Randomness properties of $\mathbb{Z}_v$ ElGamal sequences

Daniel Panario* Lucas Pandolfo Perin† Brett Stevens*

*Carleton University — Canada
†Universidade Federal de Santa Catarina — Brazil
†Technical Innovation Institute — United Arab Emirates

2020-08-05

Outline

Outline

ElGamal Permutations

For p prime, $\mathbb{Z}_p^* = \{1, \ldots, p-1\}$ is a cyclic group of order $p-1$ under multiplication. For g a generator, the ElGamal map $x \rightarrow g^x$ from $\mathbb{Z}_p^*$ to $\mathbb{Z}_p^*$ is a permutation

- The ElGamal function is the basis of the ElGamal Signature Scheme
- The ElGamal function used in the Welch construction of Costas Arrays

Research challenge

In 2016 Joachim von zur Gathen posed this research challenge:

- Let $a, b, c \overset{?}{\leftarrow} \mathbb{Z}_p^*$.
- DDH assumption: $(g^a, g^b, g^{ab}) \sim (g^a, g^b, g^c)$

Research challenge

In 2016 Joachim von zur Gathen posed this research challenge:

- Let $a, b, c \stackrel{?}{\leftarrow} \mathbb{Z}_p^*$.
- DDH assumption: $(g^a, g^b, g^{ab}) \sim (g^a, g^b, g^c)$

How random is the ElGamal map?

Is $(x, g^x) \sim (x, x')$ when $x, x' \stackrel{?}{\leftarrow} \mathbb{Z}_p^*$?

Research challenge

In 2016 Joachim von zur Gathen posed this research challenge:

- Let $a, b, c \overset{?}{\longleftarrow} \mathbb{Z}_p^*$.
- DDH assumption: $(g^a, g^b, g^{ab}) \sim (g^a, g^b, g^c)$

How random is the ElGamal map?

Is $(x, g^x) \sim (x, x')$ when $x, x' \overset{?}{\longleftarrow} \mathbb{Z}_p^*$?

We proceed showing some evidence from (Niehues et al., 2020)

Cycles in ElGamal Permutations

Example: The generators of $\mathbb{Z}_5^*$ are 2 and 3.

x	g^x
1	$2^1 = 2$
2	$2^2 = 4$
3	$2^3 = 3$
4	$2^4 = 1$

$\gamma = (1, 2, 4)(3)$

x	g^x
1	$3^1 = 3$
2	$3^2 = 4$
3	$3^3 = 2$
4	$3^4 = 1$

$\gamma = (1, 2, 3, 4)$

Cycles in ElGamal Permutations

Example: The generators of $\mathbb{Z}_5^*$ are 2 and 3.

