Leap generators for composition schemes

Éric Fusy

CNRS, LIGM, Université Gustave Eiffel, Marne-la-vallée

Joint work with Carine Pivoteau
& work in progress with Konstantinos Panagiotou

AofA'26, Munich

Uniform random generation in a class

Combinatorial class $=$ set $\mathcal{C}$ of objects

$\&$ size-function $\gamma \in \mathcal{C} \;\to\; |\gamma| \in \mathbb{N}$

such that $\mathcal{C}_n = \{\gamma \in \mathcal{C} \;:\; |\gamma| = n\}$ is finite for $n \geq 0$

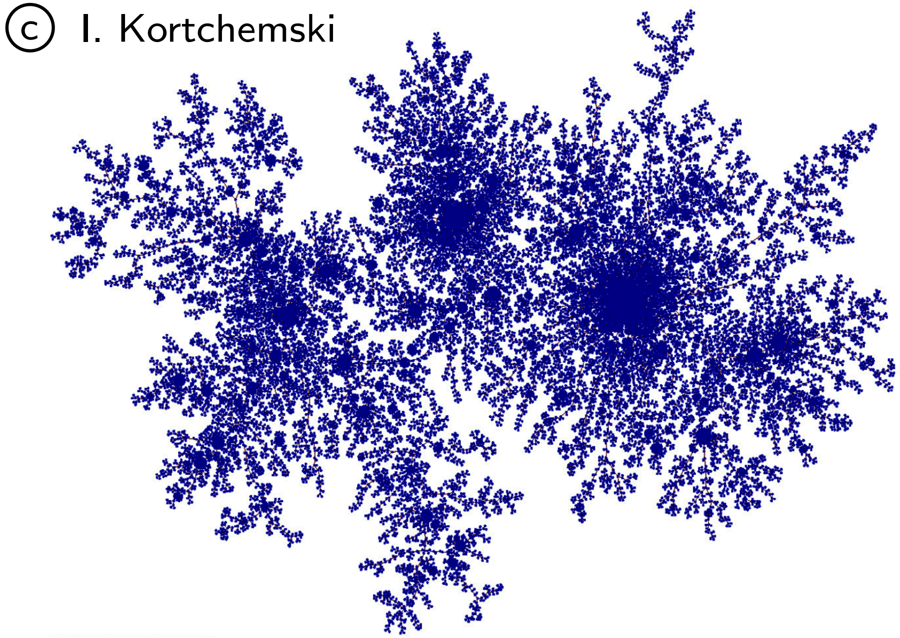
(Exact-size uniform) random generator for $\mathcal{C}$:

algorithm $\Gamma C[n]$ that given input size $n \geq 0$,

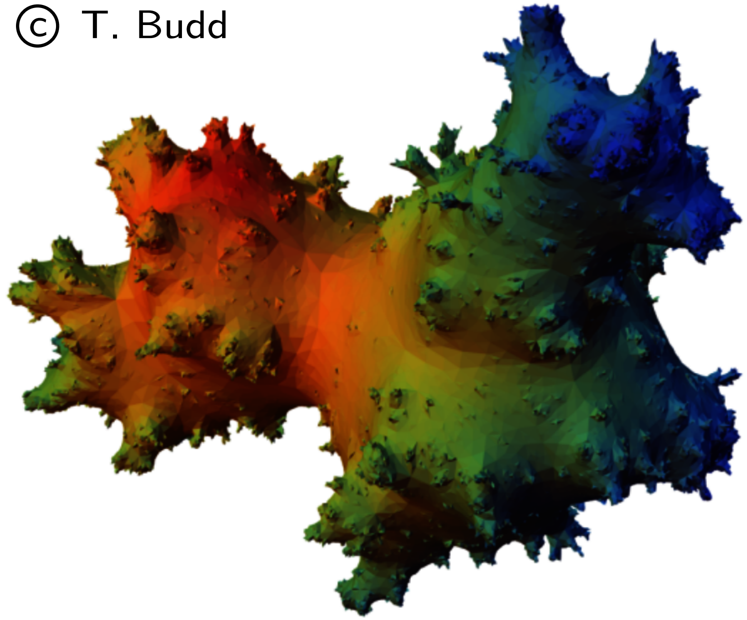
outputs an object $\gamma \in \mathcal{C}_n$ under the uniform distribution

Rk: strives for polynomial complexity in n, ideally $O(n)$

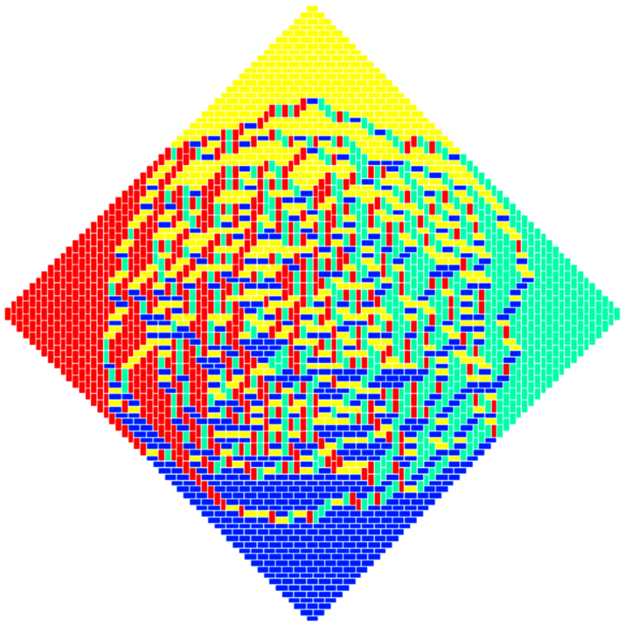
Simulations of large random objects



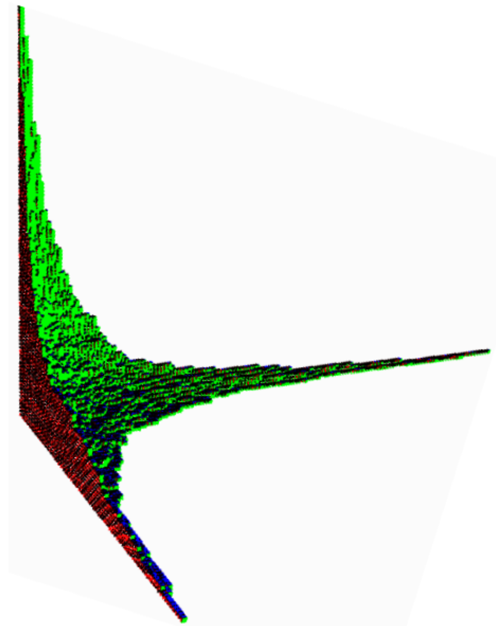
random (stable) tree



random (Schnyder-wood-decorated) planar map



random domino tiling of aztec diamond



random plane partition

Main methods

- Markov chains

  random update operations giving symmetric Markov chain on $\mathcal{C}_n$

  (challenge to analyse the mixing time, superlinear complexity)

- Combinatorial methods

  bijections/correspondences

  possibly combined with rejection techniques

- Automatic methods in symbolic combinatorics

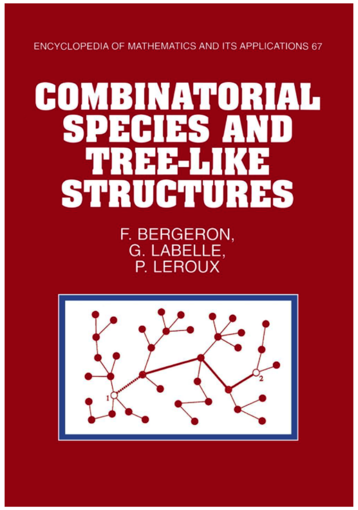
  - recursive method (based on the counting coefficients)

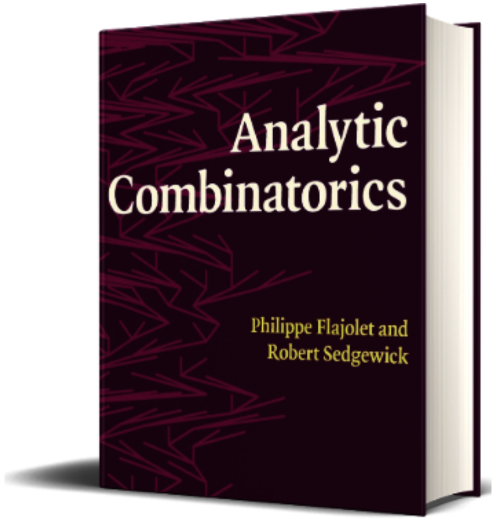
    [Nijenhuis-Wilf'78, Flajolet-Zimmermann, Van Cutsem'94]

  - Boltzmann samplers (based on evaluations of generating functions)

    [Duchon-Flajolet-Louchard-Schaeffer'04]

Symbolic combinatorics





For a combinatorial class $\mathcal{C}_n$:

Counting coefficients	$c_n = \mathrm{Card}(\mathcal{C}_n)$

Generating function $C(z) = \begin{cases} \sum_{\gamma \in \mathcal{C}} z^{|\gamma|} = \sum_n c_n z^n & \text{unlabeled} \\ \sum_{\gamma \in \mathcal{C}} \frac{1}{|\gamma|!} z^{|\gamma|} = \sum_n \frac{1}{n!} c_n z^n & \text{labeled} \end{cases}$

Radius of convergence ρ is given by $\rho^{-1} = \limsup \left([z^n]C(z)\right)^{1/n}$

Transfer theorem: $C(z) \sim \kappa(1 - z/\rho)^{-s} \;\Rightarrow [z^n]C(z) \sim \frac{\kappa}{\Gamma(s)} \rho^{-n} n^{s-1}$

Symbolic combinatorics

Basic classes

Neutral class	$\mathbf{1}$
Atom	$\mathcal{Z}$

Some of the basic constructions

Sum	$\mathcal{C} = \mathcal{A} + \mathcal{B}$	$C(z) = A(z) + B(z)$
Product	$\mathcal{C} = \mathcal{A} \times \mathcal{B}$	$C(z) = A(z) \times B(z)$
Set (labeled)	$\mathcal{C} = \mathrm{Set}(\mathcal{A})$	$C(z) = \exp(A(z))$
Multiset (unlabeled)	$\mathcal{C} = \mathrm{MSet}(\mathcal{A})$	$C(z) = \exp\left(\sum_{i\geq 1} \frac{1}{i} A(z^i)\right)$

Examples:

binary trees	Cayley trees	Pólya trees	Permutations
$\mathcal{C} = \mathcal{Z} + \mathcal{C}^2$	$\mathcal{C} = \mathcal{Z} \times \mathrm{Set}(\mathcal{C})$	$\mathcal{C} = \mathcal{Z} \times \mathrm{MSet}(\mathcal{C})$	$\mathcal{C} = \mathrm{Set}(\mathrm{Cyc}(\mathcal{Z}))$

Boltzmann samplers [Duchon, Flajolet, Louchard, Schaeffer'04]

For $\mathcal{C}$ a combinatorial class, and $x > 0$ such that $C(x)$ converges

Boltzmann distribution on $\mathcal{C}$ assigns probability

$$\mathbb{P}_x^{\mathcal{C}}(\gamma) = \begin{cases} \dfrac{1}{C(x)} x^{|\gamma|} & \text{unlabeled case} \\ \dfrac{1}{C(x)} \dfrac{x^{|\gamma|}}{|\gamma|!} & \text{labeled case} \end{cases}$$

Rk: induced distribution on $\mathcal{C}_n$ is uniform for each $n \geq 0$

Def: Boltzmann sampler $\Gamma C(x)$ = algorithm that draws an object in $\mathcal{C}$ at random under distribution $\mathbb{P}_x^{\mathcal{C}}$

Sampling rules

[Duchon, Flajolet, Louchard, Schaeffer'04]

For $\mathcal{C} = \mathrm{op}(\mathcal{A}_1, \ldots, \mathcal{A}_k)$

design a rule to assemble $\Gamma C(x)$ from $\Gamma A_1(x), \ldots, \Gamma A_k(x)$

Sampling rules

[Duchon, Flajolet, Louchard, Schaeffer'04]

For $\mathcal{C} = \mathrm{op}(\mathcal{A}_1, \ldots, \mathcal{A}_k)$

design a rule to assemble $\Gamma C(x)$ from $\Gamma A_1(x), \ldots, \Gamma A_k(x)$

- Sum $\mathcal{C} = \mathcal{A} + \mathcal{B}$

Rk: For $\gamma \in \mathcal{C}$ drawn under $\mathbb{P}_x^{\mathcal{C}}$, one has $\mathbb{P}(\gamma \in \mathcal{A}) = \frac{A(x)}{C(x)}$

Sampling rules

[Duchon, Flajolet, Louchard, Schaeffer'04]

For $\mathcal{C} = \mathrm{op}(\mathcal{A}_1, \ldots, \mathcal{A}_k)$

design a rule to assemble $\Gamma C(x)$ from $\Gamma A_1(x), \ldots, \Gamma A_k(x)$

- Sum $\mathcal{C} = \mathcal{A} + \mathcal{B}$

Rk: For $\gamma \in \mathcal{C}$ drawn under $\mathbb{P}_x^{\mathcal{C}}$, one has $\mathbb{P}(\gamma \in \mathcal{A}) = \frac{A(x)}{C(x)}$

$\Gamma C(x)$: if $\mathrm{Bern}(A(x)/C(x))$ then return $\Gamma A(x)$
else return $\Gamma B(x)$

Sampling rules

[Duchon, Flajolet, Louchard, Schaeffer'04]

For $\mathcal{C} = \mathrm{op}(\mathcal{A}_1, \ldots, \mathcal{A}_k)$

design a rule to assemble $\Gamma C(x)$ from $\Gamma A_1(x), \ldots, \Gamma A_k(x)$

- Sum $\mathcal{C} = \mathcal{A} + \mathcal{B}$

  Rk: For $\gamma \in \mathcal{C}$ drawn under $\mathbb{P}_x^{\mathcal{C}}$, one has $\mathbb{P}(\gamma \in \mathcal{A}) = \frac{A(x)}{C(x)}$

  $\Gamma C(x)$: if $\mathrm{Bern}(A(x)/C(x))$ then return $\Gamma A(x)$
  else return $\Gamma B(x)$

- Product $\mathcal{C} = \mathcal{A} \times \mathcal{B}$

  $\Gamma C(x)$: return $< \Gamma A(x), \Gamma B(x) >$

Sampling rules

[Duchon, Flajolet, Louchard, Schaeffer'04]

For $\mathcal{C} = \mathrm{op}(\mathcal{A}_1, \ldots, \mathcal{A}_k)$

design a rule to assemble $\Gamma C(x)$ from $\Gamma A_1(x), \ldots, \Gamma A_k(x)$

- Sum $\mathcal{C} = \mathcal{A} + \mathcal{B}$

  Rk: For $\gamma \in \mathcal{C}$ drawn under $\mathbb{P}_x^{\mathcal{C}}$, one has $\mathbb{P}(\gamma \in \mathcal{A}) = \frac{A(x)}{C(x)}$

  $\Gamma C(x)$: if $\mathrm{Bern}(A(x)/C(x))$ then return $\Gamma A(x)$
  else return $\Gamma B(x)$

- Product $\mathcal{C} = \mathcal{A} \times \mathcal{B}$

  $\Gamma C(x)$: return $< \Gamma A(x), \Gamma B(x) >$

- Sequence $\mathcal{C} = \mathrm{Seq}(\mathcal{A})$

  $\Gamma C(x) : k \leftarrow \mathrm{Geom}(A(x))$

  $\gamma_1 \leftarrow \Gamma A(x), \ldots, \gamma_k \leftarrow \Gamma A(x)$

  return $< \gamma_1, \ldots, \gamma_k >$

Sampling rules

[Duchon, Flajolet, Louchard, Schaeffer'04]

For $\mathcal{C} = \mathrm{op}(\mathcal{A}_1, \ldots, \mathcal{A}_k)$

design a rule to assemble $\Gamma C(x)$ from $\Gamma A_1(x), \ldots, \Gamma A_k(x)$

- Sum $\mathcal{C} = \mathcal{A} + \mathcal{B}$

  Rk: For $\gamma \in \mathcal{C}$ drawn under $\mathbb{P}_x^{\mathcal{C}}$, one has $\mathbb{P}(\gamma \in \mathcal{A}) = \frac{A(x)}{C(x)}$

  $\Gamma C(x)$: if $\mathrm{Bern}(A(x)/C(x))$ then return $\Gamma A(x)$
  else return $\Gamma B(x)$

- Product $\mathcal{C} = \mathcal{A} \times \mathcal{B}$

  $\Gamma C(x)$: return $< \Gamma A(x), \Gamma B(x) >$

- Sequence $\mathcal{C} = \mathrm{Seq}(\mathcal{A})$

  $\Gamma C(x) : k \leftarrow \mathrm{Geom}(A(x))$
  $\gamma_1 \leftarrow \Gamma A(x), \ldots, \gamma_k \leftarrow \Gamma A(x)$
  return $< \gamma_1, \ldots, \gamma_k >$

- Set $\mathcal{C} = \mathrm{Set}(\mathcal{A})$

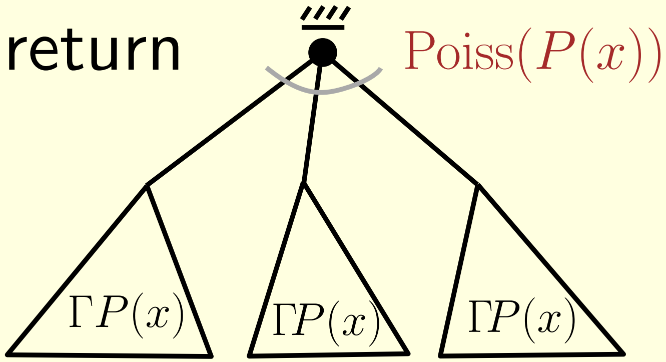
  $\Gamma C(x) : k \leftarrow \mathrm{Pois}(A(x))$
  $\gamma_1 \leftarrow \Gamma A(x), \ldots, \gamma_k \leftarrow \Gamma A(x)$
  return $< \gamma_1, \ldots, \gamma_k >$

