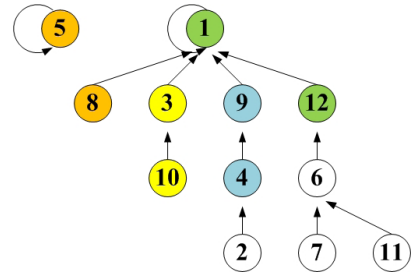
Variations on a Theme of DLP

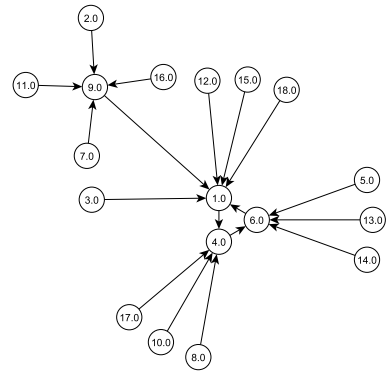


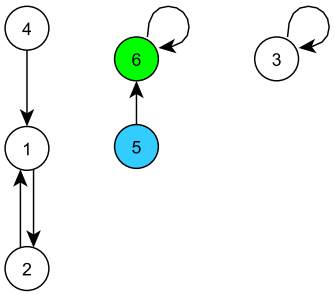
Joshua Holden

Joint work with JingJing Chen, Matthew Friedrichsen, Brian Larson, Mark Lotts, Emily McDowell, and Alex Wood (RHIT REU)

http://www.rose-hulman.edu/~holden

$$x \mapsto x^x$$





$$x \mapsto xg^x$$

$$x \mapsto g^{x^2}$$

We are investigating the functional graph induced by maps related to the discrete exponentiation map.

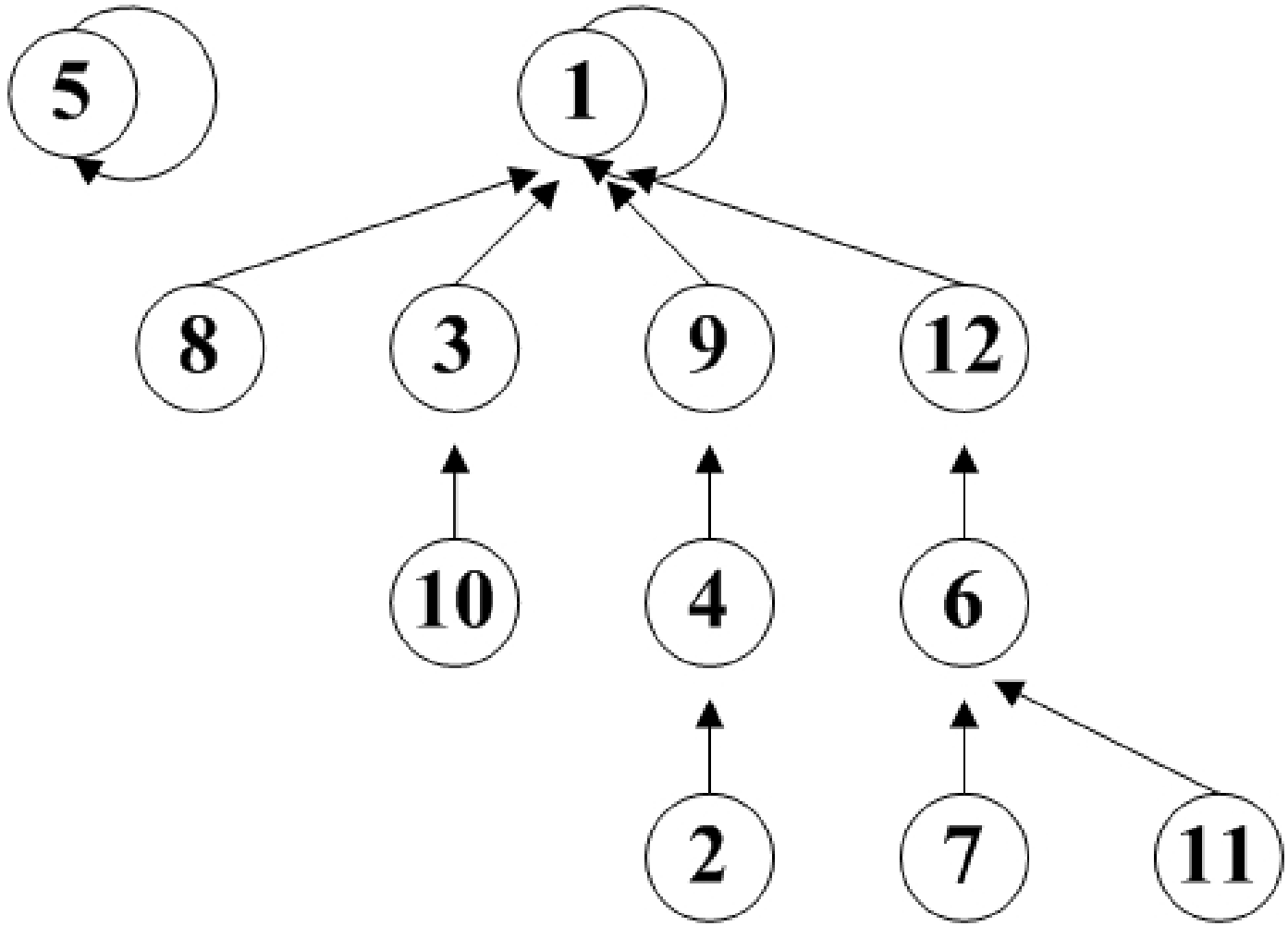
The discrete exponentiation map is the map $x \mapsto g^x \bmod p$.

Definition A *functional graph* is a directed graph such that each vertex must have exactly one edge directed out from it.

For each x, make an edge from the vertex x to the vertex $f(x)$.

Question How much does one of our functional graphs look like a "random graph"?

We are investigating the functional graph induced by maps related to the discrete exponentiation map.



Example: The self-power map $x \mapsto x^x \bmod 13$

The problem of inverting the discrete exponentiation map is called the Discrete Logarithm Problem.

$$x \mapsto y \equiv g^x \bmod p$$

Discrete exponentiation: easy.

$$x \mapsfrom y \equiv g^x \bmod p$$

Discrete logarithm: thought to be hard.

The assumption that the DLP is hard underlies the (assumed) security of several cryptographic protocols.

- Diffie-Hellman Key Agreement
- Blum-Micali Cryptographically Secure Pseudorandom Number Generator
- ElGamal Encryption
- ElGamal Digital Signature Scheme
- Elliptic Curve Pairing-Based Cryptography
- (Elliptic Curve DLP) Elliptic Curve Cryptography
- Etc.

My students and I have so far looked at three variations on the discrete exponentiation map.

The self-power map is the map $x \mapsto x^x \bmod p$.

The assumption that inverting the self-power map is hard underlies the security of a variation of the ElGamal Digital Signature Scheme.

My students and I have so far looked at three variations on the discrete exponentiation map.

The self-power map is the map $x \mapsto x^x \bmod p$.

$$x \mapsto y \equiv x^x \bmod p \qquad \text{(Easy)}$$

The assumption that inverting the self-power map is hard underlies the security of a variation of the ElGamal Digital Signature Scheme.

$$x \not\mapsto y \equiv x^x \bmod p \qquad \text{(Hard?)}$$

My students and I have so far looked at three variations on the discrete exponentiation map.

The self-power map is the map $x \mapsto x^x \bmod p$.

$$x \mapsto y \equiv x^x \bmod p \qquad \text{(Easy)}$$

The assumption that inverting the self-power map is hard underlies the security of a variation of the ElGamal Digital Signature Scheme.

$$x \mapsfrom y \equiv x^x \bmod p \qquad \text{(Hard?)}$$

Matthew Friedrichsen, Brian Larson, and Emily McDowell started investigating this in 2010.