x	g^x
1	$2^1 = 2$
2	$2^2 = 4$
3	$2^3 = 3$
4	$2^4 = 1$

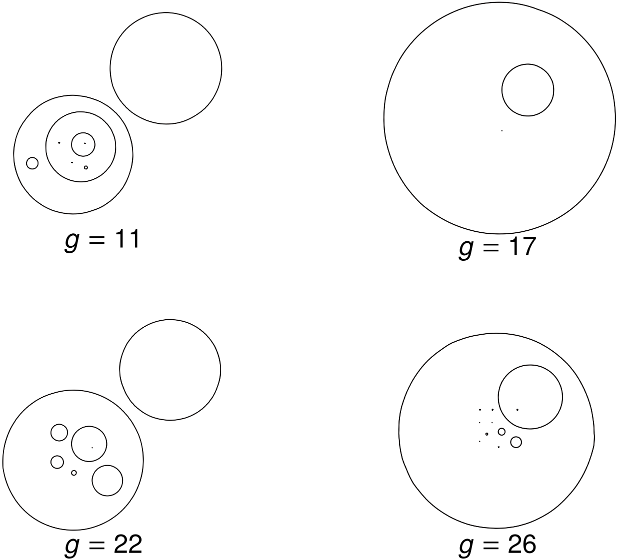
$$\gamma = (1, 2, 4)(3)$$

x	g^x
1	$3^1 = 3$
2	$3^2 = 4$
3	$3^3 = 2$
4	$3^4 = 1$

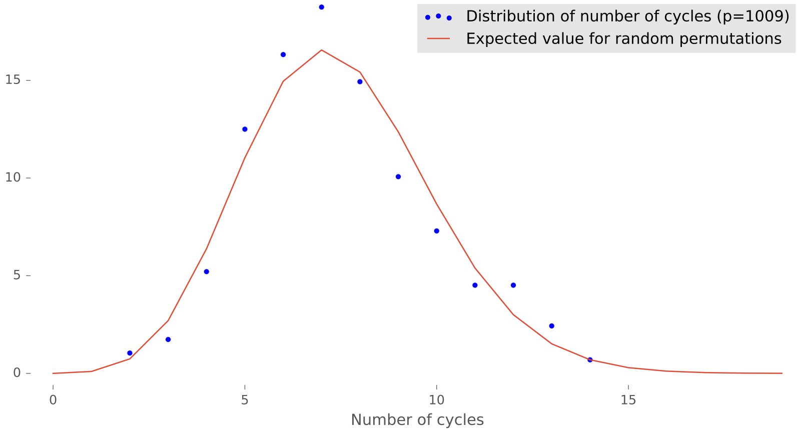
$$\gamma = (1, 2, 3, 4)$$

- Distinct g produce distinct permutations;
- Distinct g affect the cyclic structures.

$p = 1009$

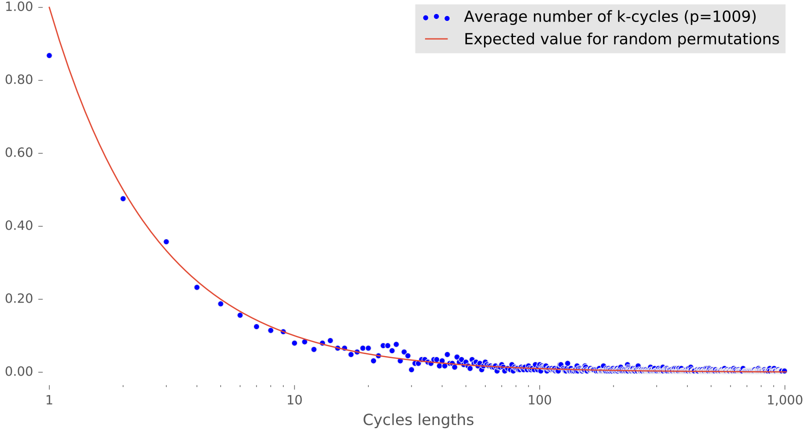


Number of cycles



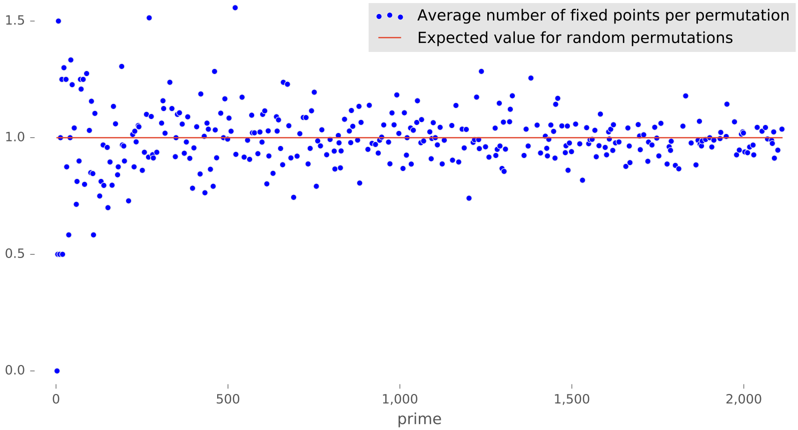
Distribution of number of cycles for all 288 generators of $\mathbb{F}_{1009}$

Number of k-cycles



Average number of k-cycles in $\mathbb{F}_{1009}$

Number of fixed points ($k = 1$)



Average number of fixed points in the generators of $\mathbb{F}_p$

Results with Sidon Sets

Let $S = \{(x, g^x) : x \in \mathbb{Z}_p^*\}$ be the graph of the ElGamal permutation. Because S is a Sidon Set,

Theorem (Niehues et al., 2020)

Let

$$B = [h_1, \ldots, h_2] \times [k_1, \ldots k_2] \subset \mathbb{Z}_p^* \times \mathbb{Z}_p.$$

Then

$$\left| \#S \cup B - \frac{\#B}{p} \right| \leq 50 p^{1/2} \log^2 p$$

Other randomness properties

- Drakakis et al. prove the ElGamal function is *Almost Perfect Nonlinear*
- Closer to PN than most APN functions in differential uniformity
- More linear than most Costas functions on a log-ratio test
- Less linear than random functions with a phase modulation test

Sequences from permutations

How about sequences?