Examples for tree-families

Plane trees

$$\mathcal{P} = \mathcal{Z} \times \mathrm{Seq}(\mathcal{P})$$

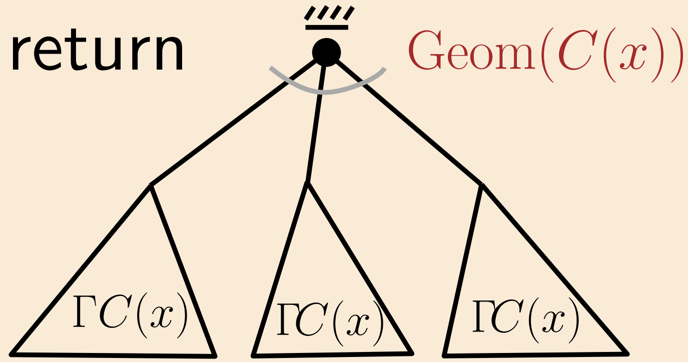
$\Gamma P(x)$: return



Cayley trees

$$\mathcal{C} = \mathcal{Z} \times \mathrm{Set}(\mathcal{C})$$

$\Gamma C(x)$: return

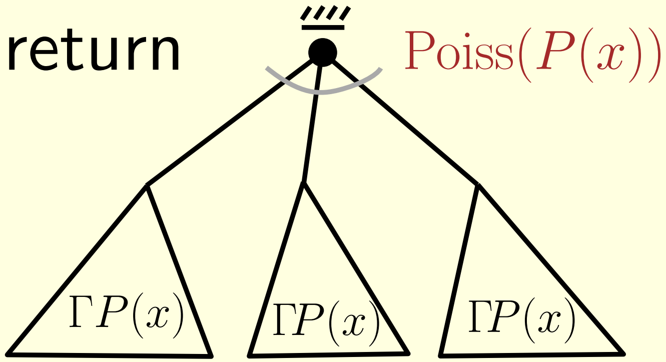


Examples for tree-families

Plane trees

$$\mathcal{P} = \mathcal{Z} \times \mathrm{Seq}(\mathcal{P})$$

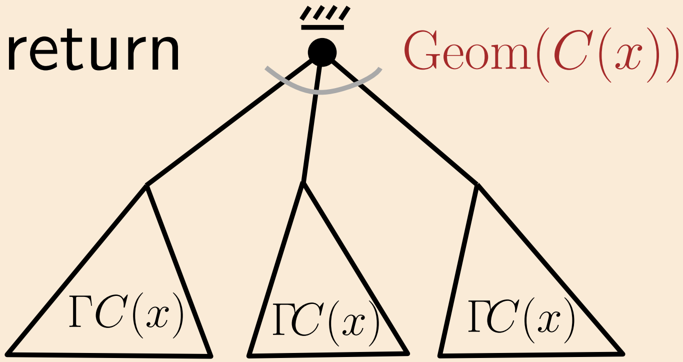
$\Gamma P(x)$: return



Cayley trees

$$\mathcal{C} = \mathcal{Z} \times \mathrm{Set}(\mathcal{C})$$

$\Gamma C(x)$: return

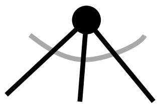


For a **simply generated** tree-family $\mathcal{C}$, with generating function equation

$$C(z) = z\,\Phi(C(z))$$

Boltzmann-generated tree $=$ Galton-Watson tree

for each drawn node



number of children is μ-distributed,

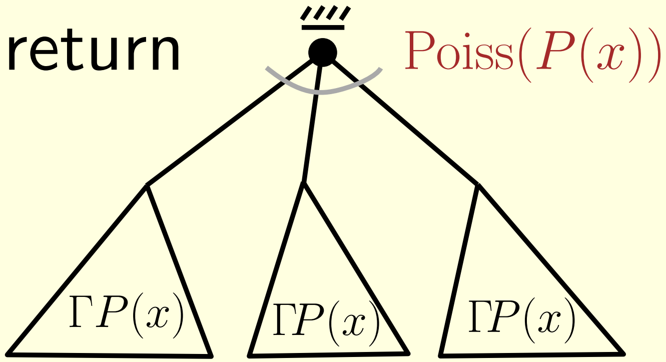
for μ given by $\sum_k \mu(k) u^k = \frac{\Phi(uC(x))}{\Phi(C(x))}$

Examples for tree-families

Plane trees

$$\mathcal{P} = \mathcal{Z} \times \mathrm{Seq}(\mathcal{P})$$

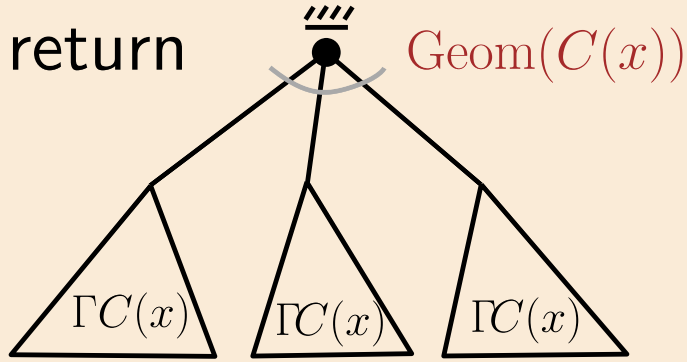
$\Gamma P(x)$: return



Cayley trees

$$\mathcal{C} = \mathcal{Z} \times \mathrm{Set}(\mathcal{C})$$

$\Gamma C(x)$: return

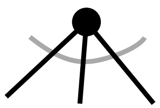


For a **simply generated** tree-family $\mathcal{C}$, with generating function equation

$$C(z) = z\,\Phi(C(z))$$

Boltzmann-generated tree $=$ Galton-Watson tree

for each drawn node



number of children is μ-distributed,

for μ given by $\sum_k \mu(k) u^k = \frac{\Phi(uC(x))}{\Phi(C(x))}$

$x < \rho \;\Rightarrow$ GW subcritical ($\mathbb{E}(\mu) < 1$)

$x = \rho \;\Rightarrow$ GW critical ($\mathbb{E}(\mu) = 1$)

Sampling rules for constructions with symmetries

[Flajolet, F, Pivoteau'06]

$\textsc{MSet}_2(\mathcal{A}) \cong$ unordered pairs of objects de $\mathcal{A}$

$$\begin{aligned} \mathcal{C} &= \textsc{MSet}_2(\mathcal{A}) \\ C(z) &= \frac{1}{2}A^2(z) + \frac{1}{2}A(z^2) \end{aligned} \qquad [\ 2\textsc{MSet}_2(\mathcal{A}) = \mathcal{A}^2 + \Delta\mathcal{A}^2\ ]$$

Algorithm: $\Gamma C(x)$

if $Bern\left(\frac{1}{2}\frac{A^2(x)}{C(x)}\right) = 1$ **then**
  Return $(\ \Gamma A(x), \Gamma A(x)\ )$
else
  $a \leftarrow \Gamma A(x^2)$;
  Return (a, a);
end if

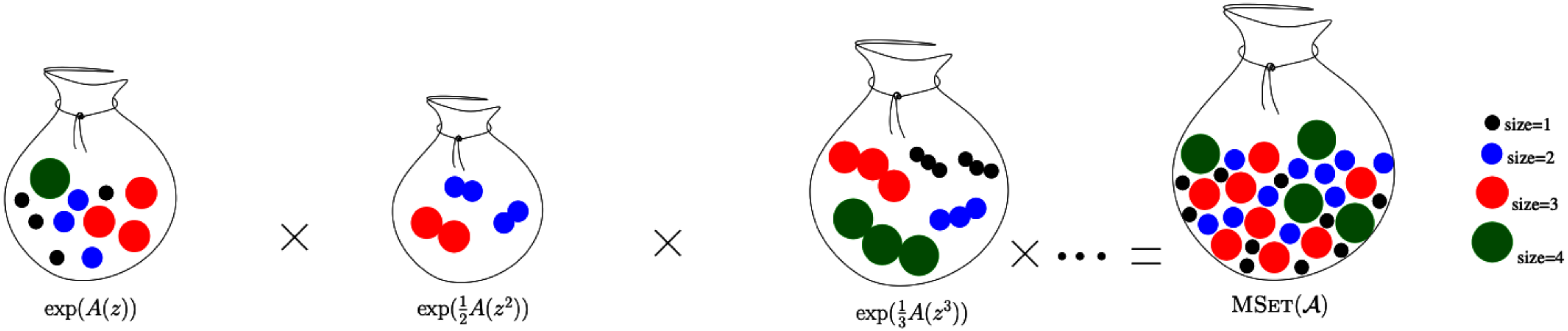
$\Rightarrow$ Boltzmann sampler for non-embedded binary trees (**not** GW-trees):

$$\mathcal{C} = \mathcal{Z} \times \textsc{MSet}_2(\mathcal{C}) \qquad C(z) = z\left(1 + \tfrac{1}{2}\left(C(z)^2 + C(z^2)\right)\right)$$