My students and I have so far looked at three variations on the discrete exponentiation map.

The discrete Lambert map is the map $x \mapsto xg^x \bmod p$.

$$x \mapsto y \equiv xg^x \bmod p \qquad \text{(Easy)}$$

The assumption that inverting the discrete Lambert map is hard underlies the security of the ElGamal Digital Signature Scheme.

$$x \mapsfrom y \equiv xg^x \bmod p \qquad \text{(Hard?)}$$

JingJing Chen and Mark Lotts started investigating this in 2011.

My students and I have so far looked at three variations on the discrete exponentiation map.

The square discrete exponentiation map is the map $x \mapsto g^{x^2} \bmod p$.

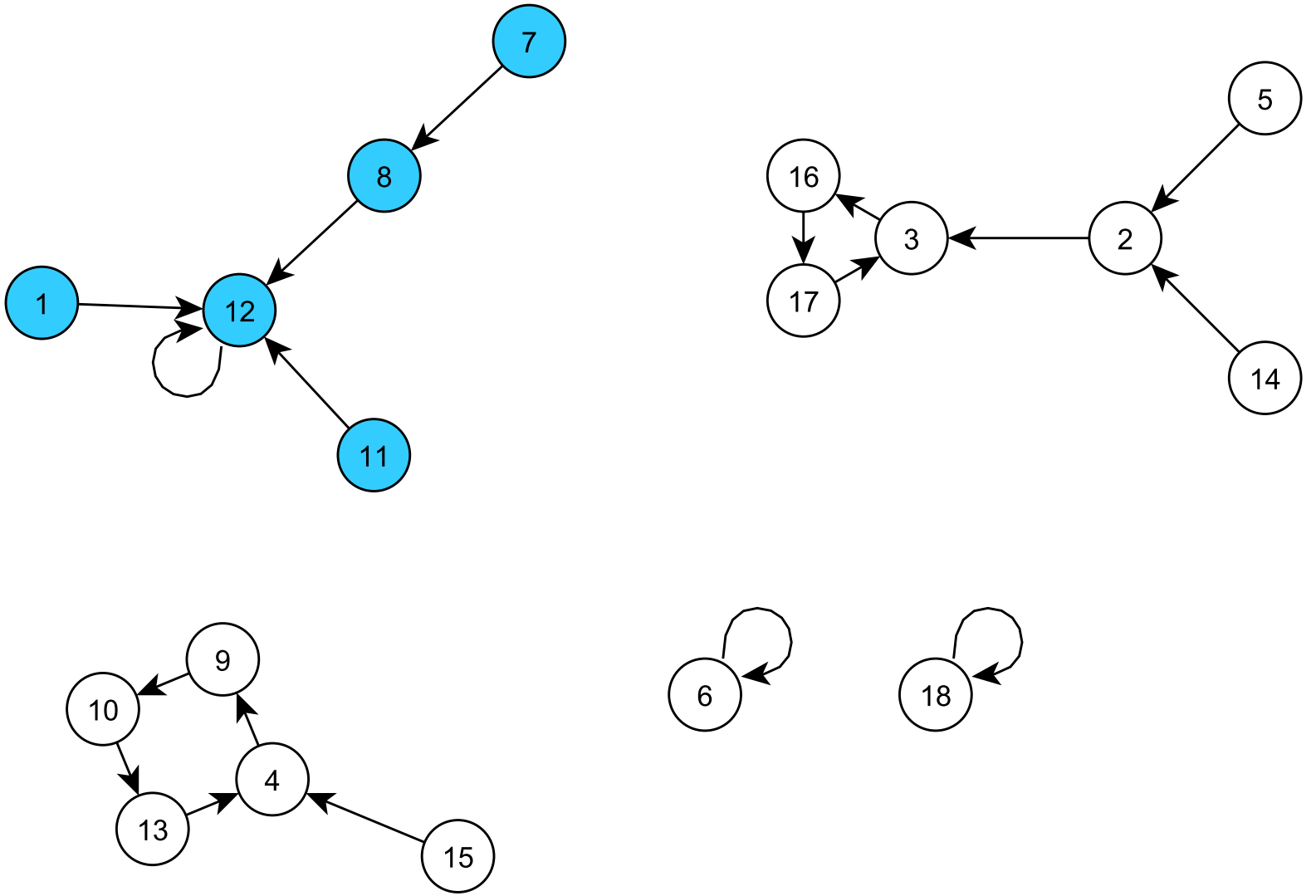
$$x \mapsto y \equiv g^{x^2} \bmod p \qquad \text{(Easy)}$$

The assumption that inverting the square discrete exponentiation map is hard underlies the security of the Camenisch-Stadler group signature scheme.

$$x \hookleftarrow y \equiv g^{x^2} \bmod p \qquad \text{(Hard?)}$$

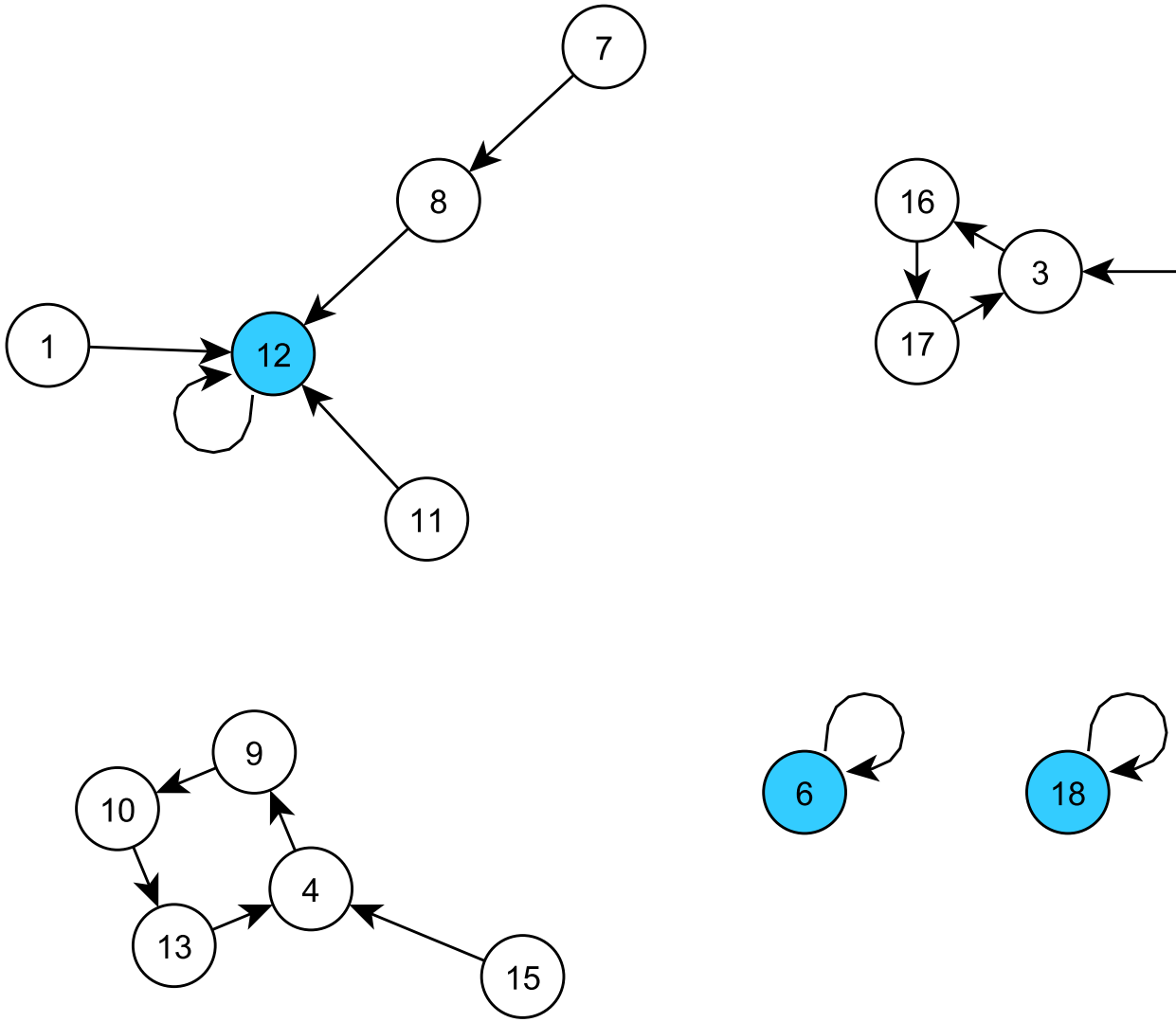
Alex Wood started investigating this in 2011.