Sequences from permutations

How about sequences?

For any permutation π in $\mathbb{Z}_p^*$, make a sequence

$$\pi_v = (\pi_1 \% v, \ldots, \pi_{p-1} \% v).$$

Sequences from permutations

How about sequences?

For any permutation π in $\mathbb{Z}_p^*$, make a sequence

$$\pi_v = (\pi_1 \% v, \ldots, \pi_{p-1} \% v).$$

Example: $p = 5$ and $g = 2$

$$\begin{aligned}
\gamma &= ((2^0)\%5)\%2, \ldots, (2^3)\%5)\%2) \\
&= (1, 2, 4, 3) \\
\gamma_2 &= (1, 0, 0, 1) \in \mathbb{Z}_2^4
\end{aligned}$$

Randomness properties of ElGamal Sequences?

How closely do ElGamal sequences compare to sequences from random permutations?

- Balance
- Period length
- Distribution of fixed t-*tuples* $z \in \mathbb{Z}_v^t$:

$$\lambda(z) = \#\{i \in [0, p-1] : \gamma_v(i +_n \iota) = z(\iota),\ 0 \leq \iota < t\}$$

- Distribution of *runs* of $b \in \mathbb{Z}_v$ and of length t:

$$\begin{aligned}\rho(b,t) =& \#\{i \in [0, p-1] : \\ & \gamma_v(i -_n 1), \gamma_v(i +_n t) \neq b = \gamma_v(i +_n \iota),\ 0 \leq \iota < t\}\end{aligned}$$

- $\rho(t) = v\rho(t+1)$

Other uses of Modulo operator in sequences

- The Legendre sequence

$$(\log_g(i)\%2, \log_g(i+1)\%2, \ldots)$$

- Colbourn constructed covering arrays from the circulant matrix

$$(\log_g(i)\%v, \log_g(i+1)\%v, \ldots)$$

- Tzanakis et al. formed covering array from circulant matrices of

$$(\log_g(tr(g^i))\%v, \log_g(tr(g^{i+1}))\%v, \ldots)$$

Outline

Balance

Proposition

Let π be a permutation in $\mathbb{Z}_p^*$, then π_v is a balanced sequence over $\mathbb{Z}_v$ if and only if $v \mid p-1$.

Proof.

The number of $x \equiv a \mod v$ in $[1, p-1]$ is

$$|\pi_v|_a = \lceil (p - 1 - ((a-1) \bmod v))/v \rceil$$

$\square$

Period

Lemma

If $p \equiv \alpha \neq 1 \pmod{v}$, then π_v has period $N = p - 1$ for any π permutation of $\mathbb{Z}_p^*$.

Proof.

The difference in the number of occurrences of any two symbols must be a multiple of $(p-1)/N$. But

$$|\pi_v|_a = \begin{cases} \lceil (p-1)/v \rceil & 0 \leq a < \alpha - 1, \\ \lfloor (p-1)/v \rfloor & \text{otherwise.} \end{cases}$$

$\square$

Period

Theorem

For every $\epsilon > 0$ there exists an n_ϵ so that for all $p \geq n_\epsilon$, the number T of balanced sequences π_v with period $p-1$ satisfies

$$(p-1)!(1-\epsilon) \leq T \leq (p-1)!. \tag{1}$$

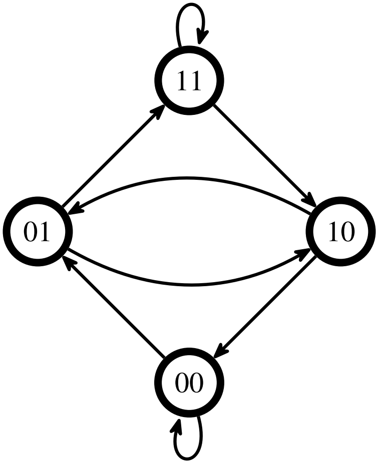
Special case

When q is prime and $p = vq + 1$,

$$\frac{(p-1)! - T}{(p-1)!} = \frac{v!(q!)^v}{(p-1)!}$$

This includes the case of Sophie Germain primes.

