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

Contextualization

Bounds for random v-ary sequences

Bounds for ElGamal *v*-ary sequences

Experimental results

Final Remarks

References

Outline

Contextualization

Bounds for random *v*-ary sequences

Bounds for ElGamal *v*-ary sequences

Experimental results

Final Remarks

References

ElGamal Permutations

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

How random is the map $\gamma(x) = g^x$?

Cycle sizes in ElGamal Permutations

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

		5			
X	g^x		_ X	g^{x}	
1	$2^1 = 2$		1	$3^1 = 3$	
2	$2^2 = 4$		2	$3^2 = 4$	
3	$2^3 = 3$		3	$3^3 = 2$	
4	$2^4 = 1$		4	$3^4 = 1$	
$\gamma = (1, 2, 4)(3)$			γ =	= (1, 2, 3, 4)	

Cycle sizes in ElGamal Permutations

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

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

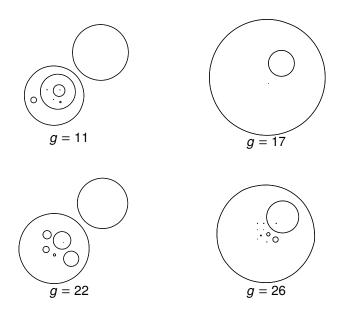
X	g^{x}
1	$3^1 = 3$
2	$3^2 = 4$
3	$3^3 = 2$
4	$3^4 = 1$

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

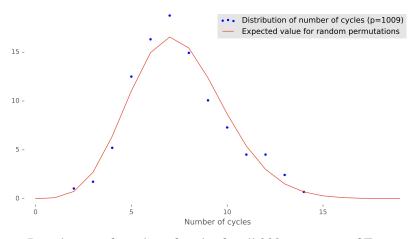
$$\gamma = (1, 2, 3, 4)$$

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

p = 1009

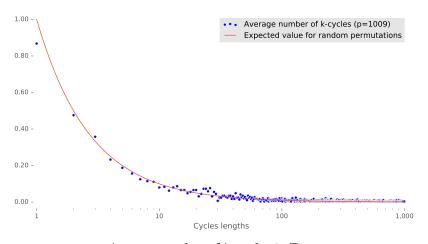


Number of cycles



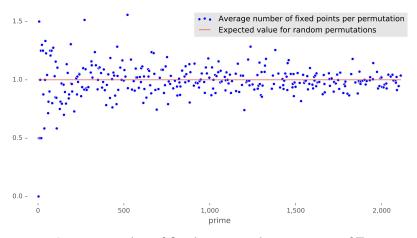
Distribution of number of cycles for all 288 generators of \mathbb{F}_{1009}

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, von zur Gathen, P, Zumalacárregui, 2020)

Let

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

Then

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

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

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

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

Sequences from permutations

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

$$\pi_{v} = (\pi_{1}\%v, \dots, \pi_{p-1}\%v).$$

Example: p = 5 and g = 2

$$\gamma = ((2^{0})\%5)\%2, \dots, (2^{3})\%5)\%2)$$
$$= (1, 2, 4, 3)$$
$$\gamma_{2} = (1, 0, 0, 1) \in \mathbb{Z}_{2}^{4}$$

Randomness properties of ElGamal Sequences?

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

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

$$\lambda(z) = \#\{i \in [0, p-1] : \gamma_{\nu}(i+_{n}\iota) = z(\iota), \ 0 \le \iota < t\}$$

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

$$\begin{split} \rho(b,t) = & \#\{i \in [0,p-1]: \\ \gamma_{\nu}(i-_{n}1), \gamma_{\nu}(i+_{n}t) \neq b = \gamma_{\nu}(i+_{n}\iota), \ 0 \leq \iota < t\} \end{split}$$

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

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

Outline

Contextualization

Bounds for random *v*-ary sequences

Bounds for ElGamal *v*-ary sequences

Experimental results

Final Remarks

References

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

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

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

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

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

$$(p-1)!(1-\epsilon) \le T \le (p-1)!.$$
 (1)

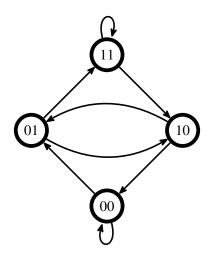
Special case

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

$$T = \begin{cases} 00 & 01 & 10 & 11 \\ 00 & ux_0 & ux_0 & 0 & 0 \\ 0 & 0 & x_0 & x_0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{cases}$$

$$C = \begin{cases} 01 & 00 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{cases}$$

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

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

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

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

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

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

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

$$\lesssim VAR(\lambda) \lesssim$$

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

Runs

$$E(\rho(b,t)) = \frac{(I(v-1)-1)(v-1)I(I)_t}{(n-1)_{t+1}},$$

$$VAR(\rho(b,t)) = \frac{(I(v-1)-1)(v-1)I(I)_t}{(n-1)_{t+1}} + \frac{(v-1)I(I)_{2t}(I(v-1)-1)^2(I(v-1)-2)}{(n-1)_{2t+2}} - \left(\frac{(I(v-1)-1)(v-1)I(I)_t}{(n-1)_{t+1}}\right)^2.$$

Where I = n/v.