There are many graph measurements that we could use to compare our graphs to random graphs.



Example: Number of connected components

There are many graph measurements that we could use to compare our graphs to random graphs.



Example: Number of fixed points

The number of components of a random functional graph has a known formula.

By “random graph” we mean a randomly chosen graph of a specified type on a specified number of nodes.

The number of components of a random functional graph has a known formula.

By “random graph” we mean a randomly chosen graph of a specified type on a specified number of nodes.

Theorem (standard result)

The expected mean number of components in a random function graph of size n is asymptotic to

$$\frac{\ln(2n) + \gamma}{2}$$

as $n \to \infty$.

Furthermore, the number of components of a randomly chosen graph is normally distributed around this mean.

The number of components of a random functional graph has a known formula.

By “random graph” we mean a randomly chosen graph of a specified type on a specified number of nodes.

Theorem (standard result)

The expected mean number of components in a random function graph of size n is asymptotic to

$$\frac{\ln(2n) + \gamma}{2}$$

as $n \to \infty$.

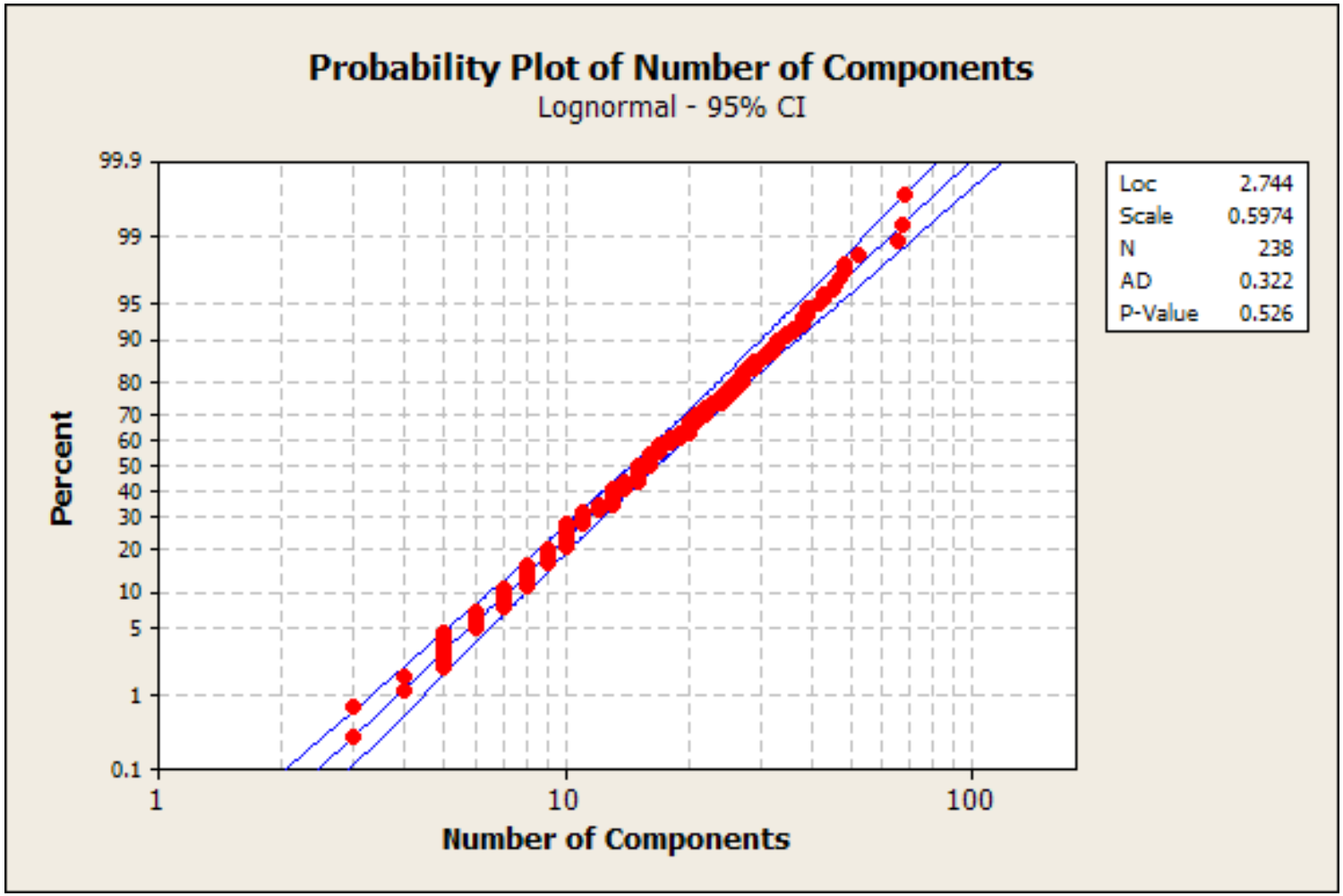
Furthermore, the number of components of a randomly chosen graph is normally distributed around this mean.

Proof: Combinatorics and Analysis

Data on self-power maps was collected for 1090 primes. (Friedrichsen, Larson, and McDowell)

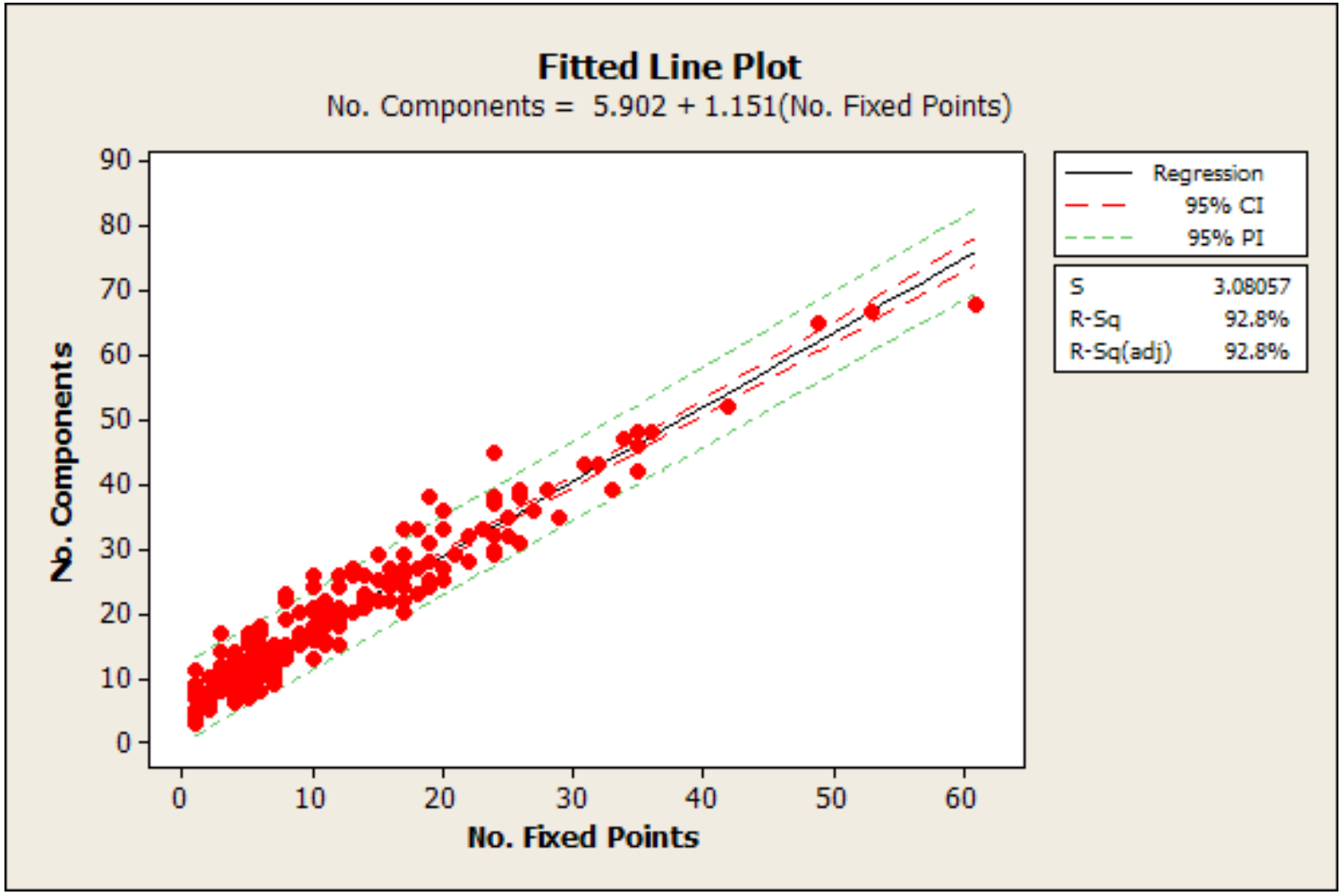
Primes were grouped into two size ranges and also by “safe primes” or not. Statistical tests were done on each category.