de Bruijn graph



Transfer Matrix

Transfer matrix is directed adjacency matrix of de Bruijn graph with variables

$$T = \begin{array}{c|c} & \begin{array}{cccc} 00 & 01 & 10 & 11 \end{array} \\ \begin{array}{c} 00 \\ 01 \\ 10 \\ 11 \end{array} & \begin{pmatrix} ux_0 & ux_0 & 0 & 0 \\ 0 & 0 & x_0 & x_0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \end{array}$$

$$C = \begin{array}{c|c} & \begin{array}{cccc} 00 & 01 & 10 & 11 \end{array} \\ \begin{array}{c} 00 \\ 01 \\ 10 \\ 11 \end{array} & \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

$$\sum_{\mathbf{k} \in \mathbb{N}^t} a_n(\mathbf{k}) x^{\mathbf{k}} = \sum_{z', z'' \in \mathbb{Z}_v^t} C_{z', z''} T^n_{z', z''}.$$

Asymptotic Normality

Theorem (Bender, Richmond, Williamson 1983)

Suppose $a_n(k)$ is admissible at 1 for $n \equiv n_0 \pmod{d}$ and that Λ is d-dimensional. Then $a_n(k)$ satisfies a central limit theorem for $n \equiv n_0 \pmod{d}$ with means and covariance matrix asymptotically proportional to n. Let q be such that $qc \in \Lambda$ for all $c \in \mathbb{Z}^v$. Then $a_n(k)$ satisfies a local limit theorem modulo Λ for $n \equiv n_0 \pmod{dq}$

Asymptotic Normality

Theorem

Let $z \in \mathbb{Z}_v^t$ and $t(\kappa)$ be the number of balanced circular sequences of length n over $\mathbb{Z}_v$ for which $\lambda(z) = \kappa$. There exists a $m_\lambda, b_\lambda, c_\lambda \in \mathbb{R}$ such that

$$\sup_\kappa \left| \frac{\sqrt{2\pi b_\lambda} t(\kappa)}{\binom{vl}{l,l,\ldots,l}} - c_\lambda e^{(\kappa - m_\lambda)^2/b_\lambda} \right| = o(1).$$

Let $b \in \mathbb{Z}_v$, $t \in \mathbb{N}$ and $r(\kappa)$ be the number of balanced circular sequences of length n over $\mathbb{Z}_v$ for which $\rho(b,t) = \kappa$. There exists a $m_\rho, b_\rho, c_\rho \in \mathbb{R}$ such that

$$\sup_\kappa \left| \frac{\sqrt{2\pi b_\rho} r(\kappa)}{\binom{vl}{l,l,\ldots,l}} - c_\rho e^{(\kappa - m_\rho^2)/b_\rho} \right| = o(1).$$

Mean for tuples

$$\frac{n}{v^t}\left(1+\frac{-(t^2-2tv+v^2-t)(v-1)}{2n}\right)+O\left(\frac{1}{n}\right)$$
$$\leq E(\lambda(z)) \leq$$
$$\frac{n}{v^t}\left(1+\frac{t(v-1)}{2n}\right)+O\left(\frac{1}{n}\right)$$

Variance for tuples

$$\frac{n}{v^{2t}}\left(\frac{2v^t}{2}+\frac{-12t^2v^t}{24n}\right)+O\left(\frac{1}{n}\right)$$

$$\lesssim \mathrm{VAR}(\lambda) \lesssim$$

$$\frac{n}{v^{2t}}\left(\frac{2v^t(v+1)}{2(v-1)}+\frac{12v^{t+2}t}{24n(v-1)}\right)+O\left(\frac{1}{n}\right)$$

Runs

$$
\begin{aligned}
E(\rho(b,t)) &= \frac{(l(v-1)-1)(v-1)l(l)_t}{(n-1)_{t+1}}, \\
\mathrm{VAR}(\rho(b,t)) &= \frac{(l(v-1)-1)(v-1)l(l)_t}{(n-1)_{t+1}} \\
&\quad + \frac{(v-1)l(l)_{2t}(l(v-1)-1)^2(l(v-1)-2)}{(n-1)_{2t+2}} \\
&\quad - \left(\frac{(l(v-1)-1)(v-1)l(l)_t}{(n-1)_{t+1}}\right)^2 .
\end{aligned}
$$

Where $l = n/v$.

