Sub-dominant Cogrowth properties and the ERR algorithm

Murray Elder and Cameron Rogers, The University of Newcastle

Odense, 2016

Recall

Given a presentation, the ERR algorithm performs a random walk on the space of trivial words. Transition probabilities are defined such that the walk converges to the stationary distribution

$$\pi(w) = \frac{(|w| + 1)^{\alpha+1}\beta^{|w|}}{Z}.$$

Recall

Given a presentation, the ERR algorithm performs a random walk on the space of trivial words. Transition probabilities are defined such that the walk converges to the stationary distribution

$$\pi(w) = \frac{(|w|+1)^{\alpha+1}\beta^{|w|}}{Z}.$$

So the probability of arriving at a given word length is

$$Pr(@n) = c_n \frac{(n+1)^{\alpha+1}\beta^{n}}{Z}.$$

BS(1,N): a case study

Recall the pathological behaviour observed for
$BS(1,N) = \langle a, t \mid tat^{-1}a^{-N} \rangle$:

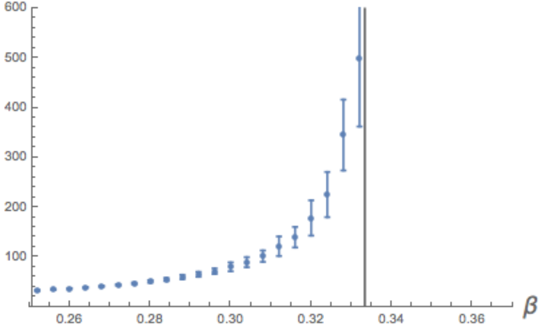


Figure: Results for ERR method on $BS(1,2)$

BS(1,N): a case study

Recall the pathological behaviour observed for
$BS(1, N) = \langle a, t \mid tat^{-1}a^{-N} \rangle$:

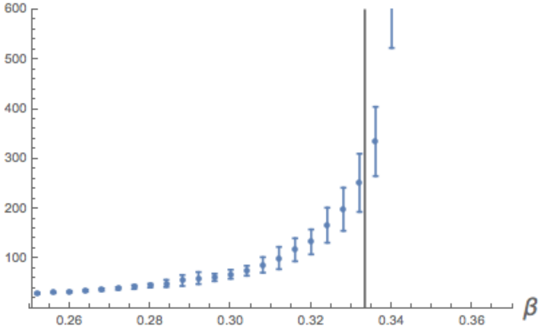


Figure: Results for ERR method on $BS(1, 3)$

BS(1,N): a case study

Recall the pathological behaviour observed for $BS(1,N) = \langle a, t \mid tat^{-1}a^{-N} \rangle$:

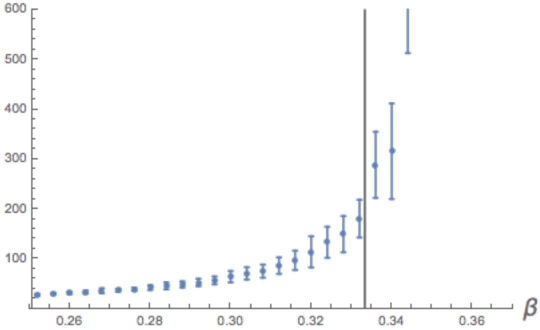


Figure: Results for ERR method on $BS(1,4)$

BS(1,N): a case study

Recall the pathological behaviour observed for $BS(1,N) = \langle a, t \mid tat^{-1}a^{-N} \rangle$:

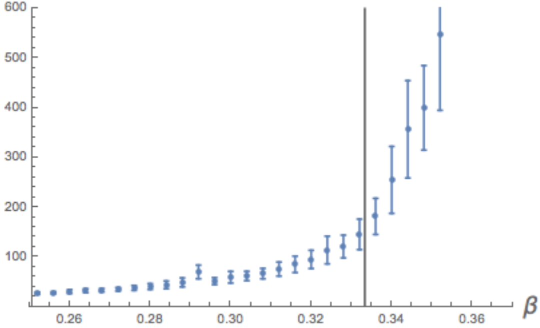


Figure: Results for ERR method on $BS(1,5)$

BS(1,N): a case study

Recall the pathological behaviour observed for
$BS(1,N) = \langle a, t \mid tat^{-1}a^{-N} \rangle$:

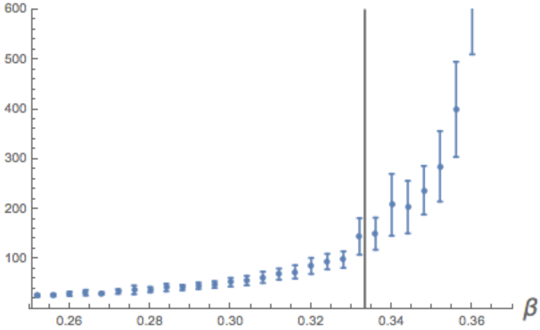


Figure: Results for ERR method on $BS(1,6)$

BS(1,N): a case study

Recall the pathological behaviour observed for $BS(1,N) = \langle a, t \mid tat^{-1}a^{-N} \rangle$:

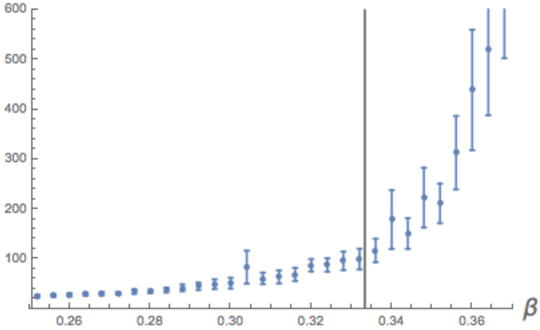


Figure: Results for ERR method on $BS(1,7)$

BS(1,N): a case study

Recall the pathological behaviour observed for
$BS(1,N) = \langle a, t \mid tat^{-1}a^{-N} \rangle$:

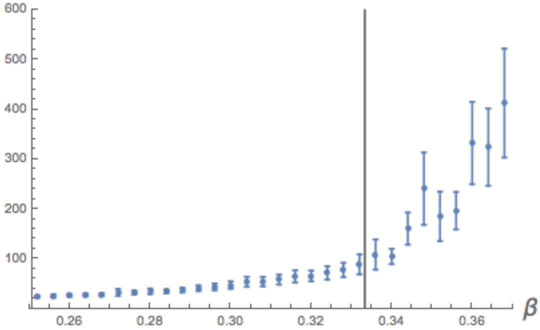


Figure: Results for ERR method on $BS(1,8)$

Questions to Answer

1. Why, for the groups $BS(1, N)$, does the accuracy of the ERR method decrease as N increases?

Why?

The ERR method only records the average word length of words visited. To unpack this pathological behaviour we must look at the distribution as a whole.

Why?

The ERR method only records the average word length of words visited. To unpack this pathological behaviour we must look at the distribution as a whole.

What should

$$Pr(@n) = c_n \frac{(n+1)^{\alpha+1}\beta^n}{Z}$$

look like?

When $\beta > \beta_c$

Recall, for an amenable group c_n grows like $(|S| - 1)^n$.

When $\beta > \beta_c$

Recall, for an amenable group c_n grows like $(|S| - 1)^n$.
β exceeds β_c precisely when

$$(|S| - 1)\beta > 1.$$

Hence the distribution

$$Pr(@n) = c_n \frac{(n+1)^{\alpha+1}\beta^n}{Z}$$

grows exponentially.

When $\beta < \beta_c$

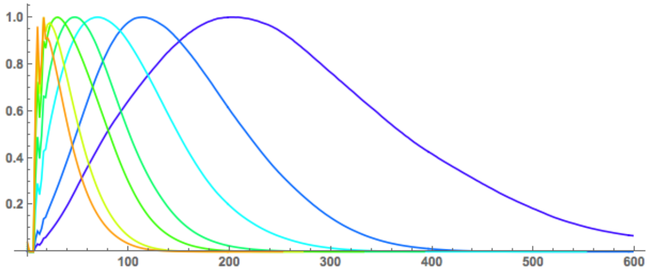
When $\beta < \beta_c$ the polynomial term in

$$Pr(@n) = c_n \frac{(n+1)^{\alpha+1}\beta^n}{Z}$$

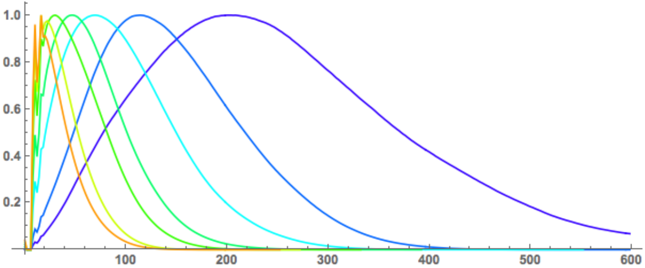
may cause a peak at small initial word lengths, but the exponential decay from the other terms will eventually dominate.

As β approaches β_c the peak moves toward longer word lengths, and the standard deviation of the distribution increases.

When $\beta < \beta_c$

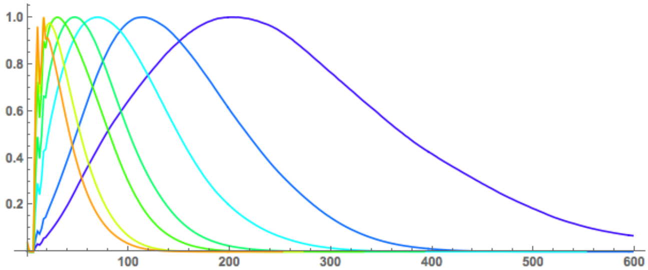


When $\beta < \beta_c$



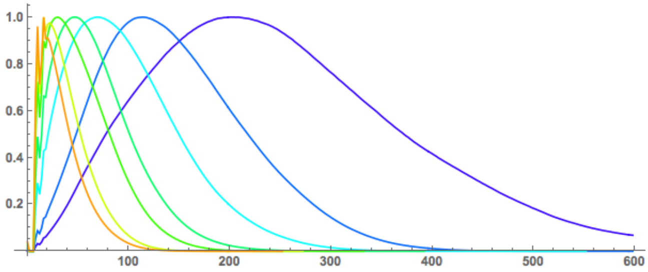
This is a figure of the distributions of ERR random walks on $BS(1,7)$ where $\alpha = 3$ and

?????



This is a figure of the distributions of ERR random walks on $BS(1,7)$ where $\alpha = 3$ and $\beta = 0.335,\ 0.34,\ 0.345,\ 0.35,\ 0.355,\ 0.36,\ 0.365$

?????