Sampling rules for constructions with symmetries

Let $\mathcal{C} := \textsc{MSet}(\mathcal{A})$ be the class of all multisets of objects of $\mathcal{A}$.

$$C(z) = \exp\left(\sum_{k=1}^{\infty} \frac{1}{k} A(z^k)\right) = \prod_{k=1}^{\infty} \exp\left(\frac{1}{k} A(z^k)\right)$$



repeat $\text{Poiss}(A(x))$ times $\gamma \leftarrow \Gamma A(x)$ add γ	repeat $\text{Poiss}(\frac{A(x^2)}{2})$ times $\gamma \leftarrow \Gamma A(x^2)$ add 2 copies of γ	repeat $\text{Poiss}(\frac{A(x^3)}{3})$ times $\gamma \leftarrow \Gamma A(x^3)$ add 3 copies of γ

Sampling complexity of Boltzmann samplers

- Preprocessing time is small (advantage over the recursive method):

evaluation at x of finitely many generating functions
(those involved in the decomposition grammar of $\mathcal{C}$)

[Pivoteau, Salvy, Soria'12] [Pivoteau, Salvy'25]

- Sampling complexity: linear in the size of the output object

[Duchon, Flajolet, Louchard, Schaeffer'04] [Flajolet, F, Pivoteau'06]

Sampling complexity of Boltzmann samplers

- Preprocessing time is small (advantage over the recursive method):

evaluation at x of finitely many generating functions
(those involved in the decomposition grammar of $\mathcal{C}$)

[Pivoteau, Salvy, Soria'12] [Pivoteau, Salvy'25]

- Sampling complexity: linear in the size of the output object

[Duchon, Flajolet, Louchard, Schaeffer'04] [Flajolet, F, Pivoteau'06]

In practice, aim for a (large) size n, exactly or approximately: $\left|\frac{\mathrm{size}}{n} - 1\right| \leq \epsilon$

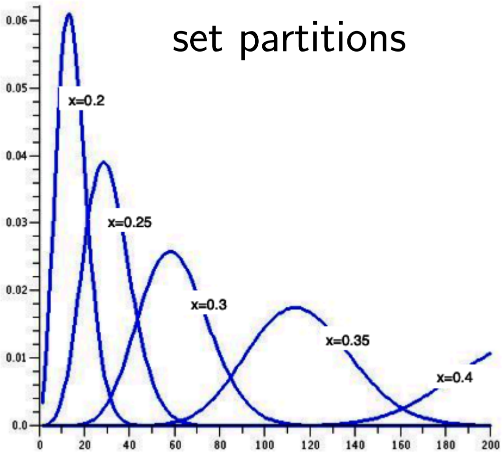
repeat calling $\Gamma C(x)$ until size in desired target-set

(rk: early-abort when size $>$ target-set)

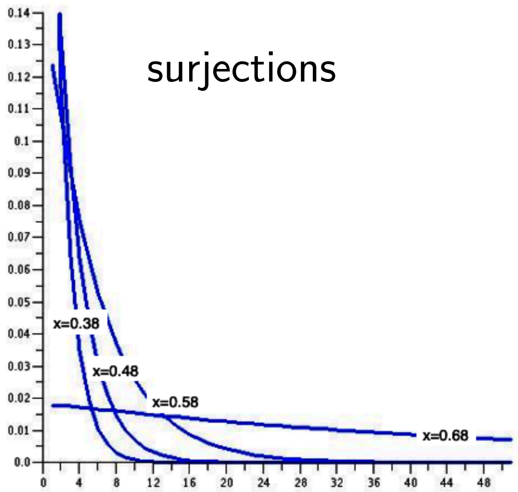
Size-distribution: $\mathbb{P}_x(N = n) = \frac{c_n x^n}{C(x)}$ $\mathbb{E}_x(N) = \frac{xC'(x)}{C(x)}$

natural choice: take x_n such that $\mathbb{E}_x(N) = n$

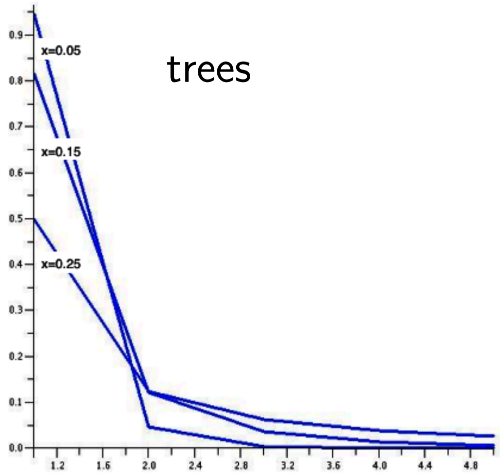
Complexity of approximate/exact size sampling



$$C(x) = e^{(e^x - 1)}$$



$$C(x) = \frac{1}{2 - e^x}$$



$$C(x) = \frac{1 - \sqrt{1 - 4x}}{2}$$

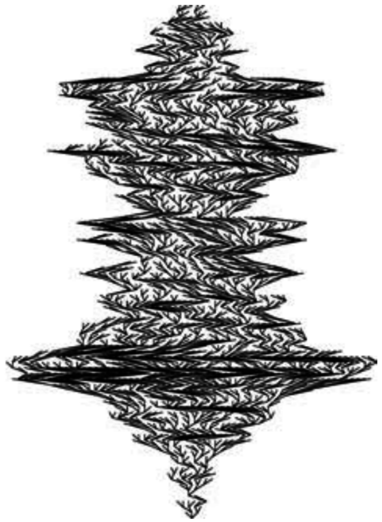
singular sampler + early-abort

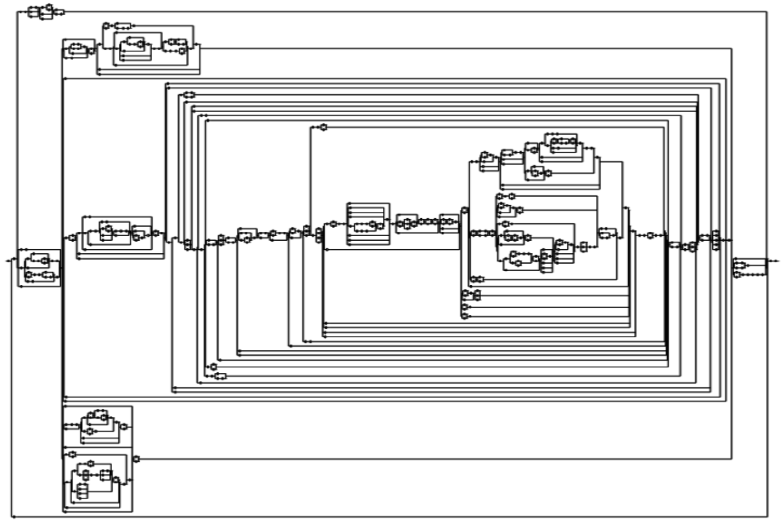
typical complexity $O(n/\epsilon)$ for approximate-size sampling