Runs

$$E(\rho(b,t)) = \frac{n(v-1)}{v^{t+2}}\left((v-1) - \frac{(v-1)^2 t^2 - (v+3)(v-1)t + 2}{2n}\right) + O\left(\frac{1}{n}\right)$$

$$\mathrm{VAR}(\rho(b,t)) \approx \frac{n(v-1)^2}{v^{t+2}}\left(1 + \frac{-(v-1)t^2}{2n}\right) + O\left(\frac{1}{n}\right)$$

Outline

Contextualization

Bounds for random v-ary sequences

Bounds for ElGamal v-ary sequences

Experimental results

Final Remarks

Balance

Proposition

Let π be a permutation in $\mathbb{Z}_p^*$, then π_v is a balanced sequence over $\mathbb{Z}_v$ if and only if $v \mid p-1$.

Period

Theorem

The ElGamal sequence γ_v has period $N = p - 1$.

Proof.

1. $p \not\equiv 1 \pmod{v}$: Use balance
2. $p \equiv 1 \pmod{v}$: Suppose period $N < p - 1$: $g^{i+N}\%p \equiv_v g^i\%p$
3. Let $i = 0$: $g' = g^N\%p \equiv_v 1$.
4. Let $p = kg' + r$, $x = k + 1$ $(p < xg' < 2p)$. Let $i = \log_g(x)$:

$$x \equiv_v xg'\%p = xg' - p \equiv_v xg' - 1$$

5. $x(g' - 1) \equiv_v 1 \equiv_v g'$ is a contradiction.

□

Tuples

Theorem

Let γ_v be an ElGamal sequence and $p = qg^{t-1} + r$, then

$$\left\lfloor \frac{g}{v} \right\rfloor^{t-1} \left\lfloor \frac{q}{v} \right\rfloor \leq \lambda(z) \leq \left\lceil \frac{g}{v} \right\rceil^{t-1} \left(\left\lfloor \frac{q}{v} \right\rfloor + 1 \right).$$

Proof

$$X = \left\{x \in [1, p-1] : (g^i x)\%p \equiv_v z_i, 0 \le i < t\right\}$$

Let $c_i = g^i z_0 - z_i$, $0 \le i < t$.

$D = \{d \in \mathbb{Z}^t : d_0 = 0, d_i \equiv_v \alpha^{-1} c_i$ and $g d_{i-1} \le d_i < g(d_{i-1}+1)$ for $0 < i < t\}$.

For $d \in D$, let

$$X_d = \left\{x \in \mathbb{Z} : x \equiv_v z_0,\ \frac{d_i p}{g^i} \le x < \frac{(d_i + 1)p}{g^i},\ \text{for } 0 \le i < t\right\}.$$

Claim: $$X = \bigcup_{d \in D} X_d$$

$X_d \subset X$

If $x \in X_d$, then $x \equiv_v z_0$ and

$$d_i p \leq g^i x < (d_i + 1)p$$

Thus

$$g^i x \% p = g^i x - d_i p \equiv_v g^i x - \alpha d_i \equiv_v g^i z_0 - c_i \equiv_v g^i z_0 - (g^i z_0 - z_i) = z_i,$$

So $x \in X$.

