in \mathcal{N}} c_b = N \end{cases} \tag{2.1}$$

and the marginal distribution of each $J_{i,i+1}(b)$ is a $(N, \frac{1}{N})$ binomial

$$p_{J_{i,i+1}}(c) = \binom{N}{c} \frac{1}{N^c} \left(1 - \frac{1}{N}\right)^{N-c} \tag{2.2}$$

On the other hand we can write $J_{i,j}(b)$ recursively as

$$J_{i,j}(b) = \sum_{a \in [f_{i+1}^j]^{-1}(\{b\})} J_{i,i+1}(a) \tag{2.3}$$

and since $J_{i,i+1}(a)$ has a $(N, \frac{1}{N})$ binomial distribution for all $a \in \mathcal{N}$ we can say that, conditioned on the cardinality $|[f_{i+1}^{j}]^{-1}(\{b\})| = J_{i+1,j}(b)$, the distribution of $J_{i,j}(b)$ is a $(N, \frac{J_{i+1,j}(b)}{N})$ binomial.
In the following section we will show that $J_{i,j}(b)$ is time homogeneous with respect to j. This fact allows us to omit the initial time i and write the probability of $J_\ell(b)$ conditioned to $J_{\ell-1}(b)$ as

$$p_{J_\ell}(c) = \binom{N}{c} \frac{J_{\ell-1}(b)}{N^c} \left(1 - \frac{J_{\ell-1}(b)}{N}\right)^{N-c} \tag{2.4}$$

It is therefore natural consider $J_\ell(b)$ as a discrete time Markov chain with binomial transition probability, i.e.

$$\begin{cases} J_0(b) = 1 \quad \forall\, b \in \mathcal{N} \\ J_j(b)|J_{j-1}(b) \sim \mathcal{B}(N, \frac{J_{j-1}(b)}{N}) \end{cases} \tag{2.5}$$

In the same way we can write the vector $\overline{\boldsymbol{J}}_\ell$ as a Markov chain defined as

$$\begin{cases} \overline{\boldsymbol{J}}_0 = (1, \dots, 1) \\ \overline{\boldsymbol{J}}_j|\overline{\boldsymbol{J}}_{j-1} \sim \mathcal{M}(N, \frac{1}{N}\overline{\boldsymbol{J}}_{j-1}) \end{cases} \tag{2.6}$$

where $\mathcal{M}(N, \frac{\overline{\boldsymbol{J}}_{j-1}}{N})$ is the multinomial distribution.

However we are not interested in studying the distribution of the preimage of the entire set $\mathcal{N}$, but instead we will only consider a small subset $\{b_1, \dots, b_m\}$ of it. Then we can write a similar random vector $\boldsymbol{J_j} = [J_j(b_1), \dots, J_j(b_m)]$ that consists of the first m entries of $\overline{\boldsymbol{J}}_j$. The corresponding Markov hain will be then

$$\begin{cases} \boldsymbol{J}_0 = (1, \dots, 1) \\ p_{\boldsymbol{J}_j|\boldsymbol{J}_{j-1}}(\boldsymbol{a}|\boldsymbol{b}) = \frac{N!}{a_1! \dots a_m!(N-a)!} (\frac{b_1}{N})^{a_1} \cdots (\frac{b_m}{N})^{a_m} (1 - \frac{b}{N})^{N-a} \end{cases} \tag{2.7}$$

where $a = \sum_{i=1}^{m} a_i$ and $b = \sum_{i=1}^{m} b_i$.
To conclude we want to consider the random variable $\|\boldsymbol{J_j}\|_1 = \sum_{i=1}^{m} J_j(b_i)$. Since $J_j(b_i)|J_{j-1}(b_i)$ has a $(N, \frac{J_{j-1}(b)}{N})$ binomial distribution, their sum will have a $(N, \frac{\sum_{i=1}^{m} J_j(b_i)}{N})$ binomial distribution, therefore we can write $\|\boldsymbol{J_\ell}\|_1$ as another Markov chain given by

$$\begin{cases} \|\boldsymbol{J}_0\|_1 = m \\ \|\boldsymbol{J}_j\|_1 | \|\boldsymbol{J}_{j-1}\|_1 \sim \mathcal{B}(N, \frac{\|\boldsymbol{J}_{j-1}\|_1}{N}) \end{cases} \tag{2.8}$$

We can see that $\|\boldsymbol{J}_\ell\|_1$ and J_ℓ are the same Markov chain with different starting point, so for simplicity we will use J_ℓ to indicate both $|[f_i^j]^{-1}(\{b\})|$ and $\|\boldsymbol{J}_\ell\|_1$.

2.2 Properties

In this section we will investigate the basic properties of some of the Markov chain previously defined, in particular of J_ℓ.

2.2.1 Time homogeneity and absorbing states

The chain $J_{i,i+\ell}$ with i fixed is time-homegeneous with respect to ℓ since $J_{i,i+\ell} = |[f_i^\ell]^{-1}(\{b\})|$ for some $b \in \mathcal{N}$ and we assumed f_j mutually independent for all $j \in \mathbb{N}$.

On the other hand when ℓ is fixed $J_{i,i+\ell}$ is stationary with respect to i for the same reason. Then from now on we will omit i and write only the index ℓ.

Moreover, since the initial state J_0 is equal to 1 the chain is non-stationary. We observe that the two initial states that make the chain stationary are 0 and N. In particular 0 and N are the only two absorbing state of the chain while $1, \dots, N-1$ are all transient. This is easily derived from the transition probability of J_ℓ.

We can make similar observations for $\boldsymbol{J}_\ell$. The absorbing states of this chain are the vectors $[0, \dots, 0]$ and $[x_1, \dots, x_m]$ with $x_{\bar{\imath}} = N$,$x_i = 0$ for all $1 \le i \le m, i \ne \bar{\imath}$.

To visualize what these states represent in our model, we recall that $\boldsymbol{J}_\ell = [J_\ell(b_1), \dots, J_\ell(b_m)]$ is the vector of the cardinality of the preimage of some $b_1, \dots, b_m \in \mathcal{N}$ through $f_i \circ f_{i+1} \circ \cdots \circ f_{\ell-1}$. Therefore the zero vector represents the case in which $b_1, \dots, b_m$ have an empty preimage, while $[x_1, \dots, x_m]$ with $x_{\bar{\imath}} = N$,$x_i = 0$ for all $1 \le i \le m, i \ne \bar{\imath}$ the case in which the preimage of $b_{\bar{\imath}}$ is the entire set $\mathcal{N}$.

2.2.2 Martingale

It is easy to notice that all the Markov chains defined in section 2.1 are Martingales, i.e. we have that, using for example J_ℓ,

$$E[J_\ell | J_{\ell-1}] = J_{\ell-1} \tag{2.9}$$

This follows from the fact that the distribution of J_ℓ given $J_{\ell-1}$ is a $(N, \frac{J_{\ell-1}}{N})$ binomial distribution, from which we can derive the expected value. Then, if $J_1 = m$, $m < N$ we can write that

$$E[J_\ell | J_1 = m] = m \quad \forall \ell \in \mathbb{N} \tag{2.10}$$

Similar results can be stated for $\boldsymbol{J}_\ell$ and $\overline{\boldsymbol{J}}_\ell$ with the same argument.