$O(n^2)$ for exact-size sampling

Drawback of approximate size sampling

efficient to obtain a simulation of a very large random object

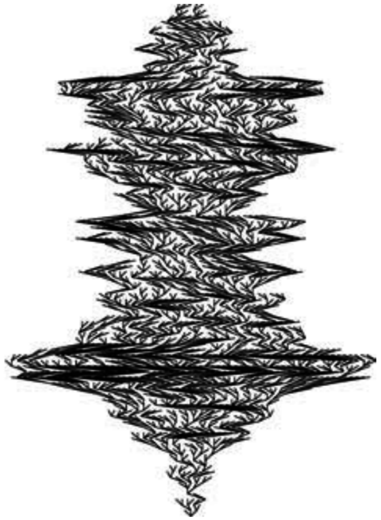


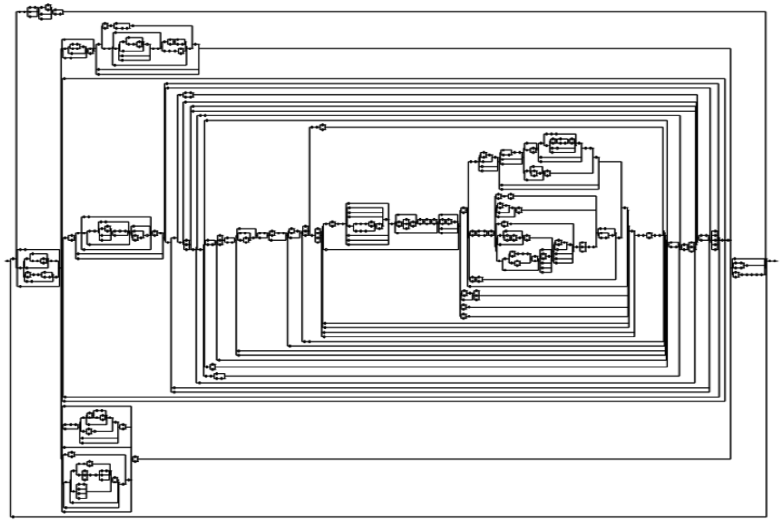


but not to obtain a histogram for a parameter at a given size

Drawback of approximate size sampling

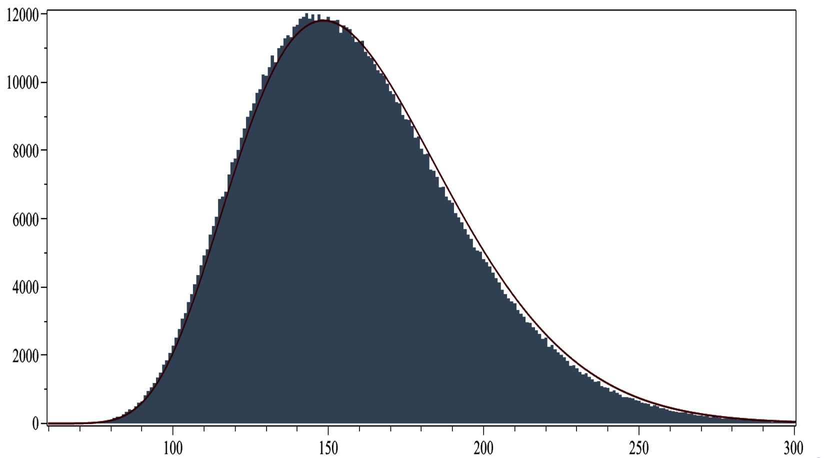
efficient to obtain a simulation of a very large random object





but not to obtain a histogram for a parameter at a given size

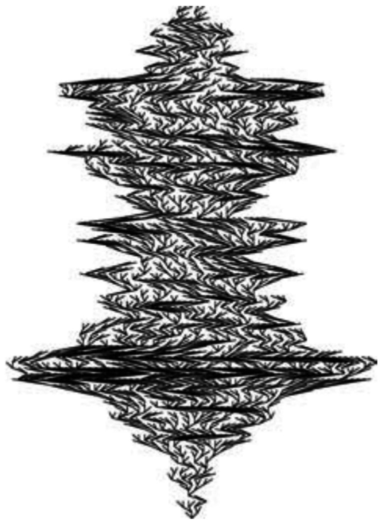
height in random Pólya trees

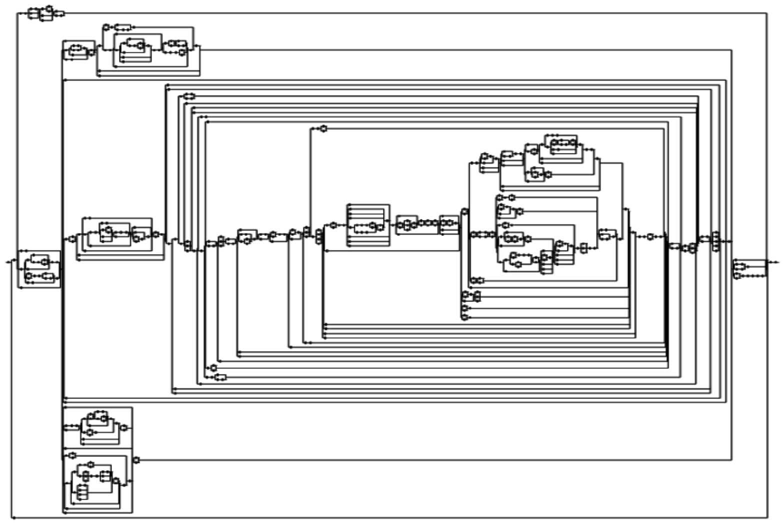


$n = 5000$, sample-size $N = 10^6$

Drawback of approximate size sampling

efficient to obtain a simulation of a very large random object

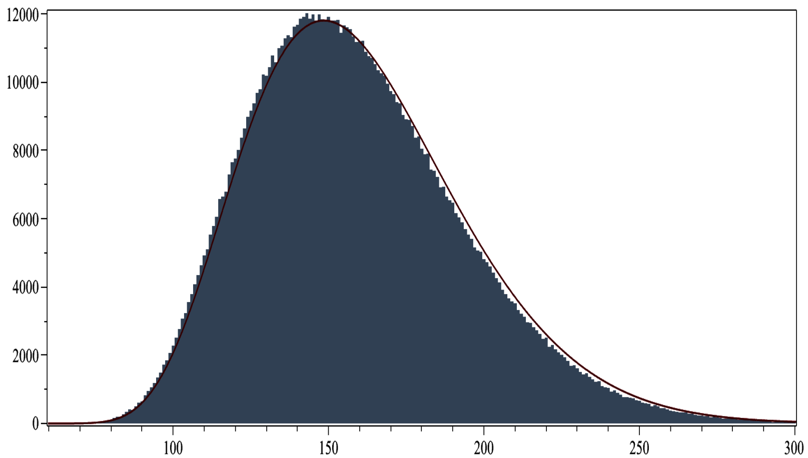




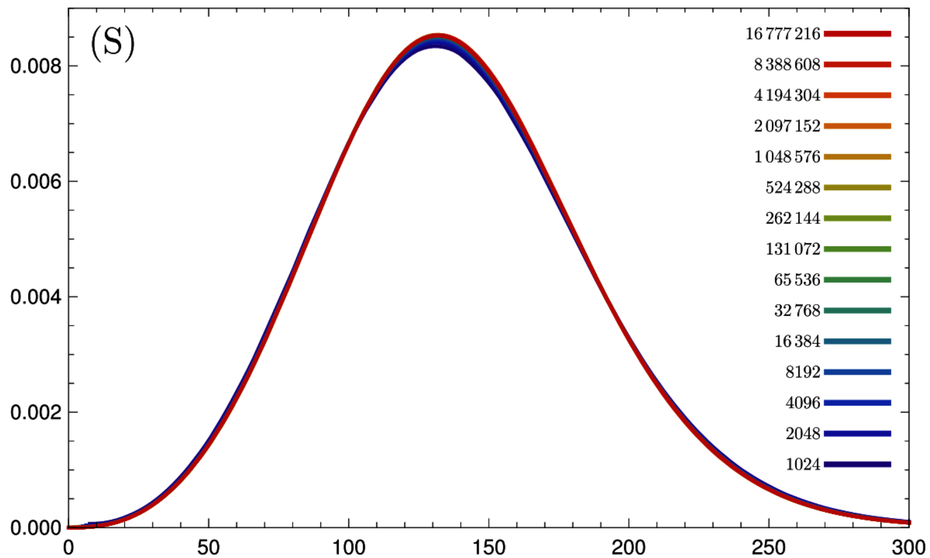
but not to obtain a histogram for a parameter at a given size

[Barkley-Budd'19] curve-fitting of histograms to estimate rescaling exponent n^α of distances in models of random decorated planar maps

height in random Pólya trees



$n = 5000$, sample-size $N = 10^6$



Cases with exact-size sampling in linear time

Two cases where exact-size sampling has linear time implementation

(i) Supercritical sequences $\mathcal{C} = \mathrm{Seq}(\mathcal{B})$

Leap generators [Duchon, Flajolet, Louchard, Schaeffer'04]

(ii) Galton-Watson trees $C(z) = z\,\Phi(C(z))$ [Devroye'12]

Cases with exact-size sampling in linear time

Two cases where exact-size sampling has linear time implementation

(i) Supercritical sequences $\mathcal{C} = \mathrm{Seq}(\mathcal{B})$

Leap generators [Duchon, Flajolet, Louchard, Schaeffer'04]

(ii) Galton-Watson trees $C(z) = z\,\Phi(C(z))$ [Devroye'12]

Motivations for this talk:

extend setting where (i) (resp. (ii)) applies

joint work with C. Pivoteau

joint work (in progress) with K. Panagiotou

Classification for sequence classes

$$\mathcal{C} = \mathrm{Seq}(\mathcal{B}) \qquad C(z) = \frac{1}{1 - B(z)} \qquad \rho = \text{singularity of } C(z)$$

Classification for sequence classes

$$\mathcal{C} = \mathrm{Seq}(\mathcal{B}) \qquad C(z) = \frac{1}{1 - B(z)} \qquad \rho = \text{singularity of } C(z)$$

supercritical	critical	subcritical
$B(\rho_B) > 1$	$B(\rho_B) = 1$	$B(\rho_B) < 1$
$\rho < \rho_B$ solution of $B(\rho) = 1$ pole at ρ	$\rho = \rho_B$ source of ρ: pole and B	$\rho = \rho_B$ source of ρ: B