$X \subset \cup X_d$

For $x \in X$, define $g^i x = q_i p + r_i$:

$$
\begin{aligned}
q_0 &= 0 \\
r_i &= g^i x - q_i p = (g^i x)\%p \equiv_v z_i \\
\frac{q_i p}{g^i} &\leq x < \frac{(q_i+1)p}{g^i}
\end{aligned}
$$

So $x \in X_{(q_0,\ldots,q_{t-1})}$

$X \subset \cup X_d$

$$q_i \equiv_v \alpha^{-1} q_i p = \alpha^{-1}(g^i x - r_i) \equiv_v \alpha^{-1}(g^i z_0 - z_i) = \alpha^{-1} c_i.$$

Then,

$$\begin{aligned} q_i &= \frac{g^i x - r_i}{p} = \frac{g(g^{i-1}x) - r_i}{p} = \frac{g(q_{i-1}p + r_{i-1}) - r_i}{p} \\ &= g q_{i-1} + g\frac{r_{i-1}}{p} - \frac{r_i}{p} < g(q_{i-1} + 1), \end{aligned}$$

and

$$g q_{i-1} = \frac{g q_{i-1} p}{p} \leq \frac{g(q_{i-1}p + r_{i-1})}{p} = \frac{g(g^{i-1}x)}{p} = \frac{g^i x}{p} = q_i + \frac{r_i}{p}.$$

Since $g q_{i-1}, q_i \in \mathbb{Z}$ and $r_i/p < 1$, $\Rightarrow q_i \geq g q_{i-1}$.
Thus $(q_0, \ldots, q_{t-1}) \in D$.

Final step

$$X = \bigcup_{d \in D} X_d = \bigcup_{d \in D} \left(\{x \equiv_v z_0\} \bigcap \left(\bigcap_{0 \le i < t} \left\{ \frac{d_i p}{g^i} \le x < \frac{(d_i + 1)p}{g^i} \right\} \right) \right).$$

$$= \bigcup_{d \in D} \left(\{x \equiv_v z_0\} \cap \left\{ \frac{d_{t-1} p}{g^{t-1}} \le x < \frac{(d_{t-1} + 1)p}{g^{t-1}} \right\} \right).$$

$$\begin{array}{rcccl} \lfloor g/v \rfloor^{t-1} & \le & \#D & \le & \lceil g/v \rceil^{t-1} \\ q & \le & \#[d_{t-1}p/g^{t-1}, (d_{t-1}+1)p/g^{t-1}) & \le & q+1 \\ \lfloor q/v \rfloor & \le & \#X_d & \le & \lceil (q+1)/v \rceil \end{array}$$

$$\left\lfloor \frac{g}{v} \right\rfloor^{t-1} \left\lfloor \frac{q}{v} \right\rfloor \le \lambda(z) \le \left\lceil \frac{g}{v} \right\rceil^{t-1} \left(\left\lfloor \frac{q}{v} \right\rfloor + 1 \right).$$

□

Observations

- When $g = mv$ bounds differ by at most m^t
- When $g = v$, $\left\lfloor \frac{q}{v} \right\rfloor \leq \lambda(z) \leq \left\lfloor \frac{q}{v} \right\rfloor + 1$
- If $p \geq vg^{t-1}$ and $g \geq v$, then $\lambda(z) > 0$ for all $z \in \mathbb{Z}_v^t$
- If $\lambda(z) > 0$ for all $z \in \mathbb{Z}_v^t$, then $g \geq v$ and $p \geq v^t + 1$.
- Coincide when $g = v$
- $\gamma_v(i+1) \equiv_v g\gamma_v(i) - s$ for some $0 \leq s < g$.

Runs

Theorem

Let γ_v be an ElGamal sequence and $p = qg^{t-1} + r$. For $z \in \mathbb{Z}_v^t$, let

$$\mu(z) = \#\{i \in [1, p-1] : \; g^{i+j}\%p \equiv_v z_j, \; 0 \le j < t-1, \; g^{i+t-1}\%p \not\equiv_v z_{t-1}\}.$$

Then

$$\left\lfloor \frac{g}{v} \right\rfloor^{t-2} \left\lfloor \frac{(v-1)g}{v} \right\rfloor \left\lfloor \frac{q}{v} \right\rfloor \le \mu(z) \le \left\lceil \frac{g}{v} \right\rceil^{t-2} \left\lceil \frac{(v-1)g}{v} \right\rceil \left(\left\lfloor \frac{q}{v} \right\rfloor + 1 \right).$$