Runs

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

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

Outline

Contextualization

Bounds for random *v*-ary sequences

Bounds for ElGamal v-ary sequences

Experimental results

Final Remarks

References

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 $N : <math>g^{i+N} \% p \equiv_v g^i \% p$
- 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

$$X = \left\{ x \in [1, p - 1] : (g^{i}x)\%p \equiv_{V} z_{i}, 0 \leq i < t \right\}$$

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

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

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 $g^i x = q_i p + r_i$:

$$q_0 = 0$$

$$r_i = g^i x - q_i p = (g^i x) \% p \equiv_v z_i$$

$$\frac{q_i p}{g^i} \le x < \frac{(q_i + 1)p}{g^i}$$

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

$$X \subset \cup X_d$$

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

Then,

$$q_{i} = \frac{g'x - r_{i}}{p} = \frac{g(g^{i-1}x) - r_{i}}{p} = \frac{g(q_{i-1}p + r_{i-1}) - r_{i}}{p}$$
$$= gq_{i-1} + g\frac{r_{i-1}}{p} - \frac{r_{i}}{p} < g(q_{i-1} + 1),$$

and

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

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

Final step

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

$$\lfloor g/v \rfloor^{t-1} \leq \#D \leq \lceil g/v \rceil^{t-1}$$

$$q \leq \# \lceil d_{t-1}p/g^{t-1}, (d_{t-1}+1)p/g^{t-1}) \leq q+1$$

$$\lfloor q/v \rfloor \leq \#X_d \leq \lceil (q+1)/v \rceil$$

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

Observations

- When g = mv bounds differ by at most m^t
- ▶ When g = v, $\left\lfloor \frac{q}{v} \right\rfloor \le \lambda(z) \le \left\lfloor \frac{q}{v} \right\rfloor + 1$
- ▶ If $p \ge vg^{t-1}$ and $g \ge v$, 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 $g \ge v$ and $p \ge v^t + 1$.
- ightharpoonup Coincide when g = v
- $ightharpoonup \gamma_{V}(i+1) \equiv_{V} g\gamma_{V}(i) s \text{ for some } 0 \leq s < g.$

Runs

Theorem

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

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

Then

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

Corollary

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

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

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

Contextualization

Bounds for random *v*-ary sequences

Bounds for ElGamal *v*-ary sequences

Experimental results

Final Remarks

References

Experimental setting

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

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 $2 \le v \le 8$.

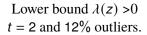
all primes: Primes where $v \mid p - 1$.

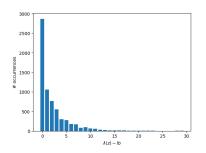
g = v primes: Primes where $v \mid p - 1$ and v is a generator.

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

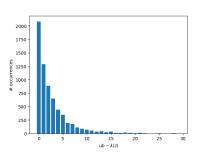
- We run experiments over *all primes* for the smallest 10 generators.
- ► If $v \in \{4, 5, 8\}$ then $v \neq g$.

ElGamal Sequences *t*-tuple bound gap distribution





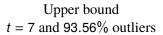
Upper bound t = 2 and 5% outliers

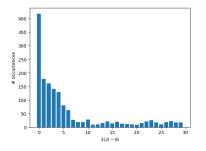


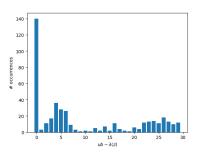
Distribution of gaps between $\lambda(z)$ and lower and upper bounds.

ElGamal Sequences *t*-tuple bound gap distribution

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

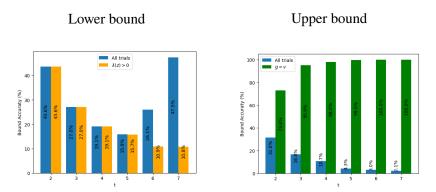






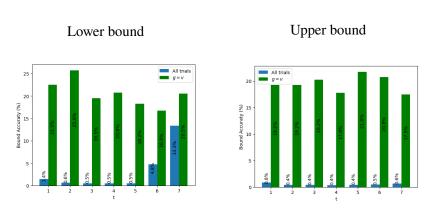
Distribution of gaps between $\lambda(z)$ and lower and upper bounds.

ElGamal Sequences *t*-tuple bound accuracy



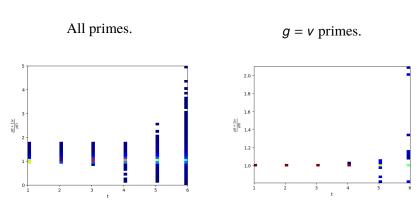
Percentage of trials with $z \in \mathbb{Z}_{V}^{t}$ s.t. $\lambda(z)$ matches lower and upper bounds.

ElGamal Sequences run bound accuracy



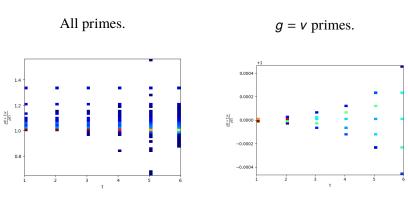
Percentage of trials with $b \in \mathbb{Z}_v$ s.t. $\rho(b,t)$ matches lower and upper bounds.

ElGamal Sequences run ratio Experiment



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

ElGamal Sequences run ratio Experiment



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

Linear Complexity

Outline

Contextualization

Bounds for random *v*-ary sequences

Bounds for ElGamal *v*-ary sequences

Experimental results

Final Remarks

References

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

شكرا لك

Thanks

Outline

Contextualization

Bounds for random *v*-ary sequences

Bounds for ElGamal *v*-ary sequences

Experimental results

Final Remarks

References

References I



Boppré Niehues, L., von zur Gathen, J., Perin, L. P., & Zumalacárregui, A. (2020). Sidon sets and statistics of the elgamal function. *Cryptologia*, 44(5), 438–450.