Classification for sequence classes

$\mathcal{C} = \mathrm{Seq}(\mathcal{B})$ $\qquad C(z) = \dfrac{1}{1 - B(z)}$ $\qquad \rho = \text{singularity of } C(z)$

supercritical	critical	subcritical
$B(\rho_B) > 1$	$B(\rho_B) = 1$	$B(\rho_B) < 1$
$\rho < \rho_B$ solution of $B(\rho) = 1$ pole at ρ	$\rho = \rho_B$ source of ρ: pole and B	$\rho = \rho_B$ source of ρ: B

Rk: for random object in $\mathcal{C}_n$

- supercritical: $\approx n/\mu$ components, where $\mu = \mathbb{E}(\text{size } \Gamma B(\rho))$
  components '$\sim$' independent draws from $\Gamma B(\rho)$
- subcritical: $\Theta(1)$ components, largest is huge (size $n - \Theta(1)$)

Classification for sequence classes

$$\mathcal{C} = \mathrm{Seq}(\mathcal{B}) \qquad C(z) = \frac{1}{1 - B(z)} \qquad \rho = \text{singularity of } C(z)$$

supercritical	critical	subcritical
$B(\rho_B) > 1$	$B(\rho_B) = 1$	$B(\rho_B) < 1$
$\rho < \rho_B$ solution of $B(\rho) = 1$ pole at ρ	$\rho = \rho_B$ source of ρ: pole and B	$\rho = \rho_B$ source of ρ: B

Rk: for random object in $\mathcal{C}_n$

- supercritical: $\approx n/\mu$ components, where $\mu = \mathbb{E}(\text{size } \Gamma B(\rho))$
  components '$\sim$' independent draws from $\Gamma B(\rho)$
- subcritical: $\Theta(1)$ components, largest is huge (size $n - \Theta(1)$)

Examples:

supercritical: $\mathcal{C} = \mathrm{Seq}(\mathcal{Z} + \cdots + \mathcal{Z}^m)$ compositions with bounded parts

subcritical: Dyck paths

Boltzmann sampler for supercritical sequences

[Duchon et al.'04]

$\mathcal{C} = \mathrm{Seq}(\mathcal{B})$ $\qquad C(z) = \dfrac{1}{1-B(z)}$ $\qquad B(\rho) = 1$ for $\rho < \rho_B$

$\Gamma C(x) : k \leftarrow \mathrm{Geom}(B(x))$
return $< \Gamma B(x), \ldots, \Gamma B(x) >$ (k indep. calls)

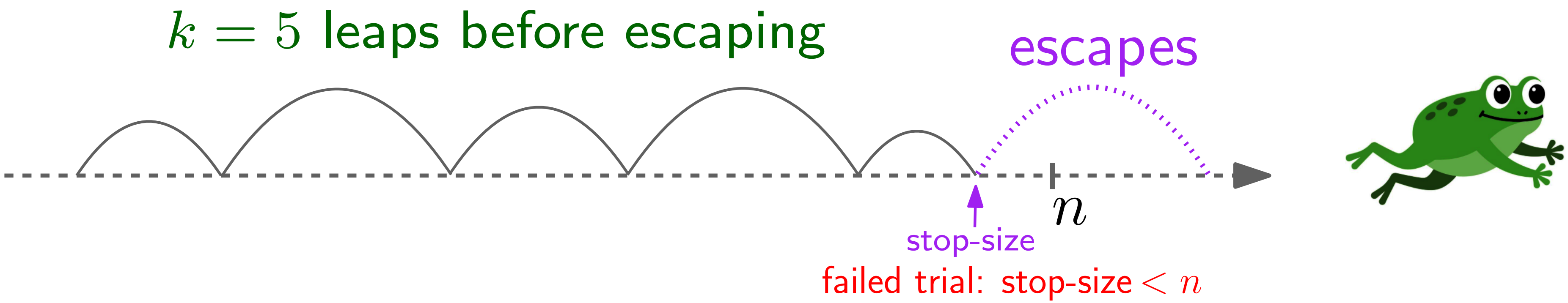
Structure in $\mathcal{C}_n$ with k components has probability

$$B(x)^k(1 - B(x))\ \frac{\rho^{n_1}}{B(x)} \cdots \frac{\rho^{n_k}}{B(x)} = \frac{\rho^n}{C(x)}$$

Leap generator for supercritical sequences

[Duchon et al'04]

Leap generator (size n) emulates singular Boltzmann sampler ($x = \rho$)



$\Gamma\widehat{C}[n]$: **Repeat**

draw $\beta_1 \leftarrow \Gamma B(\rho)$, $\beta_2 \leftarrow \Gamma B(\rho), \ldots$

up to $|\beta_1| + \cdots + |\beta_{k+1}| > n$

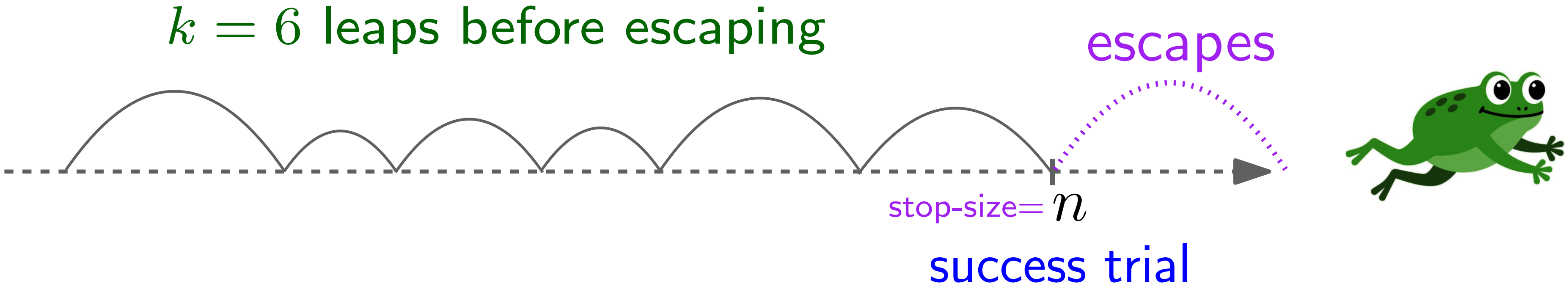
if $|\beta_1| + \cdots + |\beta_k| = n$ (success)

return $\mathrm{seq}(\beta_1, \ldots, \beta_k)$

Leap generator for supercritical sequences

[Duchon et al'04]

Leap generator (size n) emulates singular Boltzmann sampler ($x = \rho$)



$\Gamma\widehat{C}[n]$: **Repeat**

draw $\beta_1 \leftarrow \Gamma B(\rho),\ \beta_2 \leftarrow \Gamma B(\rho), \ldots$

up to $|\beta_1| + \cdots + |\beta_{k+1}| > n$

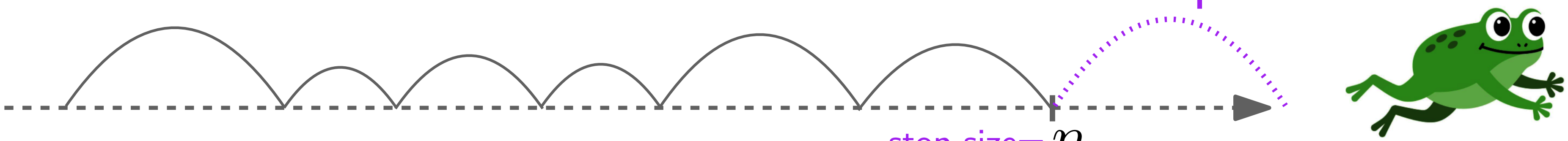
if $|\beta_1| + \cdots + |\beta_k| = n$ (success)

return $\mathrm{seq}(\beta_1, \ldots, \beta_k)$

Leap generator for supercritical sequences

[Duchon et al'04]

Leap generator (size n) emulates singular Boltzmann sampler ($x = \rho$)