2.2.3 Moments

We will now aim to get the momets of the random variable J_ℓ.

To find the higher moments of J_ℓ we can use a recursive formulation. Indeed, since $J_{\ell+1}|J_\ell$ has a binomial distribution, we have

$$E[J_{\ell+1}^d|J_\ell] = \sum_{k=1}^{d} \left\{ {d \atop k} \right\} N^{\underline{k}} \left(\frac{J_\ell}{N}\right)^k \tag{2.11}$$

where $N^{\underline{k}} = N(N-1)\cdots(N-k+1)$.

Then taking the expectation on J_ℓ we obtain

$$E[E[J_{\ell+1}^d|J_\ell]] = E[J_{\ell+1}^d] = \sum_{k=1}^{d} \left\{ {d \atop k} \right\} \frac{N^{\underline{k}}}{N^k} E[J_\ell^k] \tag{2.12}$$

We have now a formula for the generic d-moment of J_ℓ that depends on the moments of its previous steps $J_{\ell-1}, \ldots, J_0$. In order to find a closed formula let's define

$$c_i \coloneqq \frac{N^{\underline{i}}}{N^i} = \frac{N(N-1)\cdots(N-i+1)}{N^i} \tag{2.13}$$

We are now ready to give a generic formula for the d-moment of J_ℓ

Proposition 2.1. *Let J_ℓ be the previously defined Markov chain. Then*

$$E[J_\ell^d] = \sum_{i=1}^{d} a_{i,d} c_i^\ell \tag{2.14}$$

where $a_{i,j}$ is the coefficent of c_i^ℓ of the j-moment of J_ℓ and in particular

$$a_{i,d} = -\sum_{k=1}^{d-1} \left\{ {d \atop k} \right\} c_k \frac{a_{i,k}}{c_d - c_i} \quad \text{for } i \leq d-1 \tag{2.15}$$

$$a_{d,d} = m^d - \sum_{i=i}^{d-1} a_{i,d} \tag{2.16}$$

Proof. Let's prove it by induction on ℓ. For $\ell = 0$ we have $E[J_0^d] = m^d$, while our formula give us

$$E[J_0^d] = \sum_{i=1}^{d} a_{i,d} = \sum_{i=1}^{d-1} a_{i,d} + m^d - \sum_{i=i}^{d-1} a_{i,d} = m^d \tag{2.17}$$

Assume $E[J_j^d] = \sum_{i=0}^{d} a_{i,d} c_i^j$ for all $d \in \mathbb{N}$, $j < \ell$. We know that

$$E[J_\ell^d] = c_d E[J_{\ell-1}^d] + \sum_{k=1}^{d-1} \left\{ {d \atop k} \right\} c_k E[J_{\ell-1}^k] \tag{2.18}$$

We can apply the result of (1.5) using $a_\ell = J_\ell^d$, $b = c_d$ and $g_\ell = \sum_{k=1}^{d-1} \left\{ {d \atop k} \right\} c_k E[J_{\ell-1}^k]$ and obtain

$$E[J_\ell^d] = c_d^\ell \Big(m^d + \sum_{m=0}^{\ell-1} \frac{\sum_{k=1}^{d-1} \left\{ {d \atop k} \right\} c_k E[J_m^k]}{c_d^{m+1}} \Big) \tag{2.19}$$

$$= c_d^\ell \Big(m^d + \sum_{m=0}^{\ell-1} \sum_{k=1}^{d-1} \left\{ {d \atop k} \right\} \frac{c_k}{c_d^{m+1}} E[J_m^k] \Big) \tag{2.20}$$

$$= c_d^\ell \Big(m^d + \sum_{k=1}^{d-1} \sum_{m=0}^{\ell-1} \left\{ {d \atop k} \right\} \frac{c_k}{c_d^{m+1}} \sum_{i=1}^{k} a_{i,k} c_i^m \Big) \tag{2.21}$$

$$= c_d^\ell m^d + c_d^{\ell-1} \sum_{k=1}^{d-1} \sum_{m=0}^{\ell-1} \left\{ {d \atop k} \right\} c_k \sum_{i=1}^{k} a_{i,k} \Big(\frac{c_i}{c_d} \Big)^m \tag{2.22}$$

$$= c_d^\ell m^d + c_d^{\ell-1} \sum_{k=1}^{d-1} \left\{ {d \atop k} \right\} c_k \sum_{i=1}^{k} a_{i,k} \sum_{m=0}^{\ell-1} \Big(\frac{c_i}{c_d} \Big)^m \tag{2.23}$$

$$= c_d^\ell m^d + c_d^{\ell-1} \sum_{k=1}^{d-1} \left\{ {d \atop k} \right\} c_k \sum_{i=1}^{k} a_{i,k} \frac{1}{c_d^{\ell-1}} \frac{c_d^\ell - c_i^\ell}{c_d - c_i} \tag{2.24}$$

$$= c_d^\ell m^d + \sum_{k=1}^{d-1} \left\{ {d \atop k} \right\} c_k \sum_{i=1}^{k} a_{i,k} \frac{c_d^\ell - c_i^\ell}{c_d - c_i} \tag{2.25}$$

$$= c_d^\ell m^d + \sum_{i=1}^{d-1} \sum_{k=i}^{d-1} \left\{ {d \atop k} \right\} c_k \frac{a_{i,k}}{c_d - c_i} (c_d^\ell - c_i^\ell) \tag{2.26}$$

Now, if we define

$$a_{i,d} = -\sum_{k=i}^{d-1} \left\{ {d \atop k} \right\} c_k \frac{a_{i,k}}{c_d - c_i} \quad \text{for } i \le d-1 \tag{2.27}$$

$$a_{d,d} = m^d - \sum_{i=1}^{d-1} a_{i,d} \tag{2.28}$$

we can then conclude

$$E[J_\ell^d] = \sum_{i=1}^{d} a_{i,d} c_i^\ell \tag{2.29}$$

□

We write down the coefficent of the first four moments of J_ℓ

$a_{1,2} = mN \qquad a_{2,2} = m^2 - a_{1,2}$

$a_{1,3} = a_{1,2}N \quad a_{2,3} = \frac{3}{2}a_{2,2}N \qquad a_{3,3} = m^3 - a_{2,3} - a_{1,3}$

$a_{1,4} = a_{1,3}N \quad a_{2,4} = \frac{2}{3}(\frac{9N-11}{5N-6})a_{2,3}N \quad a_{3,4} = 2a_{3,3}N \qquad a_{4,4} = m^4 - \sum_{i=1}^{2} a_{i,4}$

We now prove two simple propositions that give a simpler formulation for some of these coefficents.

Proposition 2.2. *In equation* (2.14)*, we have* $a_{1,j} = mN^{j-1}$ *for all* $j \in \mathbb{N}$.