Corollary

Let $p = q_t g^t + r_t$ and $p = q_{t+1} g^{t+1} + r_{t+1}$. Then

$$\left\lfloor \frac{g}{v} \right\rfloor^{t-1} \left\lfloor \frac{(v-1)g}{v} \right\rfloor \left\lfloor \frac{q_t}{v} \right\rfloor - \left\lceil \frac{g}{v} \right\rceil^{t} \left\lceil \frac{(v-1)g}{v} \right\rceil \left\lceil \frac{q_{t+1}+1}{v} \right\rceil$$
$$\leq \rho(b,t) \leq$$
$$\left\lceil \frac{g}{v} \right\rceil^{t-1} \left\lceil \frac{(v-1)g}{v} \right\rceil \left\lceil \frac{q_t+1}{v} \right\rceil - \left\lfloor \frac{g}{v} \right\rfloor^{t} \left\lfloor \frac{(v-1)g}{v} \right\rfloor \left\lfloor \frac{q_{t+1}}{v} \right\rfloor,$$

and

$$(v-1) \left\lfloor \frac{g}{v} \right\rfloor^{t} \left\lfloor \frac{(v-1)g}{v} \right\rfloor \left\lfloor \frac{q}{v} \right\rfloor \leq \rho(b,t) \leq (v-1) \left\lceil \frac{g}{v} \right\rceil^{t} \left\lceil \frac{(v-1)g}{v} \right\rceil \left\lceil \frac{q+1}{v} \right\rceil.$$

Comparison to random balanced sequences

From theoretical results

- Balance matches exactly
- Periodicity matches very closely
- To first order, the number of tuples and runs matches
- To first order $\rho(t) \approx v\rho(t+1)$

Outline

Experimental setting

We run experiments over two distinct data sets of pairs (p, v) with $p > 1,000,000$ and $2 \leq v \leq 8$.

all primes: Primes where $v \mid p - 1$.

$g = v$ primes: Primes where $v \mid p - 1$ and v is a generator.

Experimental setting

We run experiments over two distinct data sets of pairs (p, v) with $p > 1,000,000$ and $2 \leq v \leq 8$.

all primes: Primes where $v \mid p - 1$.

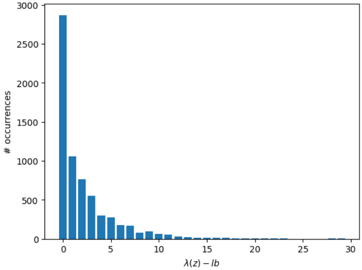
$g = v$ primes: Primes where $v \mid p - 1$ and v is a generator.

	all	$g = v$
# pairs (p, v)	715	400
# distinct v	7	4
# distinct primes	322	323
# v per prime (average)	4.51	1.48

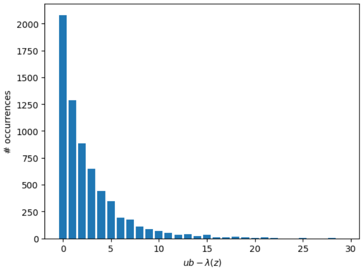
- We run experiments over *all primes* for the smallest 10 generators.
- If $v \in \{4, 5, 8\}$ then $v \neq g$.

ElGamal Sequences t-tuple bound gap distribution

Lower bound $\lambda(z)$ >0
$t = 2$ and 12% outliers.



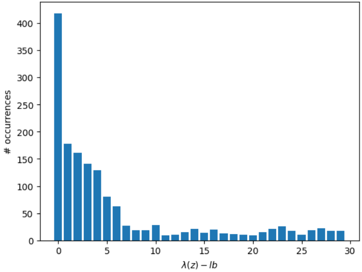
Upper bound
$t = 2$ and 5% outliers



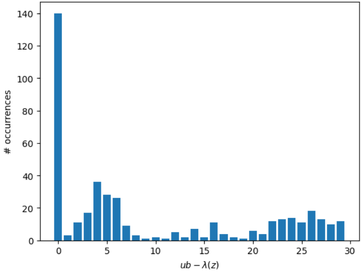
Distribution of gaps between $\lambda(z)$ and lower and upper bounds.

ElGamal Sequences t-tuple bound gap distribution

Lower bound $\lambda(z) > 0$
$t = 7$ and 59.75% outliers.