$\Gamma\widehat{C}[n]$: **Repeat**

draw $\beta_1 \leftarrow \Gamma B(\rho)$, $\beta_2 \leftarrow \Gamma B(\rho), \ldots$

up to $|\beta_1| + \cdots + |\beta_{k+1}| > n$

if $|\beta_1| + \cdots + |\beta_k| = n$ (success)

return $\mathrm{seq}(\beta_1, \ldots, \beta_k)$

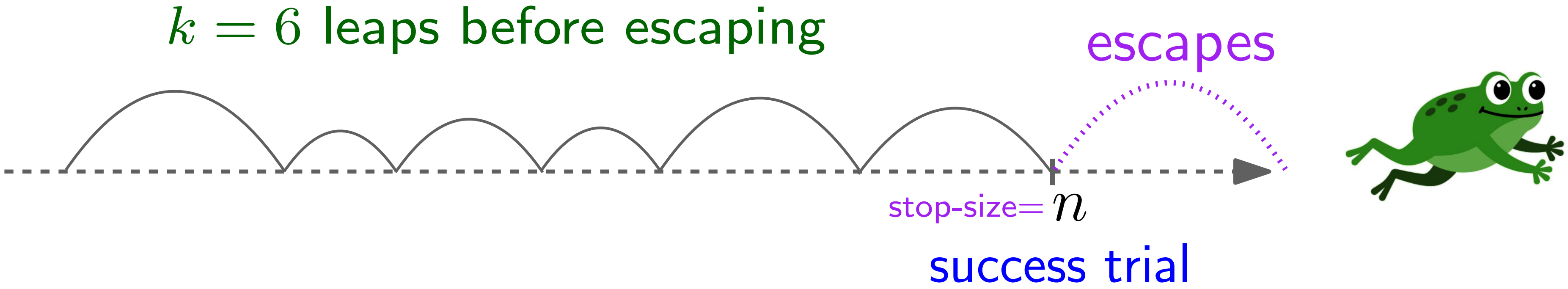
at each trial, structure in $\mathcal{C}_n$ with k components has probability

$$\frac{\rho^{n_1}}{B(\rho)} \cdots \frac{\rho^{n_k}}{B(\rho)} = \rho^n \qquad \text{(no dependency on } k \text{ !)}$$

Leap generator for supercritical sequences

[Duchon et al'04]

Leap generator (size n) emulates singular Boltzmann sampler ($x = \rho$)



$\Gamma\widehat{C}[n]$: **Repeat**

draw $\beta_1 \leftarrow \Gamma B(\rho)$, $\beta_2 \leftarrow \Gamma B(\rho), \ldots$

up to $|\beta_1| + \cdots + |\beta_{k+1}| > n$

if $|\beta_1| + \cdots + |\beta_k| = n$ (success)

return $\mathrm{seq}(\beta_1, \ldots, \beta_k)$

at each trial, structure in $\mathcal{C}_n$ with k components has probability

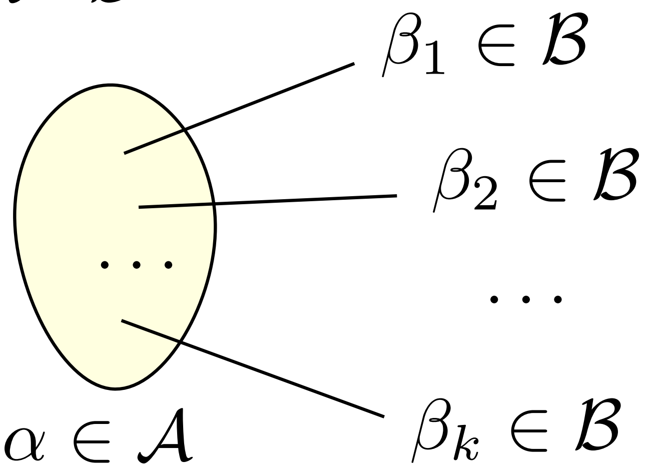
$$\frac{\rho^{n_1}}{B(\rho)} \cdots \frac{\rho^{n_k}}{B(\rho)} = \rho^n \qquad \text{(no dependency on } k \text{ !)}$$

proba success of trial is $\sim 1/\mu$ with $\mu = \mathbb{E}(\text{size of } \Gamma B(\rho))$

Composition schemes

Composition operator: $\mathcal{C} = \mathcal{A} \circ \mathcal{B}$

$\gamma \in \mathcal{C}$ is of the form

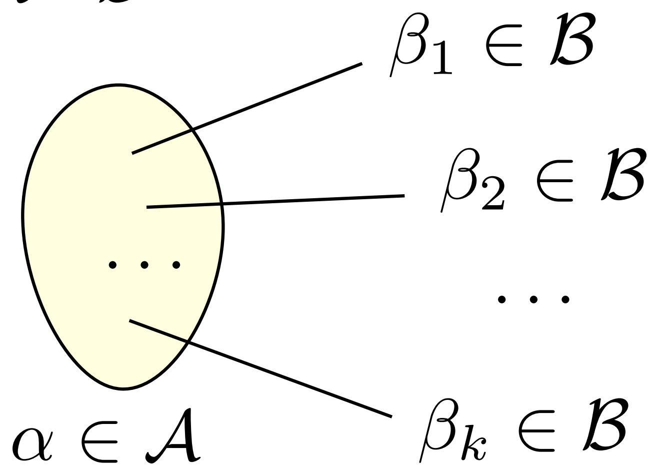


core-size $k = |\alpha|$

Composition schemes

Composition operator: $\mathcal{C} = \mathcal{A} \circ \mathcal{B}$

$\gamma \in \mathcal{C}$ is of the form



core-size $k = |\alpha|$

Generating functions: $C(z) = A(B(z))$

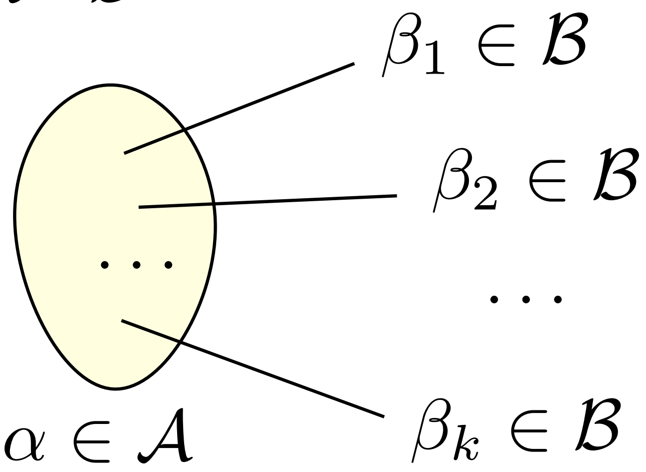
(applies in labeled/unlabeled case)

assume objects in $\mathcal{A}$ have no symmetries

Composition schemes

Composition operator: $\mathcal{C} = \mathcal{A} \circ \mathcal{B}$

$\gamma \in \mathcal{C}$ is of the form



core-size $k = |\alpha|$

Generating functions: $C(z) = A(B(z))$

(applies in labeled/unlabeled case)

assume objects in $\mathcal{A}$ have no symmetries

Rk: basic constructions (with no symmetry) can be formulated with $\circ$

e.g. $\mathrm{Seq}(\mathcal{B}) = \mathcal{A} \circ \mathcal{B}$ where $\mathcal{A} = \mathrm{Seq}$

Classification of composition schemes

$$\mathcal{C} = \mathcal{A} \circ \mathcal{B} \qquad C(z) = A(B(z)) \qquad \rho = \text{singularity of } C(z)$$

supercritical	critical	subcritical
$B(\rho_B) > \rho_A$	$B(\rho_B) = \rho_A$	$B(\rho_B) < \rho_A$
$\rho < \rho_B$ solution of $B(\rho) = \rho_A$ source of ρ: A	$\rho = \rho_B$ source of ρ: A and B	$\rho = \rho_B$ source of ρ: B

Classification of composition schemes

$\mathcal{C} = \mathcal{A} \circ \mathcal{B}$ $\qquad C(z) = A(B(z))$ $\qquad \rho = \text{singularity of } C(z)$

supercritical	critical	subcritical
$B(\rho_B) > \rho_A$	$B(\rho_B) = \rho_A$	$B(\rho_B) < \rho_A$
$\rho < \rho_B$ solution of $B(\rho) = \rho_A$ source of ρ: A	$\rho = \rho_B$ source of ρ: A and B	$\rho = \rho_B$ source of ρ: B

Rk: random object in $\mathcal{C}_n$ [Gourdon'98] [Stufler'22]

- supercritical: $\approx n/\mu$ components, where $\mu = \mathbb{E}(\text{size } \Gamma B(\rho))$
  components '$\sim$' independent draws from $\Gamma B(\rho)$
- subcritical: $\Theta(1)$ components, largest is huge (size $n - \Theta(1)$)

Classification of composition schemes

$$\mathcal{C} = \mathcal{A} \circ \mathcal{B} \qquad C(z) = A(B(z)) \qquad \rho = \text{singularity of } C(z)$$

supercritical	critical	subcritical
$B(\rho_B) > \rho_A$	$B(\rho_B) = \rho_A$	$B(\rho_B) < \rho_A$
$\rho < \rho_B$ solution of $B(\rho) = \rho_A$ source of ρ: A	$\rho = \rho_B$ source of ρ: A and B	$\rho = \rho_B$ source of ρ: B

Rk: random object in $\mathcal{C}_n$ [Gourdon'98] [Stufler'22]

- supercritical: $\approx n/\mu$ components, where $\mu = \mathbb{E}(\text{size } \Gamma B(\rho))$
  components '$\sim$' independent draws from $\Gamma B(\rho)$
- subcritical: $\Theta(1)$ components, largest is huge (size $n - \Theta(1)$)

Rk: In supercritical case,

if $A(y) \sim \kappa(1 - y/\rho_A)^{-s}$ then $C(z) \sim \kappa'(1 - z/\rho)^{-s}$