Proof. Assume $a_{1,j} = mN^{j-1}$ for all $j \leq d$. Then

$$a_{1,d+1} = -\sum_{k=1}^{d} \left\{ {d+1 \atop k} \right\} c_k \frac{a_{1,k}}{c_{d+1} - c_1} \tag{2.30}$$

$$= -\sum_{k=1}^{d} \left\{ {d+1 \atop k} \right\} \frac{N^{\underline{k}}}{N^k} \frac{mN^{k-1}}{c_{d+1} - 1} \tag{2.31}$$

$$= -m\frac{1}{c_{d+1} - 1} \sum_{k=1}^{d} \left\{ {d+1 \atop k} \right\} \frac{N^{\underline{k}}}{N} \tag{2.32}$$

Now, since

$$-\frac{1}{c_{d+1} - 1} = \frac{N^{d+1}}{N^{d+1} - N^{\underline{d+1}}}$$

and by the charaterization of the Stirling numbers of the second kind

$$\sum_{k=1}^{d+1} \left\{ {d+1 \atop k} \right\} N^{\underline{k}} = N^{d+1} \tag{2.33}$$

we have

$$\sum_{k=1}^{d+1} \left\{ {d+1 \atop k} \right\} \frac{N^{\underline{k}}}{N} = N^d \tag{2.34}$$

$$\frac{N^{\underline{d+1}}}{N} + \sum_{k=1}^{d} \left\{ {d+1 \atop k} \right\} \frac{N^{\underline{k}}}{N} = N^d \tag{2.35}$$

$$\sum_{k=1}^{d} \left\{ {d+1 \atop k} \right\} \frac{N^{\underline{k}}}{N} = \frac{N^{d+1} - N^{\underline{d+1}}}{N} \tag{2.36}$$

Therefore we can conclude that

$$a_{1,d+1} = mN^d \tag{2.37}$$

□

Proposition 2.3. *In equation* (2.14)*, we have*

$$a_{d,d+1} = \left\{ {d+1 \atop d} \right\} \frac{N}{d} a_{d,d} \tag{2.38}$$

for all $d \in \mathbb{N}$

We have indeed

$$a_{d,d+1} = -\left\{ {d+1 \atop d} \right\} c_d \frac{a_{d,d}}{c_{d+1} - c_d} \tag{2.39}$$

$$= -\left\{ {d+1 \atop d} \right\} a_{d,d} \frac{N^{\underline{d}}}{N^d} \frac{N^{d+1}}{N^{\underline{d+1}} - N N^{\underline{d}}} \tag{2.40}$$

$$= \left\{ {d+1 \atop d} \right\} \frac{N}{d} a_{d,d} \tag{2.41}$$

2.3 Attack model

Using the results of the previous section we can evaluate the probability of success of the attack described in section V of [4]. In the article they assume that J_l has a compound Poisson distribution; we will now evaluate this probability without thia assumption.

We consider an attack that aims to find $\hat{k}_{i+1}, \dots, \hat{k}_{i+\ell}$ such that $f_i(\hat{k}_{i+1}) = k_i$, $f_{i+1}(\hat{k}_{i+2}) = \hat{k}_{i+1}, \dots, f_{i+\ell-1}(\hat{k}_{i+\ell}) = \hat{k}_{i+\ell-1}$, where k_i is the last disclosed key of the chain,. To do so the attacker picks N_A random and independent values in $\mathcal{N}$ as guesses for $k_{i+\ell}$,

$$\hat{k}^j_{i+\ell} \sim \mathcal{U}(\mathcal{N}), \quad j = 1, \dots, N_A$$

and recursively apply the hash function up to

$$\hat{k}^j_i = f_i^{i+\ell}(\hat{k}^j_{i+\ell}), \quad j = 1, \dots, N_A$$

For each guess the attacker checks whether $\hat{k}^j_i$ coincides with the actual k_i, and in case of success he can use the chain $\hat{k}_{i+1}, \dots, \hat{k}_{i+\ell}$ to make the victim accept his messages as authentic.

The success event of this attack can be written as

$$\mathbb{S}(i, \ell, N_A) = \bigcup_{j=1}^{N_A} S_j(i, \ell) \tag{2.42}$$

where the success event for a single guess i is

$$S_j(i, \ell) = \{\hat{k}^j_i = k_i\} \tag{2.43}$$

Observe that, conditioned on the realization $f_i^{i+\ell}(\mathcal{N})$, the $\hat{k}_{i+\ell}^j$ are i.i.d. with probability mass distribution

$$P[\hat{k}_i^j = b | f_i^{i+\ell}] = \frac{J_{i,i+\ell}}{N} \tag{2.44}$$

Then we can evaluate the conditioned single attempt probability

$$\begin{aligned} P[S_j(i,\ell)|f_i^{i+\ell}] &= \sum_{a\in\mathcal{N}} P[k_i^j = a|f_i^L]P[\hat{k}_i^j = a|f_i^{i+\ell}] \\ &= \sum_{a\in\mathcal{N}} \frac{1}{N^2} J_{i,L}(a)J_{i,i+\ell}(a) \end{aligned} \tag{2.45}$$

and therefore obtain its average over the realization of $f_i^{i+\ell}$ as

$$P[S_j(i,\ell)] = \sum_{a\in\mathcal{N}} \frac{1}{N^2} E[J_{i,L}(a)J_{i,i+\ell}(a)] \tag{2.46}$$

$$= \frac{1}{N} E[J_{i,L}J_{i,i+\ell}] \tag{2.47}$$

Proposition 2.4. *$E[J_{i,L}J_{i,i+\ell}] = E[J_{i,i+\ell}^2]$ for all $L \geq \ell$.*

Proof. It is sufficient to prove that

$$E[J_{i,L}|J_{i,i+\ell}] = J_{i,i+\ell} \tag{2.48}$$

and then condition $E[J_{i,L}J_{i,i+\ell}]$ on $J_{i,i+\ell}$.

We have then

$$\begin{aligned} E[J_{i,L}|J_{i,i+\ell}] &= E[E[J_{i,L}|J_{i,L-1},\ldots,J_{i,i+\ell}]|J_{i,i+\ell}] \\ &= E[J_{i,L-1}|J_{i,i+\ell}] \end{aligned} \tag{2.49}$$

since $J_{i,L}$ is a Martingale. Iterating this argument we can easily prove (2.48). Then

$$\begin{aligned} E[J_{i,L}J_{i,i+\ell}] =& E[E[J_{i,L}|J_{i,i+\ell}]J_{i,i+\ell}] \\ =& E[J_{i,i+\ell}^2] \end{aligned} \tag{2.50}$$