This is a figure of the distributions of ERR random walks on $BS(1,7)$ where $\alpha = 3$ and $\beta = 0.335,\ 0.34,\ 0.345,\ 0.35,\ 0.355,\ 0.36,\ 0.365$

These random walks certainly do not seem to converge to the predicted stationary distribution.

Questions to Answer

1. Why, for the groups $BS(1, N)$, does the accuracy of the ERR method decrease as N increases?

Questions to Answer

1. Why, for the groups $BS(1, N)$, does the accuracy of the ERR method decrease as N increases?
2. What group properties does this reflect, and what is the mechanism by which they slow convergence?

Questions to Answer

1. Why, for the groups $BS(1, N)$, does the accuracy of the ERR method decrease as N increases?
2. What group properties does this reflect, and what is the mechanism by which they slow convergence?
3. Are the walks "seeing" a representative sample of the space of trivial words?

The ERR-R algorithm

Instead of estimating the asymptotic growth rate, here is a way to get estimates from an ERR random walk for *initial terms* of the cogrowth sequence.

The ERR-R algorithm

The probability that a long walk will arrive at a word of length n is

$$Pr(@n) = c_n \cdot \frac{(n+1)^{1+\alpha}\beta^n}{Z}$$

The ERR-R algorithm

The probability that a long walk will arrive at a word of length n is

$$Pr(@n) = c_n \cdot \frac{(n+1)^{1+\alpha}\beta^n}{Z}$$

We can run the random walk , say Y steps, and let $W(n)$ be the number of times it hits a word of length n. Then

$$Pr(@n) \approx W(n)/Y$$

The ERR-R algorithm

The probability that a long walk will arrive at a word of length n is

$$Pr(@n) = c_n \cdot \frac{(n+1)^{1+\alpha}\beta^n}{Z}$$

We can run the random walk , say Y steps, and let $W(n)$ be the number of times it hits a word of length n. Then

$$Pr(@n) \approx W(n)/Y$$

Now for $m > n$, we have

$$\frac{W(n)/Y}{W(m)/Y} \approx \frac{c_n}{c_m} \frac{(n+1)^{1+\alpha}\beta^n/Z}{(m+1)^{1+\alpha}\beta^m/Z}$$

The ERR-R algorithm

The probability that a long walk will arrive at a word of length n is

$$Pr(@n) = c_n \cdot \frac{(n+1)^{1+\alpha}\beta^n}{Z}$$

We can run the random walk , say Y steps, and let $W(n)$ be the number of times it hits a word of length n. Then

$$Pr(@n) \approx W(n)/Y$$

Now for $m > n$, we have

$$\frac{W(n)/Y}{W(m)/Y} \approx \frac{c_n}{c_m}\frac{(n+1)^{1+\alpha}\beta^n/Z}{(m+1)^{1+\alpha}\beta^m/Z}$$

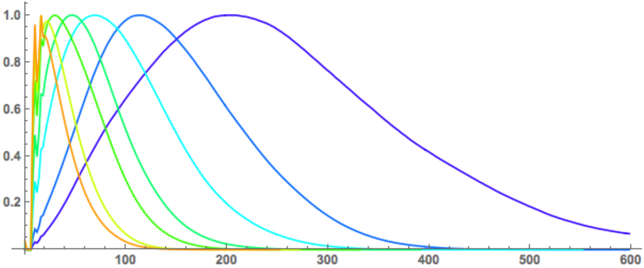
so

$$c_m \approx c_n \left(\frac{n+1}{m+1}\right)^{1+\alpha}\beta^{n-m}\frac{W(m)}{W(n)}.$$

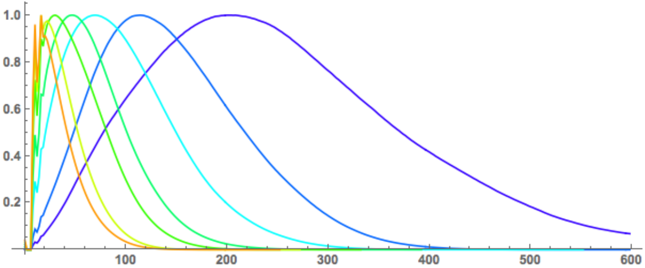
Note, if the ERR walk has not successfully converged to the stationary distribution then this will not give accurate estimates for c_m;

Note, if the ERR walk has not successfully converged to the stationary distribution then this will not give accurate estimates for c_m; so what happens with $BS(1,7)$?

Note, if the ERR walk has not successfully converged to the stationary distribution then this will not give accurate estimates for c_m; so what happens with $BS(1,7)$?



Note, if the ERR walk has not successfully converged to the stationary distribution then this will not give accurate estimates for c_m; so what happens with $BS(1,7)$?



$c(8) = 16$, what does the data predict?

Estimates for c_8

Each random walk represented was $\approx 10^{11}$ steps long.

β	$W(0)$	$W(8)$	estimate
0.335	1509336558	25134566940	16.0014
0.34	843416499	15813088349	16.002
0.345	417932474	8799949864	15.9901
0.35	170402696	4030830678	16.0105
0.355	61441282	1624947099	15.9802
0.36	10256919	303470462	15.9849
0.365	1660139	55189877	16.0843

Questions to Answer

1. Why, for the groups $BS(1, N)$, does the accuracy of the ERR method decrease as N increases?

Questions to Answer

1. Why, for the groups $BS(1, N)$, does the accuracy of the ERR method decrease as N increases?
2. What group properties does this reflect, and what is the mechanism by which they slow convergence?