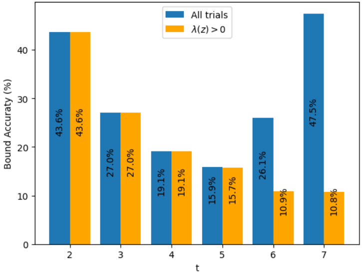
Upper bound
$t = 7$ and 93.56% outliers



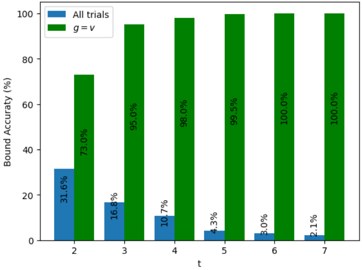
Distribution of gaps between $\lambda(z)$ and lower and upper bounds.

ElGamal Sequences t-tuple bound accuracy

Lower bound



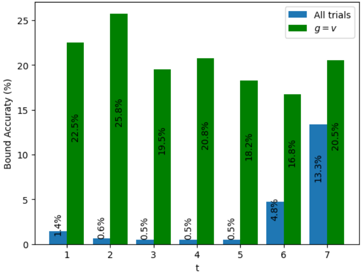
Upper bound



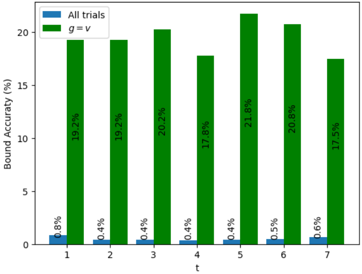
Percentage of trials with $z \in \mathbb{Z}_v^t$ s.t. $\lambda(z)$ matches lower and upper bounds.

ElGamal Sequences run bound accuracy

Lower bound



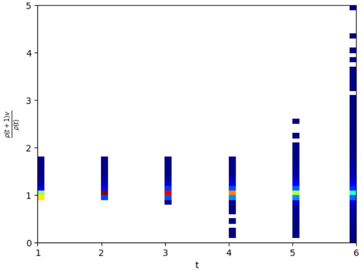
Upper bound



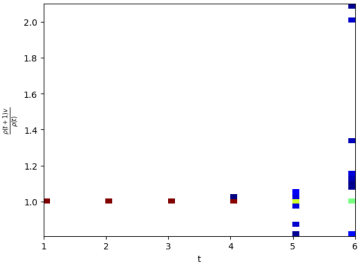
Percentage of trials with $b \in \mathbb{Z}_v$ s.t. $\rho(b, t)$ matches lower and upper bounds.

ElGamal Sequences run ratio Experiment

All primes.



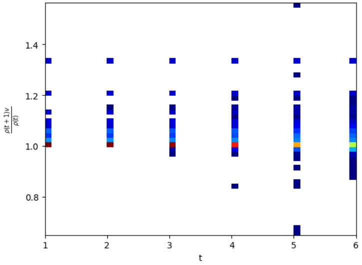
$g = v$ primes.



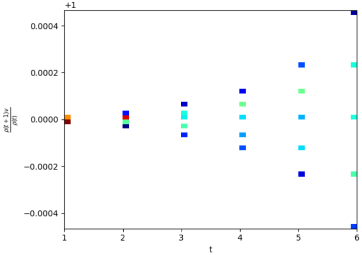
Distribution of $\rho(t+1)v/\rho(t)$ as a heat map with $2 \leq v \leq 8$

ElGamal Sequences run ratio Experiment

All primes.



$g = v$ primes.



Distribution of $\rho(t+1)v/\rho(t)$ as a heat map with $v = 2$

Outline

Contextualization

Bounds for random v-ary sequences

Bounds for ElGamal v-ary sequences

Experimental results

Final Remarks

Conclusions

- ElGamal permutations behave like random for cycle sizes and distribution of graph
- ElGamal permutations are close to random permutations for nonlinearity
- ElGamal sequences have balance and periodicity close to random
- Tuples in ElGamal sequences are distributed as in random balanced sequences
- Run lengths in ElGamal sequences satisfy Golomb's Randomness Postulate

Next steps

- Experiments indicate that $\lambda(z)$ bounds are tight. So any improvements will be conditional
- Prove properties of the distribution of $\lambda(z)$
- Prove linear complexity results for ElGamal sequences
- Determine expected linear complexity for random balanced random sequences
- Further investigate auto-correlation
- Will these be enough to justify cryptographic utility?

Obrigado
Thanks
ᑭᓇᓈᐢᑯᒥᑎᐣ

شكرا لك