□

Therefore the success probability of a single attempt can be rewritten as

$$P[S_j(i,\ell)] = \frac{1}{N} E[J_{i,i+\ell}^2] \tag{2.51}$$

In the same way we can write the joint success probability for two distinct guesses j, j' on the same chain $k_i, \ldots, k_{i+\ell}$

$$P[S_j(i,\ell) \cap S_{j'}(i,\ell)|f_i^{i+\ell}] = \sum_{a\in\mathcal{N}} P[k_i^j = a|f_i^L]P[\hat{k}_i^j = a|f_i^{i+\ell}]P[\hat{k}_i^{j'} = a|f_i^{i+\ell}] \tag{2.52}$$

$$= \sum_{a\in\mathcal{N}} \frac{1}{N^3} J_{i,L}(a)J_{i,i+\ell}(a)J_{i,i+\ell}(a) \tag{2.53}$$

$$= \sum_{a\in\mathcal{N}} \frac{1}{N^3} J_{i,L}(a)J_{i,i+\ell}^2(a) \tag{2.54}$$

And again

$$P[S_j(i,\ell) \cap S_{j'}(i,\ell)] = \frac{1}{N^3} \sum_{a\in\mathcal{N}} E[J_{i,L}(a)J_{i,i+\ell}^2(a)] \tag{2.55}$$

$$= \frac{1}{N^3} \sum_{a\in\mathcal{N}} E[J_{i,i+\ell}^3(a)] \tag{2.56}$$

$$= \frac{1}{N^2} E[J_{i,i+\ell}^3] \tag{2.57}$$

Finally we can derive lower and upper bounds for the success probability of the attack with N_A independent guessed on the same chain section, as

$$\sum_{j=1}^{N_A} P[S_j(i,\ell)] - \sum_{j=1}^{N_A}\sum_{j'<j} P[S_j(i,\ell) \cap S_{j'}(i,\ell)] \leq P[\cup_{j=1}^{N_A} S_j(i,\ell)] \leq \sum_{j=1}^{N_A} P[S_j(i,\ell)] \tag{2.58}$$

$$\frac{N_A E[J_\ell^2]}{N} - \frac{N_A(N_A-1)E[J_\ell^3]}{N^2} \leq P[\cup_{j=1}^{N_A} S_j(i,\ell)] \leq \frac{N_A E[J_\ell^2]}{N} \tag{2.59}$$

From section 2.2 we have

$$E[J_\ell^2] = (1-N)c_2^\ell + N \tag{2.60}$$

$$E[J_\ell^3] = c_3^\ell + \frac{3}{2}N(1-N)(c_2^\ell - c_3^\ell) + N^2(1-c_3^\ell) \tag{2.61}$$

and using Taylor series expansions

$$E[J_\ell^2] = \ell + 1 + O(\frac{\ell}{N}) \tag{2.62}$$

$$E[J_\ell^3] = \frac{3}{2}\ell^2 + \frac{5}{2}\ell + 1 + O(\frac{\ell^2}{N}) \tag{2.63}$$

Then, if $\ell \ll N$ we can rewrite (2.59) as

$$\frac{N_A(\ell+1)}{N} - \frac{N_A(N_A-1)(3\ell^2+5\ell+2)}{4N^2} \leq P_s \leq \frac{N_A(\ell+1)}{N} \tag{2.64}$$

The equation (2.64) is the same obtained in [9] under the assumption that J_ℓ has a compound Poisson distribution. The following graph shows the upper and lower bound on the probability of success of the attack using the approximate model (2.64) and the exact one (2.59) varying the number of possible keys N.

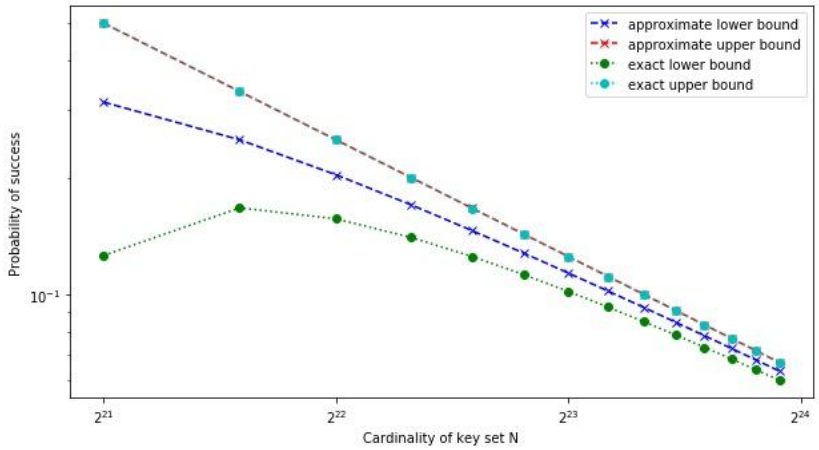

Figure 2.1: Probability of success vs N for $N_A = 2^9, \ell = 2^{11}$

We can see that the upper bounds are almost identical and that therefore the assumption that J_ℓ has a compound distribution is unnecessary.

Moreover we can calculate these bounds with a fixed N and varying the length of the hash chain ℓ.

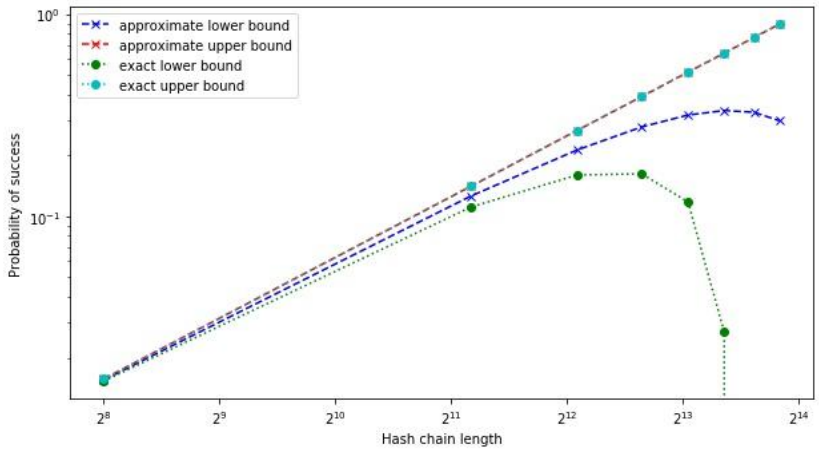

Figure 2.2: Probability of success vs ℓ for $N_A = 2^9, N = 2^{23}$

As expected increasing ℓ the bounds grow as well and for values of ℓ greater than 2^{13} the upper bound is close to 1.

Chapter 3

Approximate distribution

In this chapter we aim to find an approximation to the distribution of the random vector $\boldsymbol{J}_\ell = [J_\ell(b_1), \dots, J_\ell(b_m)]$ defined as in 2.1, where $b_1, \dots, b_m \in \mathcal{N}$ and $m < N = |\mathcal{N}|$. In particular we want to prove that, for $\ell, m \ll |\mathcal{N}|$, $\boldsymbol{J}_\ell$ is close to a random vector whose elements are independent on each other.