Data on self-power maps was collected for 1090 primes. (Friedrichsen, Larson, and McDowell)



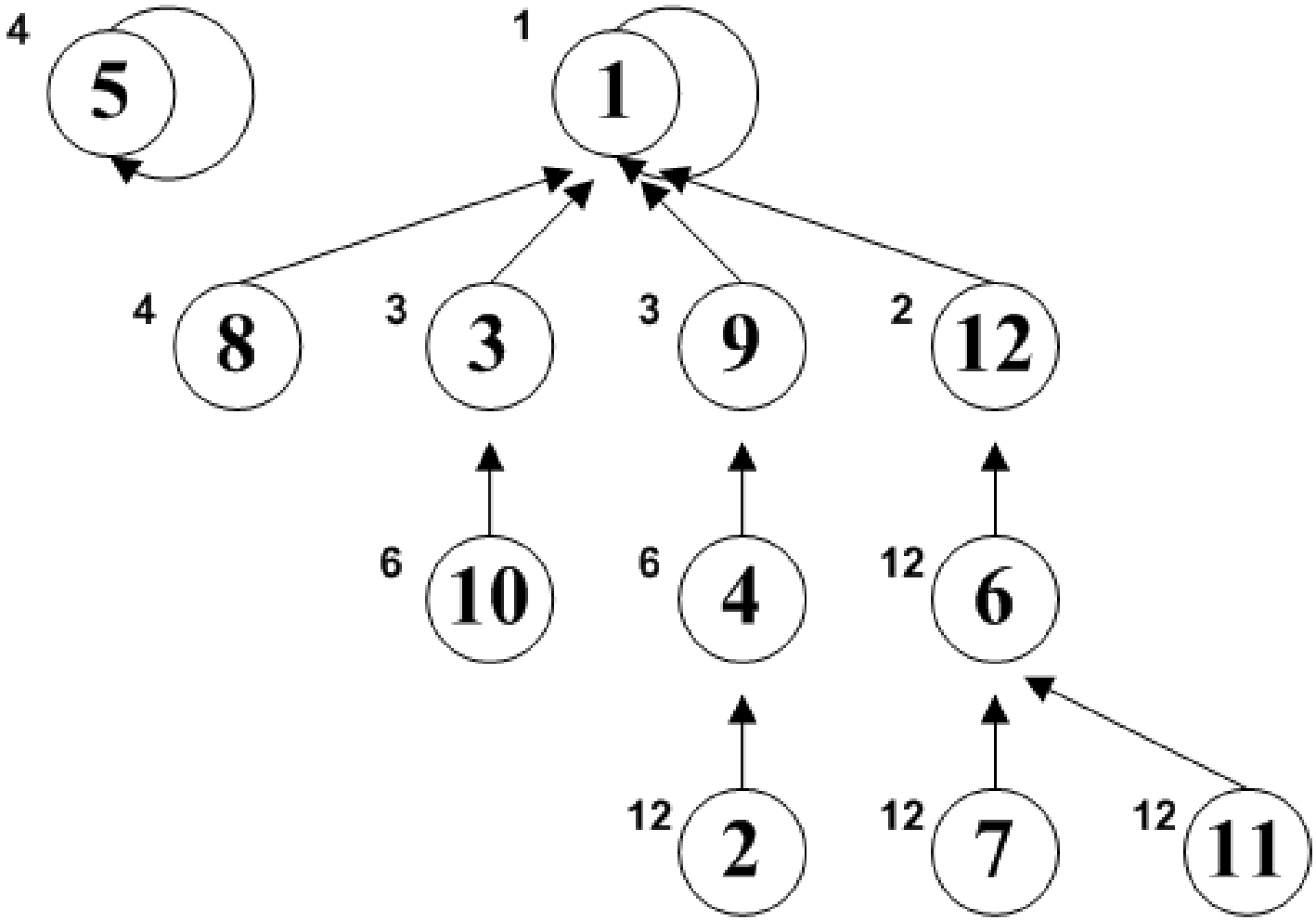
First of several surprises: measurements were not normally distributed — maybe lognormally?

The number of components is closely correlated with the number of fixed points.



Also, the proportion of fixed points to components was significantly larger than expected.

Self-power maps have structure based on order mod p. (Friedrichsen, Larson, and McDowell)



$x \mapsto x^x \bmod 13$, where superscripts denote the order of a node

Further investigation led to the following theorem:

Theorem (Holden, 2012)

Let $F(p)$ be the number of fixed points of the self-power map modulo p. Then

$$\left| F(p) - \sum_{n|p-1} \frac{\phi(n)}{n} \right| \leq d(p-1)^2 \sqrt{p}(1 + \ln p),$$

where $d(p-1)$ is the number of divisors of $p-1$.

Further investigation led to the following theorem:

Theorem (Holden, 2012)

Let $F(p)$ be the number of fixed points of the self-power map modulo p. Then

$$\left| F(p) - \sum_{n|p-1} \frac{\phi(n)}{n} \right| \leq d(p-1)^2 \sqrt{p}(1 + \ln p),$$

where $d(p-1)$ is the number of divisors of $p-1$.

Furthermore, it is known that the values of

$$\sum_{n|p-1} \frac{\phi(n)}{n}$$

are distributed lognormally, so this could explain our findings.

Similar analysis suggests the following conjecture:

Conjecture

The number of components in a self-power functional graph modulo p is approximately $(\ln(p-1))^2/2$ *for "most" p.*

Furthermore, the number of components of a randomly chosen self-power functional graph is lognormally distributed around this value.