Questions to Answer

1. Why, for the groups $BS(1, N)$, does the accuracy of the ERR method decrease as N increases?
2. What group properties does this reflect, and what is the mechanism by which they slow convergence?
3. Are the walks "seeing" a representative sample of the space of trivial words?

Questions to Answer

1. Why, for the groups $BS(1, N)$, does the accuracy of the ERR method decrease as N increases?
2. What group properties does this reflect, and what is the mechanism by which they slow convergence?
3. Are the walks "seeing" a representative sample of the space of trivial words?
4. Why do these (faulty) ERR walks consistently produce accurate estimates for initial cogrowth values?

????

Definition

Let G be a finitely generated group and let c_n denote the number of trivial words of length n equal to the identity. Then let

$$\mathcal{R}(n) = \min \left\{ k \ : \ \frac{c_{2k+2}}{c_{2k}} > (|S| - 1)^2 - \frac{1}{n} \right\}.$$

Murray gave examples of $\mathcal{R}(n)$ for different groups in his talk.

Sub-dominant cogrowth behaviour

Example

Suppose a group had a cogrowth function given by $c(n) = 3^{n-pn^{1/3}}$*.* *($BS(1,N)$ have a cogrowth function which is asymptotically equivalent to this form, and* p *increases with* N*). As* p *increases we increase the potency of the sub-dominant cogrowth behaviour and* $\mathcal{R}(n)$ *grows faster.*

Sub-dominant cogrowth behaviour

Example

Suppose a group had a cogrowth function given by $c(n) = 3^{n - pn^{1/3}}$*.*
($BS(1, N)$ *have a cogrowth function which is asymptotically equivalent to this form, and* p *increases with* N*). As* p *increases we increase the potency of the sub-dominant cogrowth behaviour and* $\mathcal{R}(n)$ *grows faster.*

What happens to $Pr(@n)$ *as* p *changes?*

Sub-dominant cogrowth behaviour

The expected distribution for an ERR random walk with $\alpha = 4$ and $\beta = 0.335$:

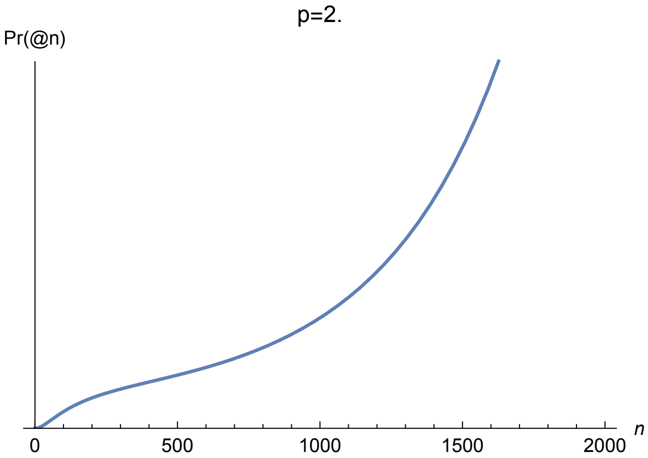


Figure: $Pr(@n) = 3^{n-pn^{1/3}}(1+n)^4(0.335)^n$

Sub-dominant cogrowth behaviour

The expected distribution for an ERR random walk with $\alpha = 3$ and $\beta = 0.335$:

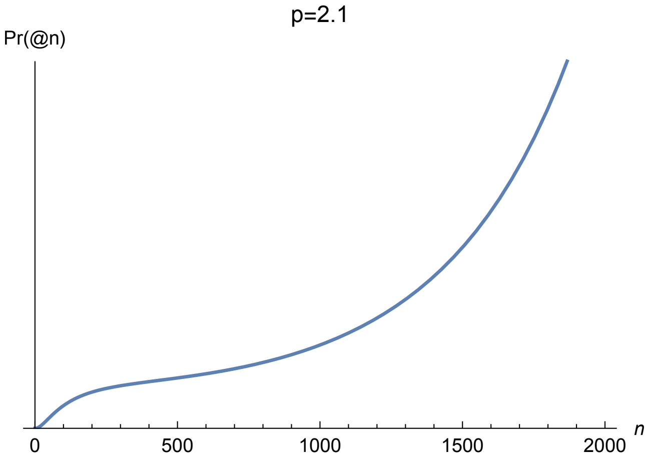


Figure: $Pr(@n) = 3^{n-pn^{1/3}}(1+n)^4(0.335)^n$

Sub-dominant cogrowth behaviour

The expected distribution for an ERR random walk with $\alpha = 3$ and $\beta = 0.335$:

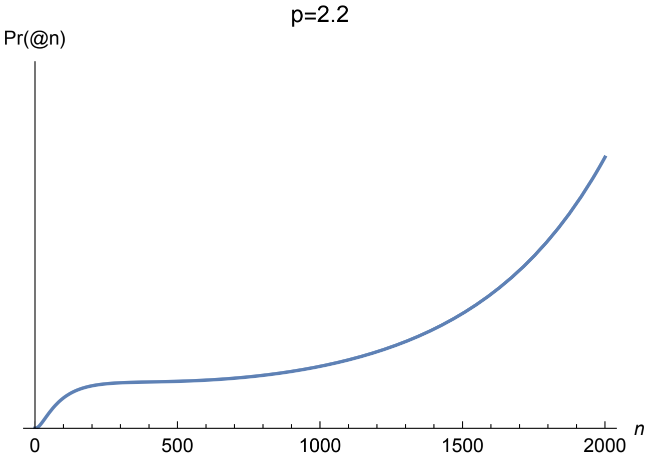


Figure: $Pr(@n) = 3^{n-pn^{1/3}}(1+n)^4(0.335)^n$

Sub-dominant cogrowth behaviour

The expected distribution for an ERR random walk with $\alpha = 3$ and $\beta = 0.335$:

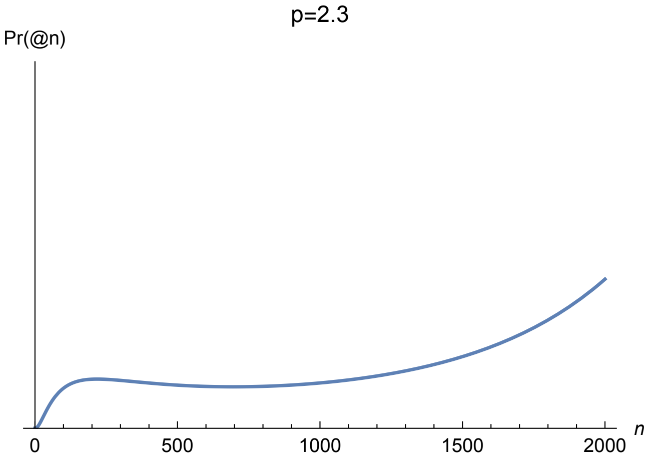


Figure: $Pr(@n) = 3^{n-pn^{1/3}}(1+n)^4(0.335)^n$

Sub-dominant cogrowth behaviour

The expected distribution for an ERR random walk with $\alpha = 3$ and $\beta = 0.335$:

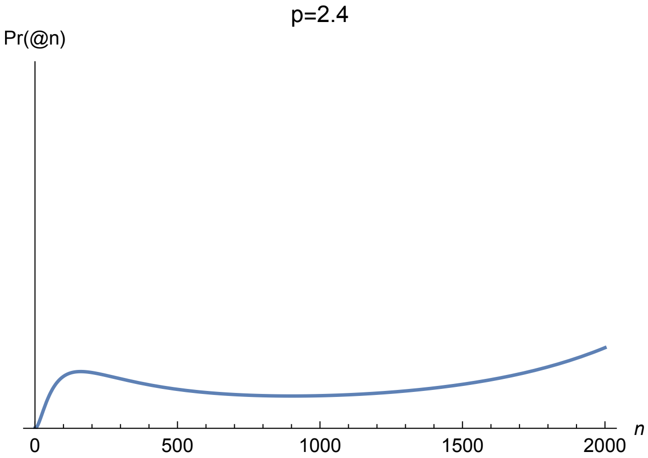


Figure: $Pr(@n) = 3^{n-pn^{1/3}}(1+n)^4(0.335)^n$

Sub-dominant cogrowth behaviour

The expected distribution for an ERR random walk with $\alpha = 3$ and $\beta = 0.335$:

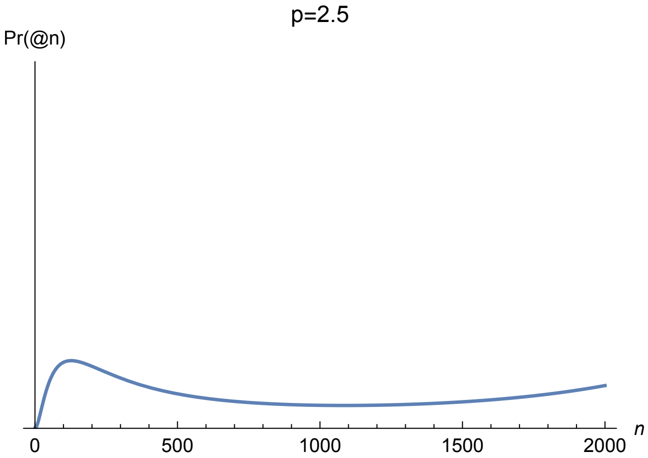


Figure: $Pr(@n) = 3^{n-pn^{1/3}}(1+n)^4(0.335)^n$

Sub-dominant cogrowth behaviour

The expected distribution for an ERR random walk with $\alpha = 3$ and $\beta = 0.335$:

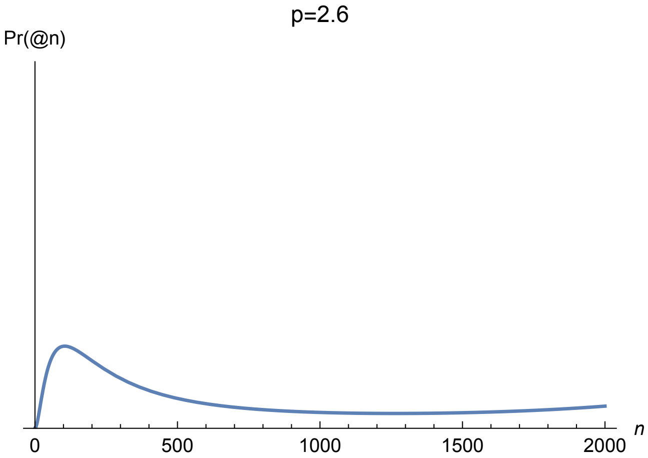


Figure: $Pr(@n) = 3^{n-pn^{1/3}}(1+n)^4(0.335)^n$

Sub-dominant cogrowth behaviour

The expected distribution for an ERR random walk with $\alpha = 3$ and $\beta = 0.335$:

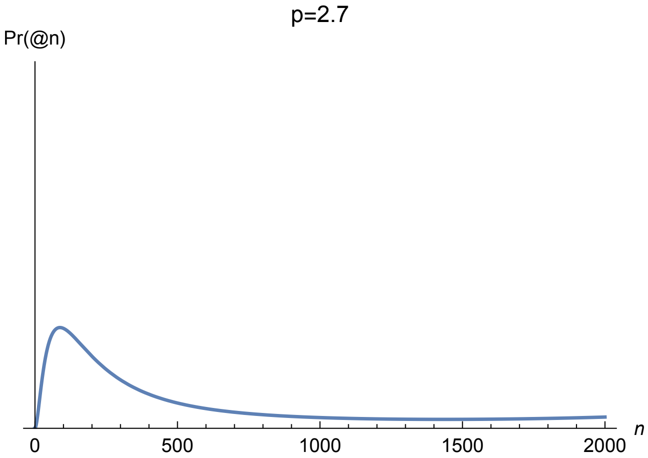


Figure: $Pr(@n) = 3^{n-pn^{1/3}}(1+n)^4(0.335)^n$

Sub-dominant cogrowth behaviour

The expected distribution for an ERR random walk with $\alpha = 3$ and $\beta = 0.335$:

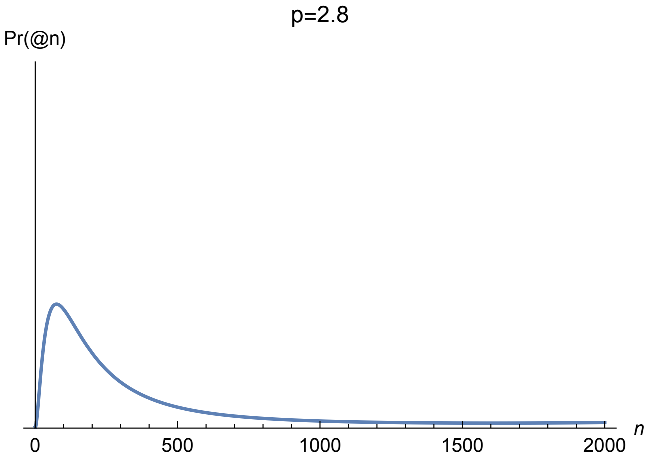


Figure: $Pr(@n) = 3^{n-pn^{1/3}}(1+n)^4(0.335)^n$

Sub-dominant cogrowth behaviour

The expected distribution for an ERR random walk with $\alpha = 3$ and $\beta = 0.335$:

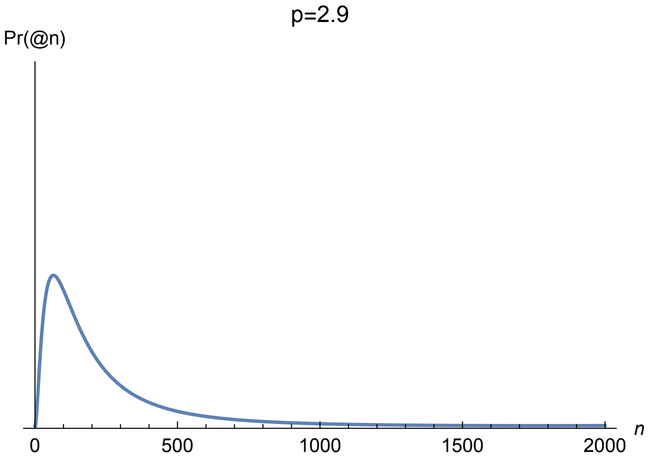


Figure: $Pr(@n) = 3^{n-pn^{1/3}}(1+n)^4(0.335)^n$

Sub-dominant cogrowth behaviour

The expected distribution for an ERR random walk with $\alpha = 3$ and $\beta = 0.335$:

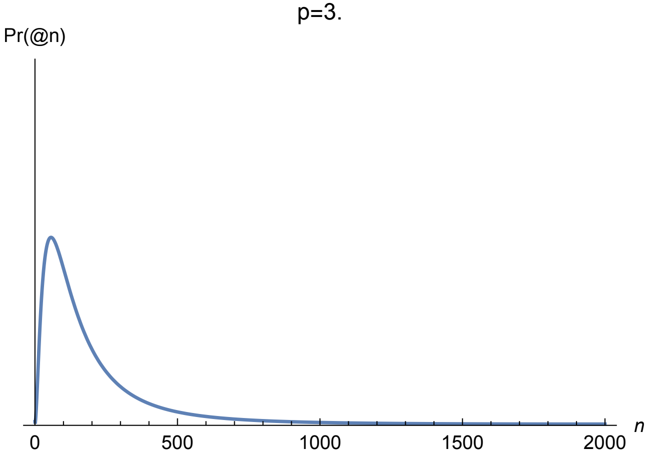


Figure: $Pr(@n) = 3^{n-pn^{1/3}}(1+n)^4(0.335)^n$

Sub-dominant cogrowth behaviour

The expected distribution for an ERR random walk with $\alpha = 3$ and $\beta = 0.335$:

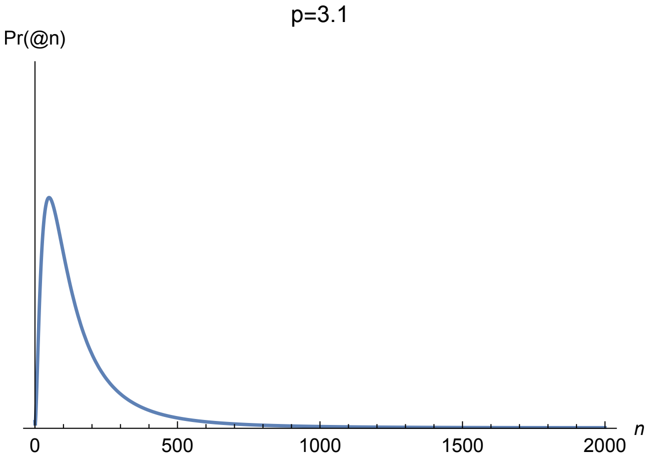


Figure: $Pr(@n) = 3^{n-pn^{1/3}}(1+n)^4(0.335)^n$

Sub-dominant cogrowth behaviour

The expected distribution for an ERR random walk with $\alpha = 3$ and $\beta = 0.335$:

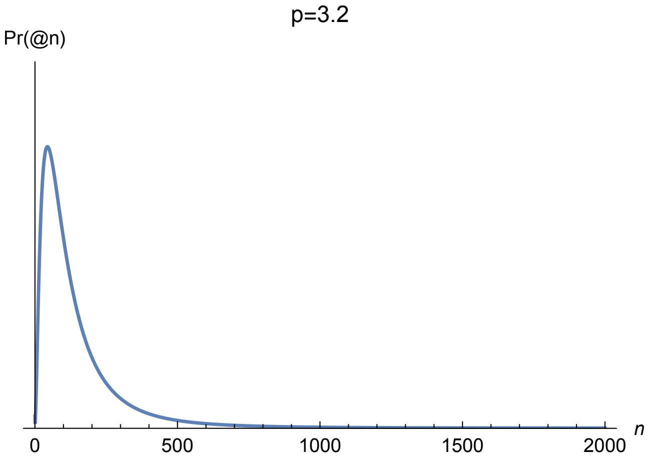


Figure: $Pr(@n) = 3^{n-pn^{1/3}}(1+n)^4(0.335)^n$

Predicting word lengths required for divergence

One solution to the problem is to start the ERR walk at some long word instead of the empty word.

Predicting word lengths required for divergence

One solution to the problem is to start the ERR walk at some long word instead of the empty word. How long is long?

Predicting word lengths required for divergence

One solution to the problem is to start the ERR walk at some long word instead of the empty word. How long is long? Suppose that G is an amenable group generated by two elements and their inverses (ie. $\beta_c = 1/3$). Suppose a random walk is run with $\beta = 0.34$. Then divergence will occur when

$$\frac{c_{2k+2}}{c_{2k}} > \frac{1}{0.34^2},$$

Predicting word lengths required for divergence

One solution to the problem is to start the ERR walk at some long word instead of the empty word. How long is long? Suppose that G is an amenable group generated by two elements and their inverses (ie. $\beta_c = 1/3$). Suppose a random walk is run with $\beta = 0.34$. Then divergence will occur when

$$\frac{c_{2k+2}}{c_{2k}} > \frac{1}{0.34^2},$$

which can be calculated from $\mathcal{R}(n)$.

Predicting word lengths required for divergence

One solution to the problem is to start the ERR walk at some long word instead of the empty word. How long is long? Suppose that G is an amenable group generated by two elements and their inverses (ie. $\beta_c = 1/3$). Suppose a random walk is run with $\beta = 0.34$. Then divergence will occur when

$$\frac{c_{2k+2}}{c_{2k}} > \frac{1}{0.34^2},$$

which can be calculated from $\mathcal{R}(n)$.

Note, however, that if the sub-dominant cogrowth properties are known to this level of precision, we could realistically expect the dominant cogrowth properties (and hence the amenability of the group) to be known also.

Sub-dominant cogrowth behaviour

The ERR paper attempts to infer information about the asymptotics of the cogrowth function, but what the walk actually measures is *local behaviour near word lengths actually visited*.

Sub-dominant cogrowth behaviour

The ERR paper attempts to infer information about the asymptotics of the cogrowth function, but what the walk actually measures is *local behaviour near word lengths actually visited*.

If the random walk gets stuck in a 'hump' at the start of predicted distribution it is still reflecting the cogrowth function *as it actually is*.

Answers to Questions

1. Why, for the groups $BS(1, N)$, does the accuracy of the ERR method decrease as N increases?

Answers to Questions

1. Why, for the groups $BS(1, N)$, does the accuracy of the ERR method decrease as N increases?

   $\mathcal{R}$ grows faster for larger N.

Answers to Questions

1. Why, for the groups $BS(1, N)$, does the accuracy of the ERR method decrease as N increases?

   $\mathcal{R}$ grows faster for larger N.

2. What group properties does this reflect, and what is the mechanism by which they slow convergence?

Answers to Questions

1. Why, for the groups $BS(1, N)$, does the accuracy of the ERR method decrease as N increases?

   $\mathcal{R}$ grows faster for larger N.

2. What group properties does this reflect, and what is the mechanism by which they slow convergence?

   The ERR random walk can only "see" the cogrowth function for word lengths actually visited by the walk.

Answers to Questions

1. Why, for the groups $BS(1, N)$, does the accuracy of the ERR method decrease as N increases?

   $\mathcal{R}$ grows faster for larger N.