We know that the transition probability at the step ℓ conditioned on the step $\ell - 1$ has a multinomial distribution

$$P_{\boldsymbol{J}_\ell|\boldsymbol{J}_{\ell-1}}(\boldsymbol{y}|\boldsymbol{x}) = \prod_{i=1}^{m} \frac{1}{y_i!}\Big(\frac{x_i}{N}\Big)^{y_i} \frac{N!}{(N-y)!}\Big(1 - \frac{x}{N}\Big)^{N-y} \tag{3.1}$$

where $x = \sum_{i=1}^m x_i$ and $y = \sum_{i=1}^m y_i$

Arenbaev in [3] and Deheuvels & Pfeifer in [5] have found some bounds on the total variational distance between the distribution $\boldsymbol{S}$ defined as

$$P_{\boldsymbol{S}}(\boldsymbol{l}) = \prod_{i=1}^{m} \frac{{p_i}^{l_i}}{l_i!} \frac{N!}{(N-l)!}(1-p)^{N-l} \tag{3.2}$$

with $l = \sum_{i=1}^m l_i \leq N, p = \sum_{i=1}^m p_i$, and

$$\boldsymbol{\pi} = (\pi_1, \dots, \pi_m),$$

which is a vector of independent Poisson random variables with parameters $np_1, \dots, np_m$. We notice that $\boldsymbol{J}_1$ has the same distribution of $\boldsymbol{S}$ when $p_i = \frac{1}{N}$ for all i.

Considering these results we will try to find some bounds on the distance between $\boldsymbol{J}_\ell$ and a new Markov chain $\boldsymbol{J}'_\ell$ defined as

$$\boldsymbol{J}'_0(\boldsymbol{x}) = \begin{cases} 1 & \text{if } \boldsymbol{x} = (1, \dots, 1) \in \mathbb{N}^m \\ 0 & \text{if } \boldsymbol{x} \neq (1, \dots, 1) \in \mathbb{N}^m \end{cases} \tag{3.3}$$

with transition probability given by

$$P_{\boldsymbol{J}'_\ell|\boldsymbol{J}'_{\ell-1}}(\boldsymbol{y}|\boldsymbol{x}) = \prod_{i=1}^{m} x_i^{y_i} \frac{e^{-x_i}}{y_i!} \tag{3.4}$$

i.e. $\boldsymbol{J}'_\ell|\boldsymbol{J}'_{\ell-1}$ is a vector of independent Poisson random variables with parameters $\boldsymbol{J}'_{\ell-1}$.

However it turned out pretty impractical to apply the results concerning the total variational distance of [3] and [5] to our distributions, so in order to find an approximation we will introduce the Kullback-Leibler divergence.

3.1 Kullback-Leibler divergence

Definition 3.1. *The* Kullback-Leibler divergence *between two probability mass distributions P and Q on $\mathcal{X}$ is defined as*

$$\begin{aligned} D(P||Q) &= \sum_{x\in\mathcal{X}} P(x)\log\frac{P(x)}{Q(x)} \\ &= E_P\big[\log\frac{P}{Q}\big] \end{aligned} \tag{3.5}$$

It is easy to prove that the Kullback-Leibler divergence is always non-negative and is equal to zero if and only if $P = Q$.

However, it is not a distance since it is not simmetric and does not satisfy the triangle inequality. Nonetheless we can derive an upper bound on the total variational distance from the KL divergence through the Pinsker's inequality, which we state here.

Theorem 3.2 (Pinsker's Inequality)**.** *Let P, Q be two discrete probability distributions, then*

$$d(P,Q) \leq \sqrt{\frac{1}{2}D(P||Q)} \tag{3.6}$$

Since we do not have an explicit description of the distribution $\boldsymbol{J}_\ell$, but only of $\boldsymbol{J}_\ell|\boldsymbol{J}_{\ell-1}$ we will introduce a slightly different type of KL divergence.

Definition 3.3. *Let X, X' and Y, Y' be probability distribution on respectively $\mathcal{X}$ and $\mathcal{Y}$ and XX', YY' their joint distribution. The* conditional Kullback-Leibler divergence *between $X'|X$ and $X'|X$ is defined as*

$$\begin{aligned} D_C\big(P_{Y|X}||P_{Y'|X'}\big) &= \sum_{x\in\mathcal{X}} P_X(x)\sum_{y\in\mathcal{Y}} P_{Y|X}(y|x)\log\Big(\frac{P_{Y|X}(y|x)}{P_{Y'|X'}(y|x)}\Big) \\ &= E_{XY}\Big[\log\Big(\frac{P_{Y|X}}{P_{Y'|X'}}\Big)\Big] \end{aligned} \tag{3.7}$$

We have that $D_C\big(P_{Y|X}||P_{Y'|X'}\big) \geq 0$ with equality if and only if $P_{Y|X}(y|x) = P_{Y'|X'}(y|x)$ for all y and x with $P_{Y|X}(y|x) > 0$.

Proposition 3.4. *Using tha same notation of definition* (3.3)*, the following relation holds*

$$D(P_{XY}||P_{X'Y'}) = D(P_X||P_{X'}) + D_C(P_{Y|X}||P_{Y'|X'}). \tag{3.8}$$

Proof. We have

$$D(P_{XY}||P_{X'Y'}) = \sum_x \sum_y P_{XY}(x,y) \log \frac{P_{XY}(x,y)}{P_{X'Y'}(x,y)} \tag{3.9}$$

$$= \sum_x \sum_y P_{XY}(x,y) \log \frac{P_X(x)P_{Y|X}(y|x)}{P_{X'}(x)P_{Y'|X'}(y|x)} \tag{3.10}$$

$$= \sum_x \sum_y P_{XY}(x,y) \log \frac{P_X(x)}{P_{X'}(x)} + \sum_x \sum_y P_{XY}(x,y) \log \frac{P_{Y|X}(y|x)}{P_{Y'|X'}(y|x)} \tag{3.11}$$

$$= D(P_X||P_{X'}) + D_C(P_{Y|X}||P_{Y'|X'}) \tag{3.12}$$

□

3.2 Bounds on the divergence

We can now write an upper bound on the KL divergence between $\boldsymbol{J_\ell}$ and $\boldsymbol{J'_\ell}$ using the conditional divergence.