Similar analysis suggests the following conjecture:

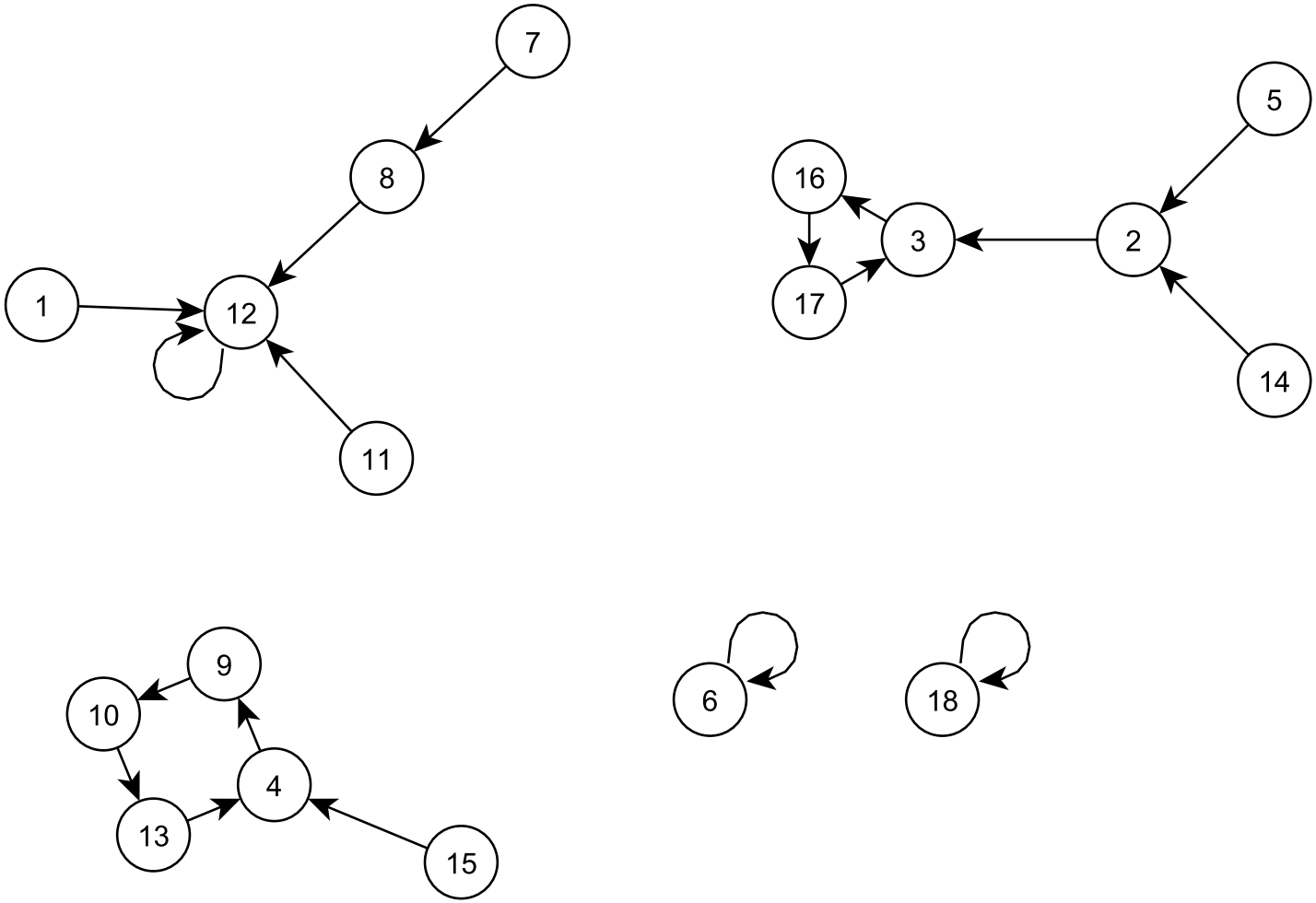
Conjecture

The number of components in a self-power functional graph modulo p is approximately $(\ln(p-1))^2/2$ *for "most" p.*

Furthermore, the number of components of a randomly chosen self-power functional graph is lognormally distributed around this value.

No one has tried to verify this conjecture yet.

Discrete Lambert map graphs have a different sort of structure.



$$x \mapsto x12^x \bmod 19$$

Discrete Lambert map graphs have a different sort of structure.

Definition

The *cosets* of $\{1, \ldots, p-1\}$ generated by g are the sets of the form $\{x, xg, xg^2, \ldots\}$ where are multiplication is modulo p.

- Any two cosets are the same size and two cosets are either equal or disjoint.
- We can predict exactly which g's (and how many) result in what number of cosets.

Discrete Lambert map graphs have a different sort of structure.

Definition

The *cosets* of $\{1, \ldots, p-1\}$ generated by g are the sets of the form $\{x, xg, xg^2, \ldots\}$ where are multiplication is modulo p.

- Any two cosets are the same size and two cosets are either equal or disjoint.
- We can predict exactly which g's (and how many) result in what number of cosets.

Example

Let $p = 19$ and $g = 12$.

$x = 1$	1	12	11	18	7	8
$x = 2$	2	5	3	17	14	16
$x = 4$	4	10	6	15	9	13

Discrete Lambert map graphs have a different sort of structure.

Definition

The *cosets* of $\{1, \ldots, p-1\}$ generated by g are the sets of the form $\{x, xg, xg^2, \ldots\}$ where are multiplication is modulo p.

- Any two cosets are the same size and two cosets are either equal or disjoint.
- We can predict exactly which g's (and how many) result in what number of cosets.

Theorem (Chen and Lotts)

If two nodes of the graph of $x \mapsto xg^x \bmod p$ *are connected, then they belong to the same coset generated by* g*.*

Discrete Lambert map graphs have a different sort of structure.

Definition

The *cosets* of $\{1, \ldots, p-1\}$ generated by g are the sets of the form $\{x, xg, xg^2, \ldots\}$ where are multiplication is modulo p.

- Any two cosets are the same size and two cosets are either equal or disjoint.
- We can predict exactly which g's (and how many) result in what number of cosets.

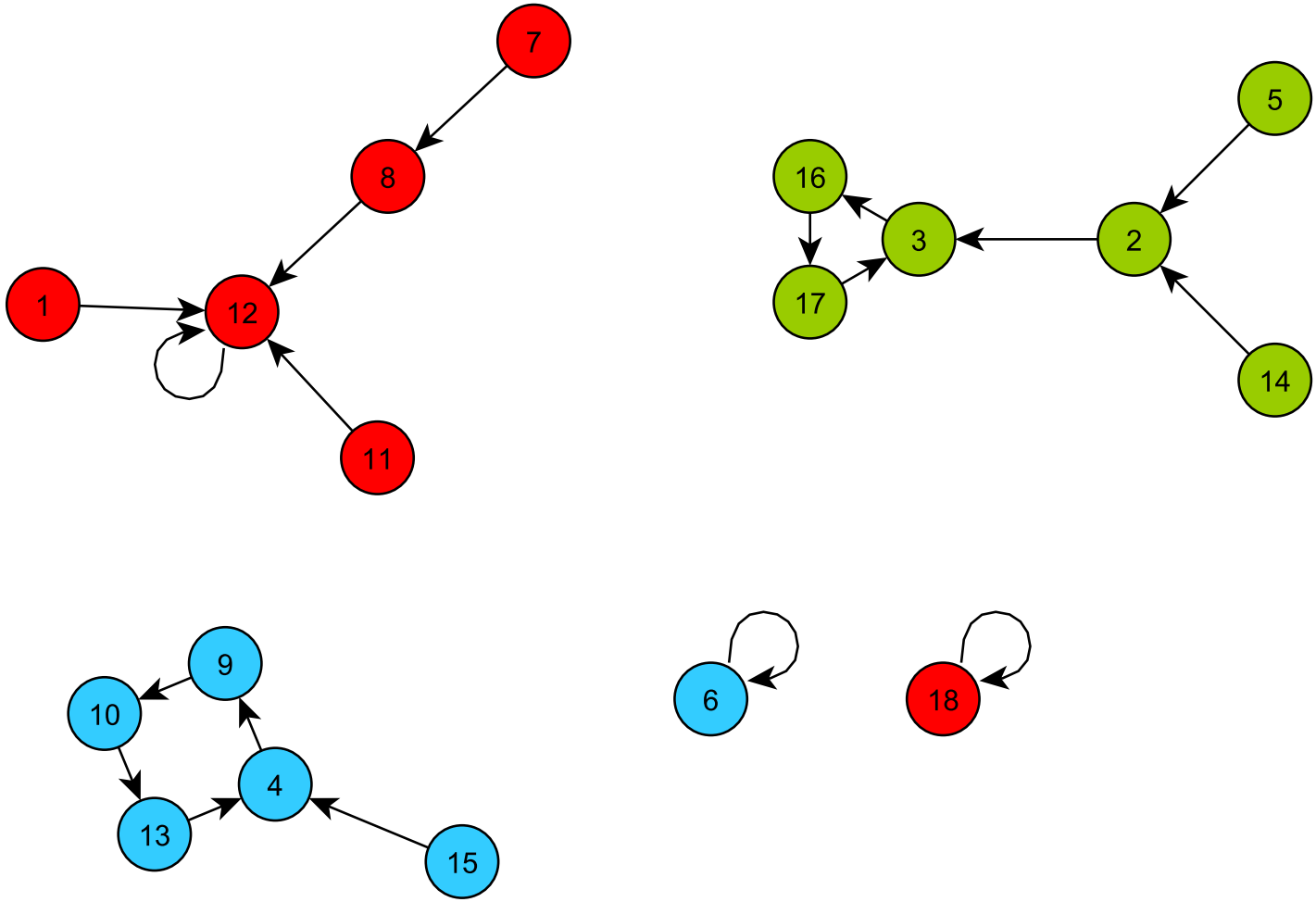
Theorem (Chen and Lotts)