2. What group properties does this reflect, and what is the mechanism by which they slow convergence?

   The ERR random walk can only "see" the cogrowth function for word lengths actually visited by the walk.

3. Are the walks "seeing" a representative sample of the space of trivial words?

Answers to Questions

1. Why, for the groups $BS(1, N)$, does the accuracy of the ERR method decrease as N increases?

   $\mathcal{R}$ grows faster for larger N.

2. What group properties does this reflect, and what is the mechanism by which they slow convergence?

   The ERR random walk can only "see" the cogrowth function for word lengths actually visited by the walk.

3. Are the walks "seeing" a representative sample of the space of trivial words?

   Yes.

Answers to Questions

1. Why, for the groups $BS(1, N)$, does the accuracy of the ERR method decrease as N increases?

   $\mathcal{R}$ grows faster for larger N.

2. What group properties does this reflect, and what is the mechanism by which they slow convergence?

   The ERR random walk can only "see" the cogrowth function for word lengths actually visited by the walk.

3. Are the walks "seeing" a representative sample of the space of trivial words?

   Yes.

4. Why do these (faulty) ERR walks consistently produce accurate estimates for initial cogrowth values?

Answers to Questions

1. Why, for the groups $BS(1, N)$, does the accuracy of the ERR method decrease as N increases?

   $\mathcal{R}$ grows faster for larger N.

2. What group properties does this reflect, and what is the mechanism by which they slow convergence?

   The ERR random walk can only "see" the cogrowth function for word lengths actually visited by the walk.

3. Are the walks "seeing" a representative sample of the space of trivial words?

   Yes.

4. Why do these (faulty) ERR walks consistently produce accurate estimates for initial cogrowth values?

   They aren't faulty - they exhibit exactly the behaviour predicted by the stationary distribution.

Application to Thompson's group F

- F has (if it is amenable) a very quickly growing Følner function, so ...

Application to Thompson's group F

- F has (if it is amenable) a very quickly growing Følner function, so ...
- existing evidence suggests $\mathcal{R}$ grows quickly for F, so ...

Application to Thompson's group F

- F has (if it is amenable) a very quickly growing Følner function, so ...
- existing evidence suggests $\mathcal{R}$ grows quickly for F, so ...
- $\beta > \beta_c$ will not necessarily imply divergence, so...

Application to Thompson's group F

- F has (if it is amenable) a very quickly growing Følner function, so ...
- existing evidence suggests $\mathcal{R}$ grows quickly for F, so ...
- $\beta > \beta_c$ will not necessarily imply divergence, so...
- The ERR method should not be used to estimate the amenability of F.

Application to Thompson's group F

- F has (if it is amenable) a very quickly growing Følner function, so ...
- existing evidence suggests $\mathcal{R}$ grows quickly for F, so ...
- $\beta > \beta_c$ will not necessarily imply divergence, so...
- The ERR method should not be used to estimate the amenability of F.
- BUT, this is not an obstruction to using ERR random walks to estimate initial values.

Application to Thompson's group F

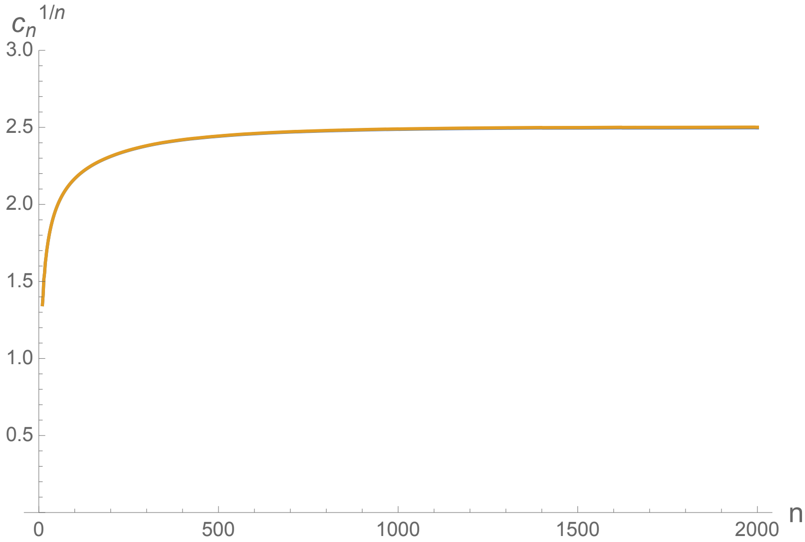
Application to Thompson's group F

n	exact	estimate
10	20	19.9996
12	64	63.9981
14	336	335.999
16	1160	1159.96
18	5896	5895.98
20	24652	24653.1
22	117628	117625
24	531136	531098
26	2559552	2558950
28	12142320	12138200

n	exact	estimate
30	59416808	59408300
32	290915560	290861000
34	1449601452	1449260000
36	7269071976	7268550000
38	36877764000	36876700000
40	1.8848×10^{11}	1.88491×10^{11}
42	9.7200×10^{11}	9.7205×10^{11}
44	5.0490×10^{12}	5.05097×10^{12}
46	2.6423×10^{13}	2.64353×10^{13}
48	1.3920×10^{14}	1.39246×10^{14}

Table: Estimate of the first 48 terms of the cogrowth function for Thompson's group F, constructed from 60 ERR random walks. Exact values from Haagerup.

Thank you



$c_n^{1/n}$
3.0
2.5
2.0
1.5
1.0
0.5
0
500
1000
1500
2000
n