Lemma 3.5. *Let* $\boldsymbol{J_\ell}$ *and* $\boldsymbol{J'_\ell}$ *defined as in* 3.1 *and* 3.4. *Then*

$$D(\boldsymbol{J}_\ell||\boldsymbol{J'_\ell}) \leq \sum_{i=1}^{l} D_C(P_{\boldsymbol{J}_i|\boldsymbol{J}_{i-1}}||P_{\boldsymbol{J'_i}|\boldsymbol{J'_{i-1}}}) \tag{3.13}$$

Proof. Rewriting proposition (3.4) with $\boldsymbol{J}_\ell$ and $\boldsymbol{J'_\ell}$ we have

$$D(P_{\boldsymbol{J}_\ell,\boldsymbol{J}_{\ell-1}}||P_{\boldsymbol{J'_\ell},\boldsymbol{J'_{\ell-1}}}) = D(P_{\boldsymbol{J}_{\ell-1}}||P_{\boldsymbol{J'_{\ell-1}}}) + D_C(P_{\boldsymbol{J}_\ell|\boldsymbol{J}_{\ell-1}}||P_{\boldsymbol{J'_\ell}|\boldsymbol{J'_{\ell-1}}}) \tag{3.14}$$

and by exchanging the roles of the distributions

$$D(P_{\boldsymbol{J}_\ell,\boldsymbol{J}_{\ell-1}}||P_{\boldsymbol{J'_\ell},\boldsymbol{J'_{\ell-1}}}) = D(P_{\boldsymbol{J}_\ell}||P_{\boldsymbol{J'_\ell}}) + D_C(P_{\boldsymbol{J}_{\ell-1}|\boldsymbol{J}_\ell}||P_{\boldsymbol{J'_{\ell-1}}|\boldsymbol{J'_\ell}}) \tag{3.15}$$

Since $D_C(P_{\boldsymbol{J}_{\ell-1}|\boldsymbol{J}_\ell}||P_{\boldsymbol{J'_{\ell-1}}|\boldsymbol{J'_\ell}}) \geq 0$ we can conclude

$$D(\boldsymbol{J}_\ell||\boldsymbol{J'_\ell}) \leq D(\boldsymbol{J}_{\ell-1}||\boldsymbol{J'_{\ell-1}}) + D_C(P_{\boldsymbol{J}_\ell|\boldsymbol{J}_{\ell-1}}||P_{\boldsymbol{J'_\ell}|\boldsymbol{J'_{\ell-1}}}) \tag{3.16}$$

Then, knowing that $\boldsymbol{J}_0 = \boldsymbol{J'_0}$ we have

$$D(\boldsymbol{J}_\ell||\boldsymbol{J'_\ell}) \leq \sum_{i=1}^{l} D_C(P_{\boldsymbol{J}_i|\boldsymbol{J}_{i-1}}||P_{\boldsymbol{J'_i}|\boldsymbol{J'_{i-1}}}) \tag{3.17}$$

□

We will now focus on $D_C(P_{\boldsymbol{J}_\ell|\boldsymbol{J}_{\ell-1}}||P_{\boldsymbol{J}'_\ell|\boldsymbol{J}'_{\ell-1}})$ for a generic ℓ.

Proposition 3.6. *Let* $\boldsymbol{J_\ell}$ *and* $\boldsymbol{J'_\ell}$ *defined as in* 3.1 *and* 3.4. *Then*

$$D_C(P_{\boldsymbol{J}_\ell|\boldsymbol{J}_{\ell-1}}||P_{\boldsymbol{J}'_\ell|\boldsymbol{J}'_{\ell-1}}) = \log\frac{N!}{N^N} - E\Big[\log\frac{((N-J_\ell)!)}{(N-J_{\ell-1})^{(N-J_\ell)}}\Big] + m \quad (3.18)$$

Proof. Let $\boldsymbol{x} = (x_1,\dots,x_m)\,, \boldsymbol{y} = (y_1,\dots,y_m)\,, x = x_1+\cdots+x_m$ and $y = y_1+\cdots y_m$. We have

$$P_{\boldsymbol{J}_\ell|\boldsymbol{J}_{\ell-1}}(\boldsymbol{y}|\boldsymbol{x}) = \prod_{i=1}^{m}\frac{1}{y_i!}\Big(\frac{x_i}{N}\Big)^{y_i}\frac{N!}{(N-y)!}\Big(1-\frac{x}{N}\Big)^{N-y} \quad (3.19)$$

$$P_{\boldsymbol{J}'_\ell|\boldsymbol{J}'_{\ell-1}}(\boldsymbol{y}|\boldsymbol{x}) = \prod_{i=1}^{m}x_i^{y_i}\frac{e^{-x_i}}{y_i!} \quad (3.20)$$

and therefore

$$\log\frac{P_{\boldsymbol{J}_\ell|\boldsymbol{J}_{\ell-1}}(\boldsymbol{y}|\boldsymbol{x})}{P_{\boldsymbol{J}'_\ell|\boldsymbol{J}'_{\ell-1}}(\boldsymbol{y}|\boldsymbol{x})} = \log\frac{N!}{(N-y)!e^{-x}}\frac{1}{N^y}\Big(1-\frac{x}{N}\Big)^{N-y} \quad (3.21)$$

Hence

$$D_C(P_{\boldsymbol{J}_\ell|\boldsymbol{J}_{\ell-1}}||P_{\boldsymbol{J}'_\ell|\boldsymbol{J}'_{\ell-1}}) = E\Big[\log\frac{P_{\boldsymbol{J}_\ell|\boldsymbol{J}_{\ell-1}}(\boldsymbol{y}|\boldsymbol{x})}{P_{\boldsymbol{J}'_\ell|\boldsymbol{J}'_{\ell-1}}(\boldsymbol{y}|\boldsymbol{x})}\Big] \quad (3.22)$$

$$= E\Big[\log\frac{N!}{(N-y)!e^{-x}}\frac{1}{N^y}\Big(1-\frac{x}{N}\Big)^{N-y}\Big] \quad (3.23)$$

$$= E[\log N! - \log(N-y)! + x + \log(N-x)^{(N-y)} - \log N^N] \quad (3.24)$$

$$= \log N! - \log N^N + E[x] - E[\log((N-y)!)] + E[\log(N-x)^{(N-y)}] \quad (3.25)$$

$$= \log\frac{N!}{N^N} - E\Big[\log\frac{(N-y)!}{(N-x)^{(N-y)}}\Big] + E[x] \quad (3.26)$$

where the expectations are meant respect to the joint probability $P_{\boldsymbol{J_\ell},\boldsymbol{J_{\ell-1}}}$.

We notice that although the expectations are on $P_{\boldsymbol{J}_\ell,\boldsymbol{J}_{\ell-1}}$ they only depend on the sums $x = x_1+\cdots+x_m = \|\boldsymbol{J}_{\ell-1}\|_1$ and $y = y_1+\cdots+y_m = \|\boldsymbol{J}_\ell\|_1$. As we saw in section 2.1, $\|\boldsymbol{J}_\ell\|_1$ has the same distribution of J_ℓ.

Therefore we will consider the expectations in equation (3.26) taken on $J_\ell, J_{\ell-1}$, so we can say that $E[x] = E[J_{\ell-1}] = m$. □

We will now give a lower bound on the second term of equation (3.18).