If two nodes of the graph of $x \mapsto xg^x \bmod p$ are connected, then they belong to the same coset generated by g.

Corollary

The cosets generated by g partition the graph into disconnected subgraphs.

The cosets generated by g partition the graph into disconnected subgraphs.



$$x \mapsto x12^x \bmod 19$$

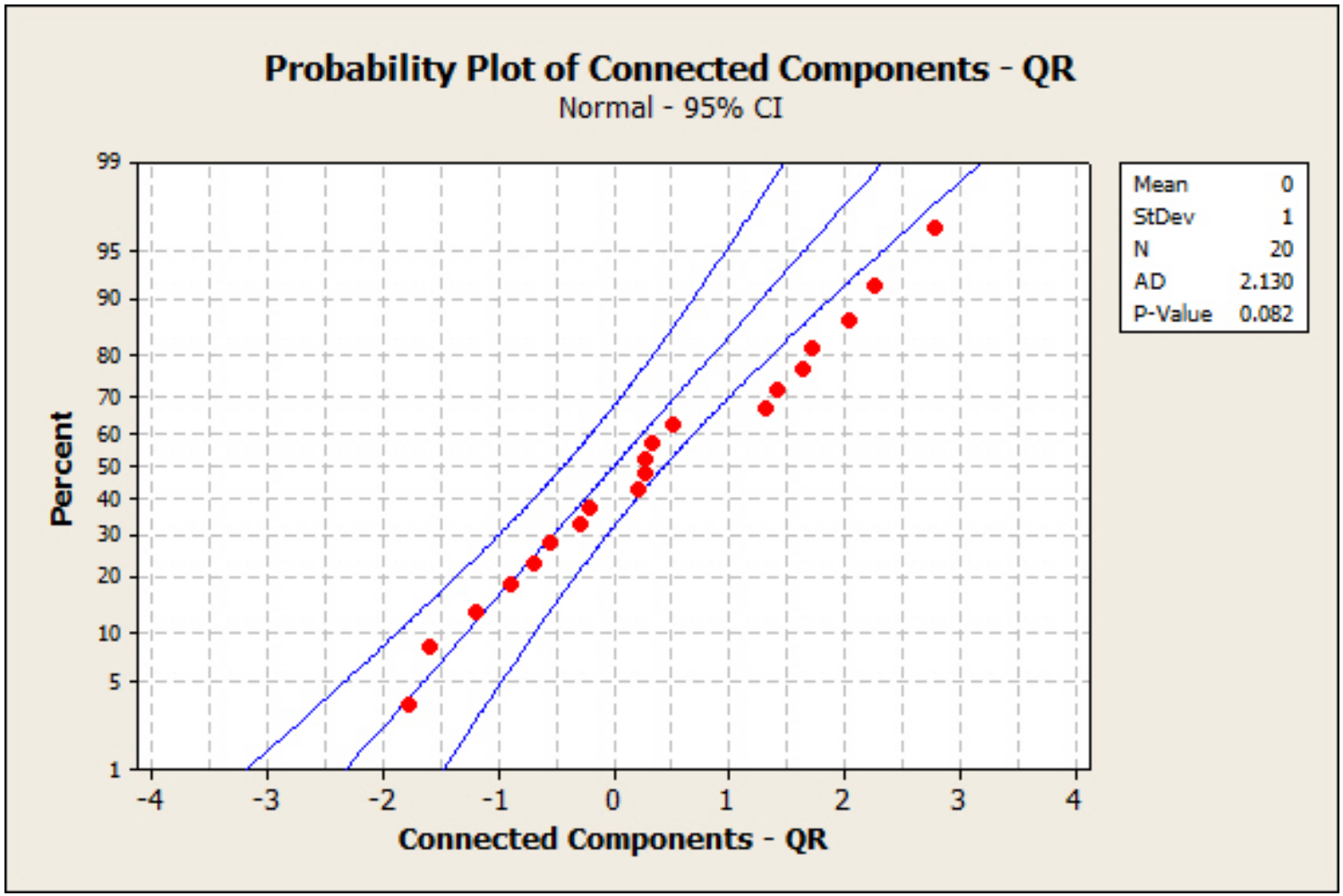
Data on discrete Lambert maps was collected for 20 safe primes greater than 40,000. (Chen and Lotts)

For each of these primes, the graph was generated for each of the $p - 3$ values of g generating either 1 or 2 cosets.

Statistical tests were done separately for the 1-coset and 2-coset groups.