Lemma 3.7. *Let* $\boldsymbol{J_\ell}$ *and* $\boldsymbol{J'_\ell}$ *defined as in* 3.1 *and* 3.4. *Then*

$$E\Big[\log\frac{(N-J_\ell)!}{(N-J_{\ell-1})^{(N-J_\ell)}}\Big] \geq E\Big[\log\frac{(N-J_{\ell-1})!}{(N-J_{\ell-1})^{(N-J_{\ell-1})}}\Big] \quad (3.27)$$

Proof. Consider $E[\log(N - J_{\ell-1})^{(N-J_\ell)}]$, we have

$$\begin{aligned}
E[\log(N - J_{\ell-1})^{(N-J_\ell)}] &= E[(N - J_\ell)\log(N - J_{\ell-1})] && (3.28)\\
&= E\big[E[(N - J_\ell)\log(N - J_{\ell-1})|J_{\ell-1}]\big] && (3.29)\\
&= E\big[E[(N - J_\ell)|J_{\ell-1}]\log(N - J_{\ell-1})\big] && (3.30)\\
&= E\big[(N - E[J_\ell|J_{\ell-1}])\log(N - J_{\ell-1})\big] && (3.31)\\
&= E[(N - J_{\ell-1})\log(N - J_{\ell-1})] && (3.32)\\
&= E[\log(N - J_{\ell-1})^{(N-J_{\ell-1})}] && (3.33)
\end{aligned}$$

To conclude we just need to prove that $E[\log((N - J_\ell)!)] \geq E[\log((N - J_{\ell-1})!)]$.

To do so we replace the factorial with the gamma function Γ and knowing that $\log\Gamma$ is convex on positive reals we can use Jensen's inequality and state the following:

$$\begin{aligned}
E[\log((N - J_\ell)!)] &= E[\log\Gamma(N - J_\ell + 1)] && (3.34)\\
&= E\big[E[\log\Gamma(N - J_\ell) + 1|J_{\ell-1}]\big] && (3.35)\\
&\geq E\big[\log\Gamma(E[(N - J_\ell + 1)|J_{\ell-1}])\big] && (3.36)\\
&= E[\log\Gamma(N - J_{\ell-1} + 1)] && (3.37)\\
&= E[\log((N - J_{\ell-1})!)] && (3.38)
\end{aligned}$$

□

Combining proposition (3.6) with lemma (3.7) we get the following result

$$D_C(P_{\boldsymbol{J}_\ell|\boldsymbol{J}_{\ell-1}}||P_{\boldsymbol{J}'_\ell|\boldsymbol{J}'_{\ell-1}}) \leq \log\frac{\Gamma(N+1)}{N^N} - E\Big[\log\frac{\Gamma(N - J_{\ell-1} + 1)}{(N - J_{\ell-1})^{(N-J_{\ell-1})}}\Big] + m \tag{3.39}$$

We notice that now the expectation depends only on $J_{\ell-1}$.

We can now prove the last bound we obtained on the conditional probability between $\boldsymbol{J}_\ell$ and $\boldsymbol{J}'_\ell$.

Proposition 3.8. *Let $\boldsymbol{J}_\ell$ and $\boldsymbol{J}'_\ell$ defined as in* 3.1 *and* 3.4. *Then*

$$D_C(P_{\boldsymbol{J}_\ell|\boldsymbol{J}_{\ell-1}}||P_{\boldsymbol{J}'_\ell|\boldsymbol{J}'_{\ell-1}}) \leq -\frac{1}{6}E\big[\log\big(1 - z_N(J_{\ell-1})\big)\big] \tag{3.40}$$

where

$$z_N(x) := \frac{8x^3 - 4(6N+1)x^2 + (24N^2 + 8N + 1)x + \frac{7}{300}}{8N^3 + 4N^2 + N + \frac{1}{30}} \tag{3.41}$$

Proof. Starting from equation (3.39), we can use an approximation of Γ given by Ramanujian in [2] and then proved by Karatsuba and Alzer in [1] and [7].

In particular we will use the following double inequality

$$\sqrt{\pi}\big(\frac{n}{e}\big)^n \sqrt[6]{8n^3+4n^2+n+\frac{1}{100}} < \Gamma(n+1) < \sqrt{\pi}\big(\frac{n}{e}\big)^n \sqrt[6]{8n^3+4n^2+n+\frac{1}{30}} \tag{3.42}$$

that holds for all $n \geq 0$.

Define the polynomial $Q(z) \coloneqq 8z^3+4z^2+z$.
We will use the right inequality for the term $\log \frac{\Gamma(N+1)}{N^N}$ and the left one for $E\Big[\log \frac{\Gamma(N-J_{\ell-1}+1)}{(N-J_{\ell-1})^{(N-J_{\ell-1})}}\Big]$ i.e.

$$\frac{\Gamma(N+1)}{N^N} < \sqrt{\pi}\big(\frac{1}{e}\big)^N \sqrt[6]{Q(N)+\frac{1}{30}} \tag{3.43}$$

$$\log \frac{\Gamma(N+1)}{N^N} < \frac{1}{2}\log \pi - N + \frac{1}{6}\log\big(Q(N)+\frac{1}{30}\big) \tag{3.44}$$

and

$$\frac{\Gamma(N-J_{\ell-1}+1)}{(N-J_{\ell-1})^{N-J_{\ell-1}}} > \sqrt{\pi}\big(\frac{1}{e}\big)^{N-J_{\ell-1}} \sqrt[6]{Q(N-J_{\ell-1})+\frac{1}{100}} \tag{3.45}$$

$$\log \frac{\Gamma(N-J_{\ell-1}+1)}{(N-J_{\ell-1})^{N-J_{\ell-1}}} > \frac{1}{2}\log \pi - N + J_{\ell-1} + \frac{1}{6}\log\big(Q(N-J_{\ell-1})+\frac{1}{100}\big) \tag{3.46}$$

$$E[\log \frac{\Gamma(N-J_{\ell-1}+1)}{(N-J_{\ell-1})^{N-J_{\ell-1}}}] > \frac{1}{2}\log \pi - N + m + \frac{1}{6}E\big[\log\big(Q(N-J_{\ell-1})+\frac{1}{100}\big)\big] \tag{3.47}$$

Using the inequalities (3.44) and (3.47) in equation (3.39) we get

$$D_C(P_{\boldsymbol{J}_\ell|\boldsymbol{J}_{\ell-1}}||P_{\boldsymbol{J}'_\ell|\boldsymbol{J}'_{\ell-1}}) \leq \frac{1}{6}E\Big[\log \frac{8N^3+4N^2+N+\frac{1}{30}}{8(N-J_{\ell-1})^3+4(N-J_{\ell-1})^2+N-J_{\ell-1}+\frac{1}{100}}\Big] \tag{3.48}$$

Rearranging the right term we get the conclusion. □

3.3 Numerical validation

In this final section we aim to find a numerical lower bound on $E[\log(1-z_N(J_{\ell-1}))]$. Knowing that $J_{\ell-1} \in \mathbb{N}, J_{\ell-1} \leq N$, we have $z_N(J_{\ell-1}) \in (0,1)$ and since $\log(1-y) \geq -y-y^2$ for $0 < y \leq \frac{1}{2}$, we can write

$$\begin{aligned} E[\log(1-z_N(J_{\ell-1}))] = \\ &\sum_{x\colon z_N(x)\leq\frac{1}{2}} P(x)\log(1-z_N(x)) + \sum_{x\colon z_N(x)>\frac{1}{2}} P(x)\log(1-z_N(x)) \\ &\geq \sum_{x\colon z_N(x)\leq\frac{1}{2}} P(x)(-z_N(x)-z_N^2(x)) + \sum_{x\colon z_N(x)>\frac{1}{2}} P(x)\log(1-z_N(x)) \end{aligned} \tag{3.49}$$

where $P(x) = P_{J_{\ell-1}}(x)$.