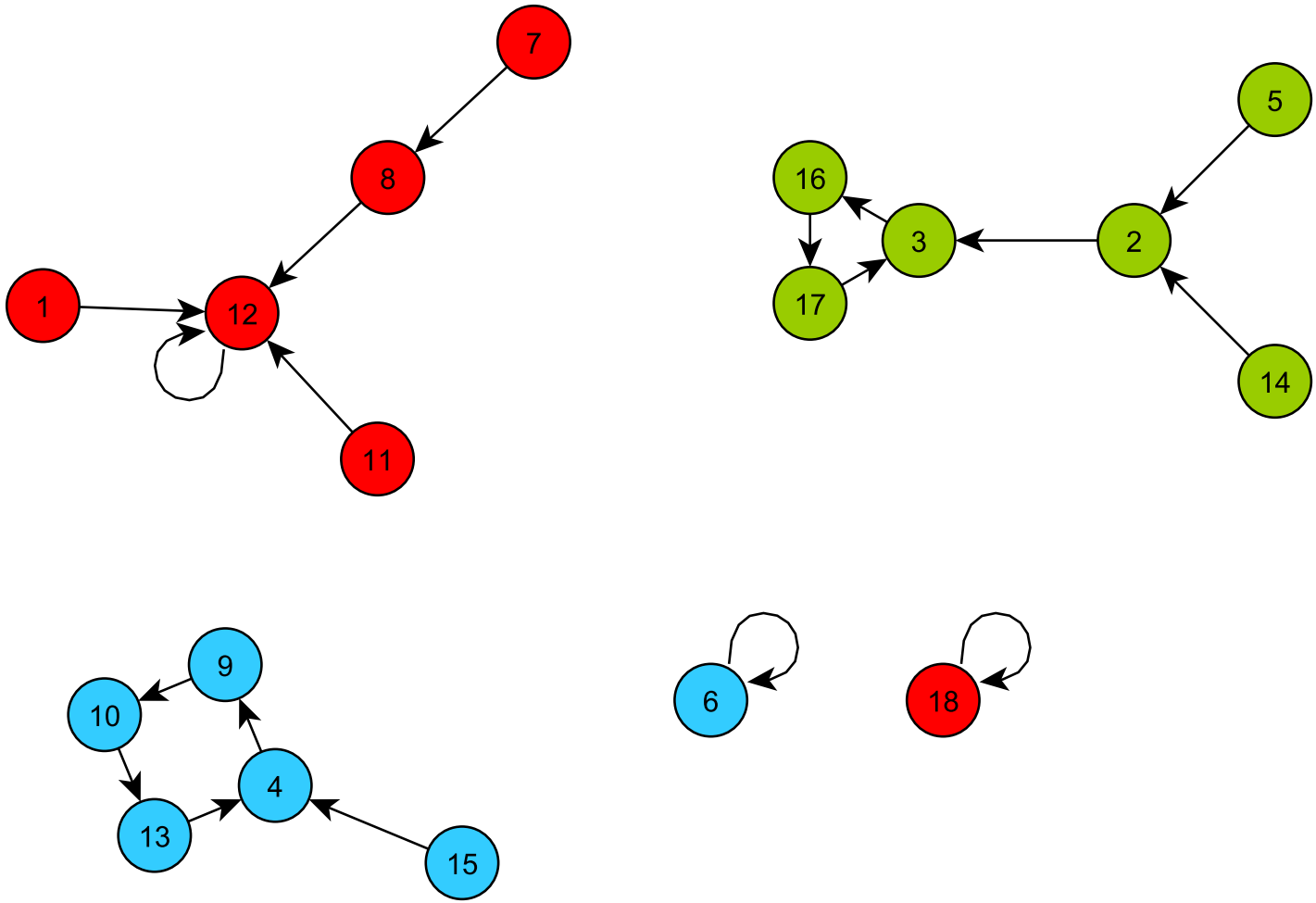
The expected number of connected components needed to be adjusted (upwards) for the 2-coset graphs, treating each coset as a random graph.

Data on discrete Lambert maps was collected for 20 safe primes greater than 40,000. (Chen and Lotts)



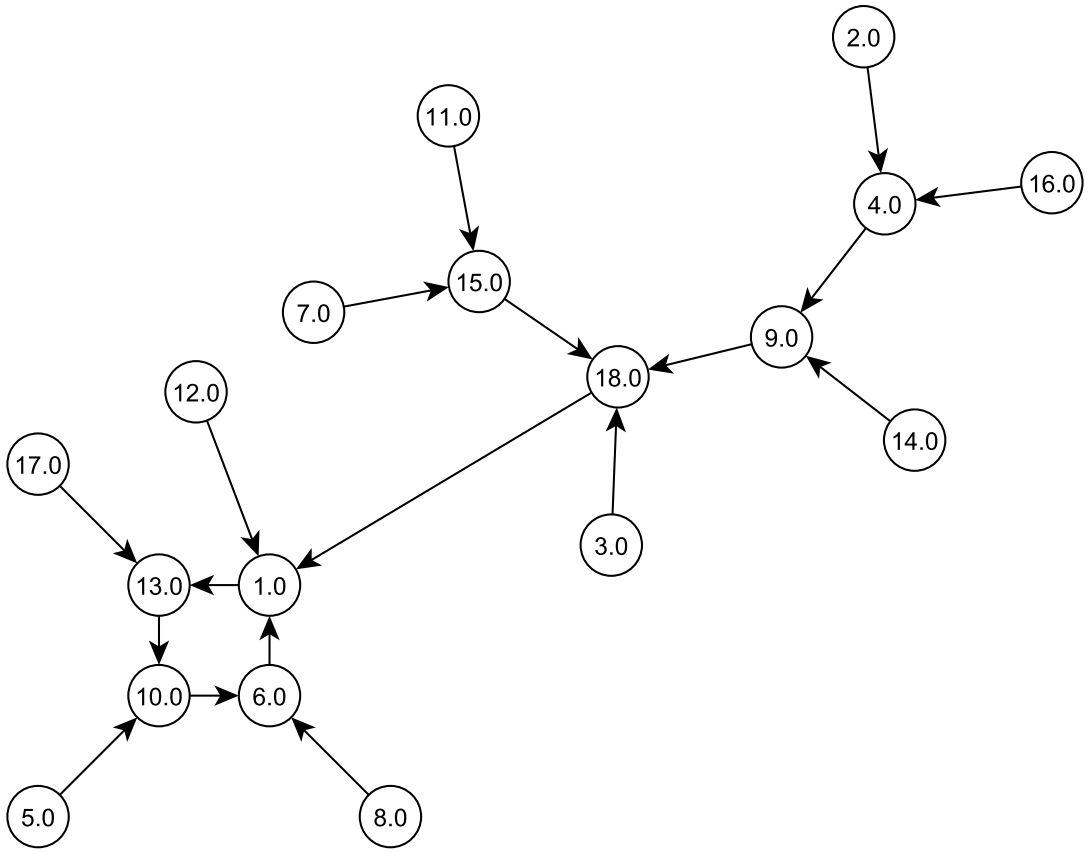
This gave good results for the “additive” statistics.
(For example, number of connected components.)

Dealing with the “maximum” statistics will require more work.



Given the expected maximum cycle size for each coset, what is the expected maximum cycle size for the graph?

Square discrete exponentiation map graphs have structure in their in-degrees, as was shown by Wood.



$$x \mapsto 13^{x^2} \bmod 19$$

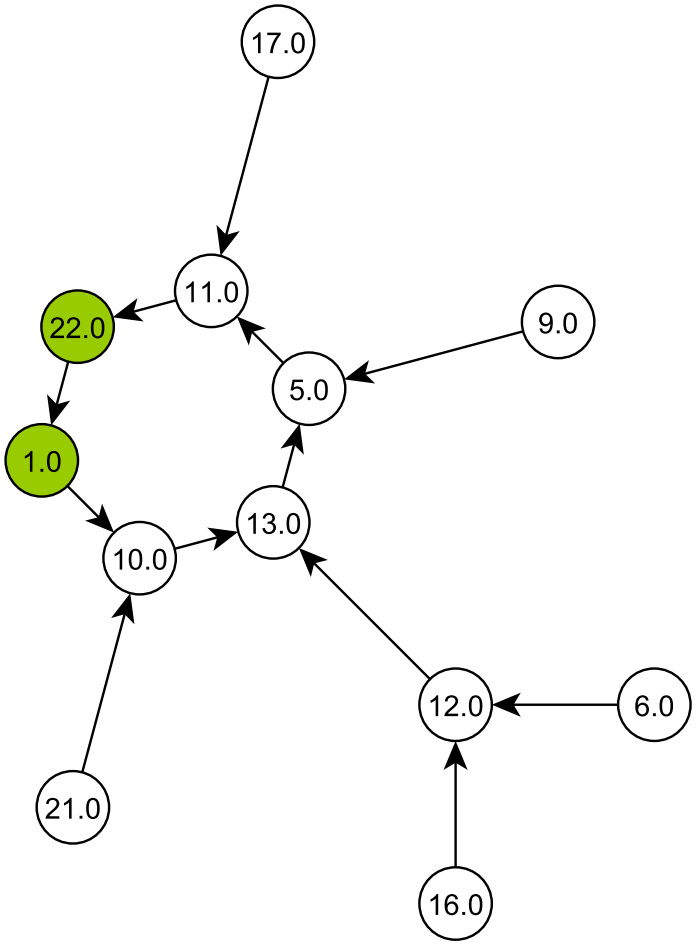
Square discrete exponentiation map graphs have structure in their in-degrees, as was shown by Wood.

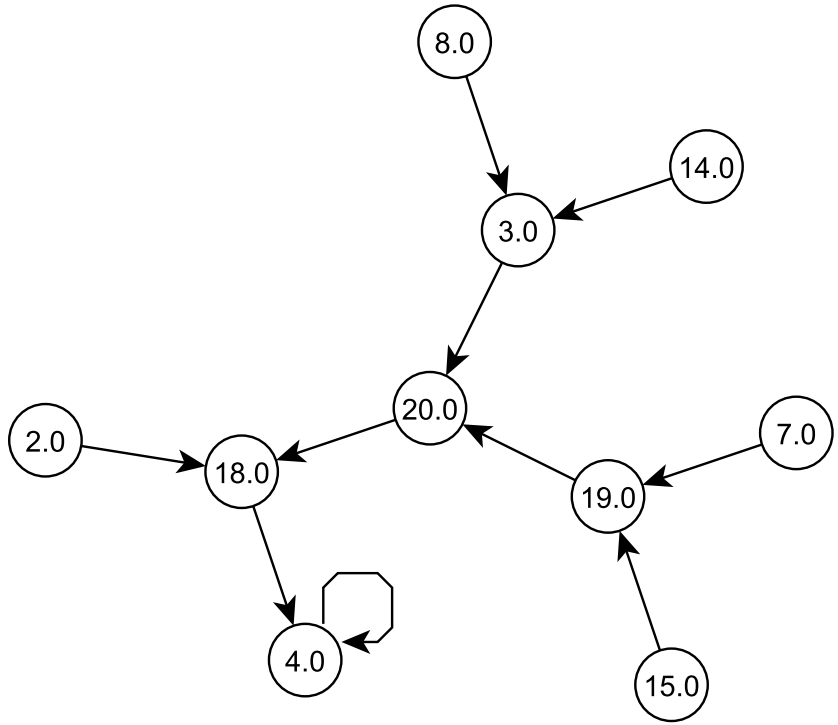
x\g	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
2	1	16	5	9	17	4	7	11	6	6	11	7	4	17	9	5	16	1
3	1	18	18	1	1	1	1	18	1	18	1	18	18	18	18	1	1	18
4	1	5	17	6	16	9	7	11	4	4	11	7	9	16	6	17	5	1
5	1	14	2	6	16	9	7	8	4	15	11	12	10	3	13	17	5	18
6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
7	1	3	14	9	17	4	7	8	6	13	11	12	15	2	10	5	16	18
8	1	17	16	4	5	6	7	11	9	9	11	7	6	5	4	16	17	1
9	1	18	18	1	1	1	1	18	1	18	1	18	18	18	18	1	1	18
10	1	17	16	4	5	6	7	11	9	9	11	7	6	5	4	16	17	1
11	1	3	14	9	17	4	7	8	6	13	11	12	15	2	10	5	16	18
12	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
13	1	14	2	6	16	9	7	8	4	15	11	12	10	3	13	17	5	18
14	1	5	17	6	16	9	7	11	4	4	11	7	9	16	6	17	5	1
15	1	18	18	1	1	1	1	18	1	18	1	18	18	18	18	1	1	18
16	1	16	5	9	17	4	7	11	6	6	11	7	4	17	9	5	16	1
17	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
18	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

The number of times y appears in a column is the in-degree of y in the map

$$x \mapsto g^{x^2} \bmod 19.$$

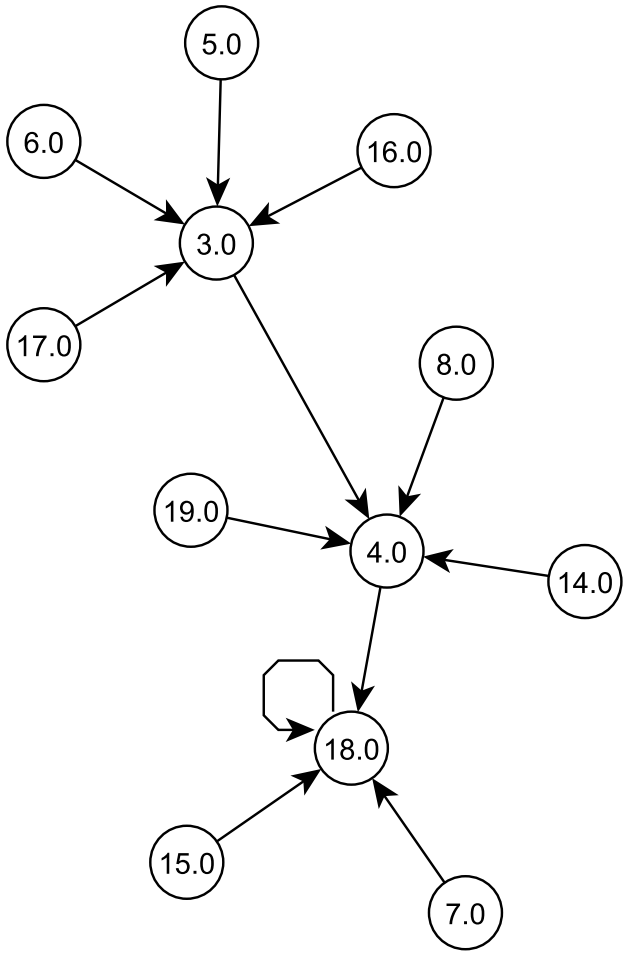
The next step is to look at random graphs of the "types" that the structure predicts.

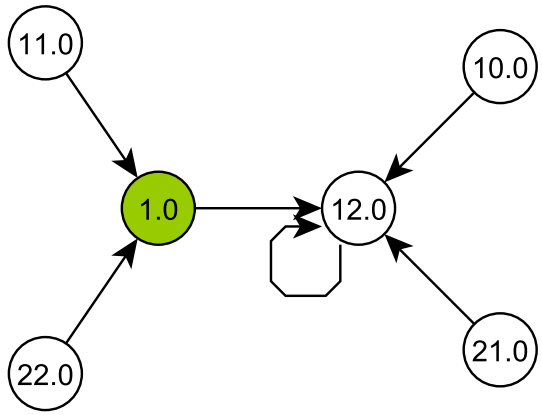


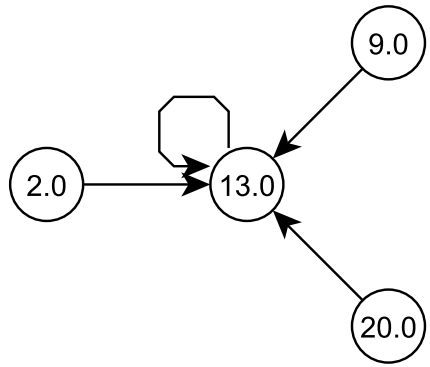


$x \mapsto 10^{x^2} \bmod 23$ is "almost binary" — all in-degrees equal either 0 or 2 except for two consecutive nodes of in-degree 1.

The next step is to look at random graphs of the "types" that the structure predicts.







$x \mapsto 12^{x^2} \bmod 23$ is "almost quaternary" — all in-degrees equal either 0 or 4 except for one node of in-degree 2.

We proposed to use a general technique which was given by Flajolet and Odlyzko:

- Explicitly define the structure.
- Convert to exponential generating functions.
- “Mark” the structures of interest.
- Compute expected value generating functions.
- Perform “automatic” singularity analysis to get asymptotic form of coefficients.
- Normalize.

We have gotten through the first two steps of this plan.

There's a lot of possibilities for future work on variations of the DLP.

- More investigation of the "types" of graphs that come up
- Predictions of more of the expected measurements
- More data collection and statistical tests
- Other maps, e.g. the "exponential Welch permutation" $x \mapsto g^{x-1+c} \bmod p$, for fixed c. (Electrical Engineering)
- "Multi-maps" such as $x \bmod p \mapsto x^x \bmod p$
- Composite moduli, e.g. prime powers, RSA numbers
- Elliptic curves, finite fields, other groups?

Thanks!

More information at:

`http://www.rose-hulman.edu/~holden/REU`