Now we observe that

$$\sum_{x\colon z_N(x)\leq\frac{1}{2}} P(x)(-z_N(x) - z_N^2(x)) \geq E[-z_N(J_{\ell-1}) - z_N^2(J_{\ell-1})] \tag{3.50}$$

and

$$\sum_{x\colon z_N(x)>\frac{1}{2}} P(x)\log(1 - z_N(x)) \geq \log(1 - \max_x z_N(x)) \sum_{x\colon z_N(x)>\frac{1}{2}} P(x) \tag{3.51}$$

It is easy to show that if $x \leq \frac{N}{5}$ we have $z_N(x) \leq \frac{1}{2}$ and that $\max_x z_N(x) = z_N(N)$, so we can conclude

$$E[\log(1-z_N(J_{\ell-1}))] \geq E[-z_N(J_{\ell-1})-z_N^2(J_{\ell-1})]+P(J_{\ell-1} > \frac{N}{5})\log(1-z_N(N)) \tag{3.52}$$

Clearly $E[-z_N(J_{\ell-1}) - z_N^2(J_{\ell-1})]$ depends only on the moments $E[J_{\ell-1}^d]$ for $d \leq 6$, for which we have already derived a closed formula. The last thing to do is give an upper bound to $P(J_{\ell-1} > \frac{N}{5})$ or equivalently a lower bound on

$$P(J_{\ell-1} \leq \frac{N}{5}) \tag{3.53}$$

If $J_0 = m$ with probability 1, let $k_0, k_1, k_2, \ldots, k_{\ell-1}$ be an increasing sequence of natural numbers with $k_0 = m$ and $k_{\ell-1} = \frac{N}{5}$. We have

$$P(J_{\ell-1} \leq \frac{N}{5}) \geq \prod_{i=1}^{\ell-1} P(J_i \leq k_i | J_{i-1} \leq k_{i-1}) \tag{3.54}$$

$$\geq \prod_{i=1}^{\ell-1} P(J_i \leq k_i | J_{i-1} = k_{i-1}) \tag{3.55}$$

We can calculate $P(J_i \leq k_i | J_{i-1} = k_{i-1})$ using the cumulative distribution function of the binomial distribution which give us

$$P(J_{\ell-1} \leq \frac{N}{5}) \geq \prod_{i=1}^{\ell-1} I_{1-\frac{k_{i-1}}{N}}(n - k_i, 1 + k_i) \tag{3.56}$$

where

$$I_x(n-k, 1+k) = (n-k)\binom{n}{k}\int_0^x t^{n-k-1}(1-t)^k dt \tag{3.57}$$

is the regularized incomplete beta function.

Then we can write

$$D_C(P_{\boldsymbol{J}_\ell|\boldsymbol{J}_{\ell-1}}||P_{\boldsymbol{J}'_\ell|\boldsymbol{J}'_{\ell-1}}) \leq \frac{1}{6}\Big(E[z_N(J_{\ell-1})+z_N^2(J_{\ell-1})]-P(J_{\ell-1} > \frac{N}{5})\log(1-z_N(N))\Big) \tag{3.58}$$

In the following graph we can see the values of the bound (3.58) for four different values of ℓ as a function of N, with a fixed m equal to 8. To calculate $P(J_{\ell-1} > \frac{N}{5})$ we used the formula (3.56) where the $k_0, k_1, \ldots, k_{\ell-1}$ are evenly spaced integers between m and N.

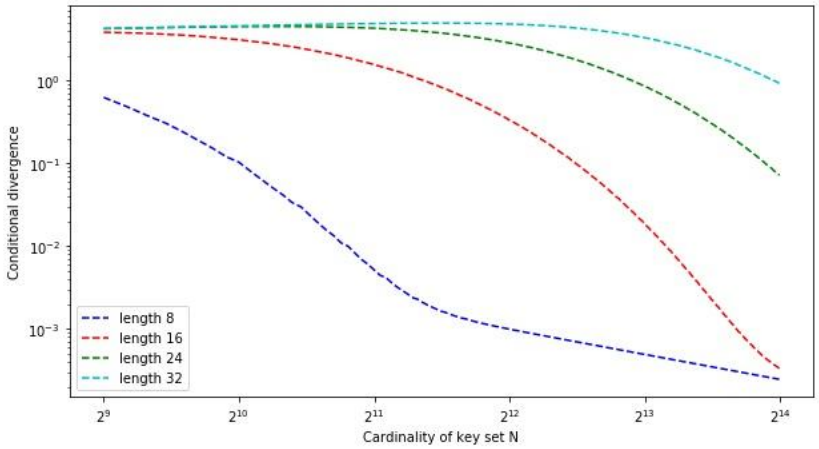

Figure 3.1: Conditional divergence vs N for $m = 8$

We can see that althought the value of the divergence decreases as N grows, even for small values of ℓ as 32 this bound on the conditional divergence is useless.

Bibliography

[1] Horst Alzer. "On Ramanujan's double inequality for the gamma function". In: *Bulletin of the London Mathematical Society* 35.5 (2003), pp. 601–607.

[2] George E Andrews and Bruce C Berndt. *Ramanujan's lost notebook*. Vol. 1. Springer, 2005.

[3] Nariman Karagaishievich Arenbaev. "Asymptotic behaviour of a multinomial distribution". In: *Teoriya Veroyatnostei i ee Primeneniya* 21.4 (1976), pp. 826–831.

[4] Gianluca Caparra et al. "Evaluating the security of one-way key chains in TESLA-based GNSS navigation message authentication schemes". In: *2016 International Conference on Localization and GNSS (ICL-GNSS)*. IEEE. 2016, pp. 1–6.

[5] Paul Deheuvels and Dietmar Pfeifer. "Poisson approximations of multinomial distributions and point processes". In: *Journal of multivariate analysis* 25.1 (1988), pp. 65–89.

[6] Yih-Chun Hu, David B Johnson, and Adrian Perrig. "SEAD: Secure efficient distance vector routing for mobile wireless ad hoc networks". In: *Ad hoc networks* 1.1 (2003), pp. 175–192.

[7] Ekatherina A Karatsuba. "On the asymptotic representation of the Euler gamma function by Ramanujan". In: *Journal of computational and applied mathematics* 135.2 (2001), pp. 225–240.

[8] Andreas Knoblauch. "Closed-form expressions for the moments of the binomial probability distribution". In: *SIAM Journal on Applied Mathematics* 69.1 (2008), pp. 197–204.

[9] Nicola Laurenti. "Correction of the collision attack success probability on TESLA protocol in the GNSS OS/CS NMA scenario". In: *private communication to V. Rijmen* (2016).

[10] Adrian Perrig et al. *Timed efficient stream loss-tolerant authentication (TESLA): Multicast source authentication transform introduction*. Tech. rep. 2005.