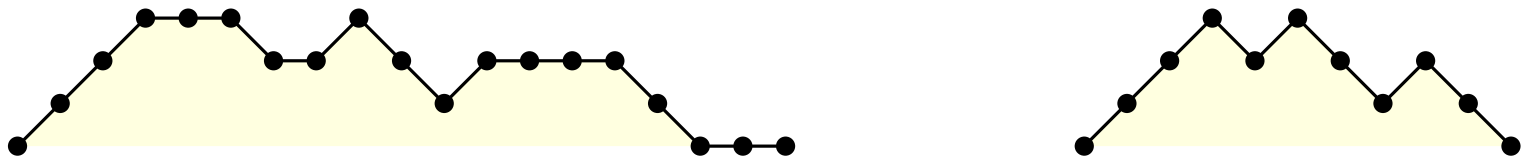
$$\kappa/\kappa' = \mu^s$$

Examples

- Supercritical

$\mathcal{C} =$ Motzkin paths $\qquad \mathcal{A} =$ Dyck paths $\qquad \mathcal{B} = \mathrm{Seq}_{\geq 1}$

(size=number of points = length $+1$)



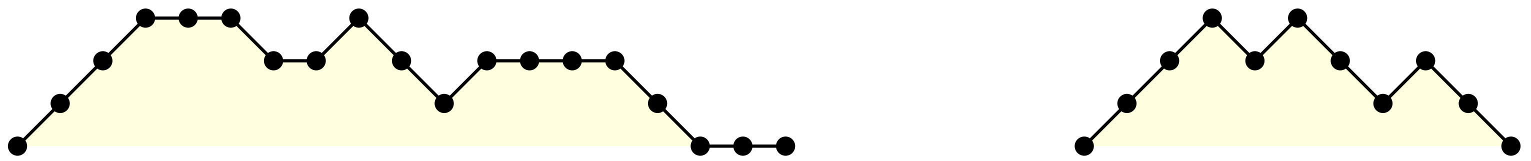
$B(z) = \frac{z}{1-z}$ reaches $\rho_A = 1/2$ at $\rho = 1/3$

Examples

- Supercritical

  $\mathcal{C}$ = Motzkin paths $\qquad$ $\mathcal{A}$ = Dyck paths $\qquad$ $\mathcal{B} = \mathrm{Seq}_{\geq 1}$

  (size=number of points = length $+1$)



$B(z) = \frac{z}{1-z}$ reaches $\rho_A = 1/2$ at $\rho = 1/3$

- Subcritical

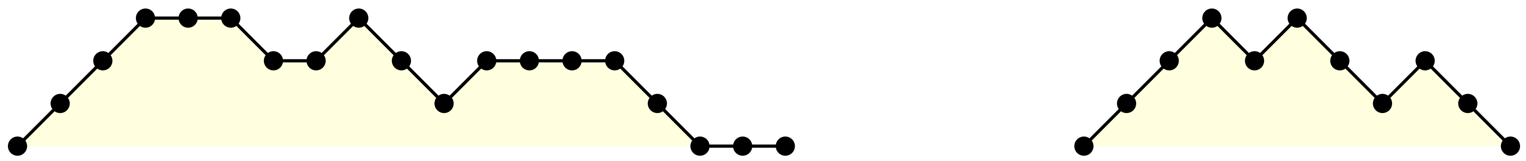
  $\mathcal{C}$ = planar graphs $\qquad$ $\mathcal{A} = \mathrm{Set}$ $\qquad$ $\mathcal{B}$ = connected planar graphs

Examples

- Supercritical

  $\mathcal{C}$ = Motzkin paths $\qquad$ $\mathcal{A}$ = Dyck paths $\qquad$ $\mathcal{B} = \mathrm{Seq}_{\geq 1}$

  (size=number of points = length $+1$)



$$B(z) = \frac{z}{1-z} \text{ reaches } \rho_A = 1/2 \text{ at } \rho = 1/3$$

- Subcritical

  $\mathcal{C}$ = planar graphs $\qquad$ $\mathcal{A} = \mathrm{Set}$ $\qquad$ $\mathcal{B}$ = connected planar graphs

- Critical (later)

  $\mathcal{C}$ = planar maps $\qquad$ $\mathcal{A}$ = 2-connected planar maps $\qquad$ $\mathcal{B} = \mathcal{Z} \times (1 + \mathcal{C})^2$

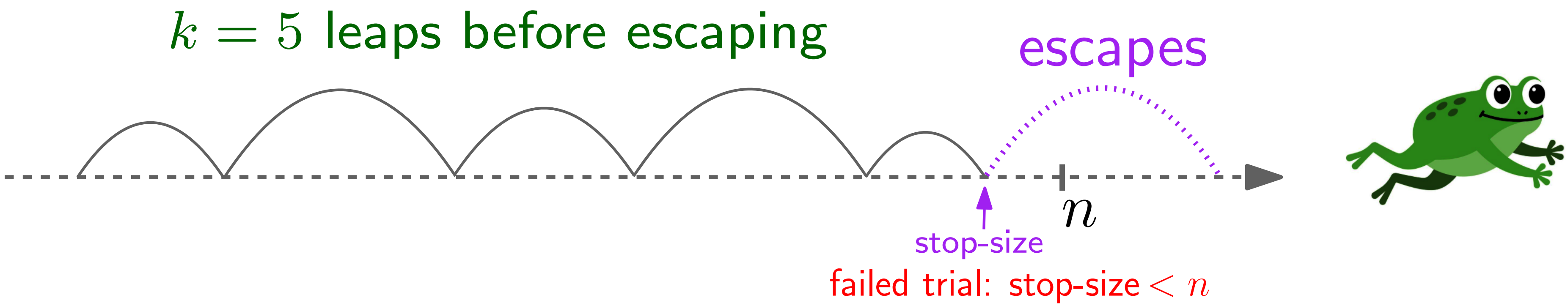
  (block-decomposition of planar maps)

Leap generator for supercritical compositions

[F, Pivoteau'26]

$\mathcal{C} = \mathcal{A} \circ \mathcal{B}$ $\qquad C(z) = A(B(z))$ $\qquad \rho$ such that $B(\rho) = \rho_A$

assumed: exact-size sampler $\Gamma A[k]$ for $\mathcal{A}$, Boltzmann sampler $\Gamma B(x)$ for $\mathcal{B}$



$\Gamma C[n]$: **Repeat**

draw $\gamma_1 \leftarrow \Gamma B(\rho)$, $\gamma_2 \leftarrow \Gamma B(\rho)$, $\ldots$

up to $|\gamma_1| + \cdots + |\gamma_{k+1}| > n$

if $|\gamma_1| + \cdots + |\gamma_k| = n$ (success)

$\alpha \leftarrow \Gamma A[k]$

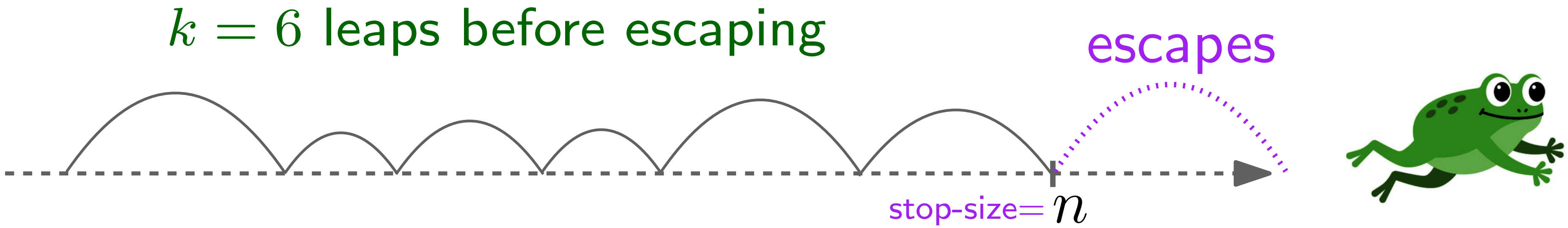
return $\alpha \circ (\beta_1, \ldots, \beta_k)$

Leap generator for supercritical compositions

[F, Pivoteau'26]

$\mathcal{C} = \mathcal{A} \circ \mathcal{B}$ $\quad C(z) = A(B(z))$ $\quad \rho$ such that $B(\rho) = \rho_A$

assumed: exact-size sampler $\Gamma A[k]$ for $\mathcal{A}$, Boltzmann sampler $\Gamma B(x)$ for $\mathcal{B}$



success trial: call $\Gamma A[k]$

$\Gamma C[n]$: **Repeat**

draw $\gamma_1 \leftarrow \Gamma B(\rho)$, $\gamma_2 \leftarrow \Gamma B(\rho)$, $\ldots$

up to $|\gamma_1| + \cdots + |\gamma_{k+1}| > n$

if $|\gamma_1| + \cdots + |\gamma_k| = n$ (success)

$\alpha \leftarrow \Gamma A[k]$

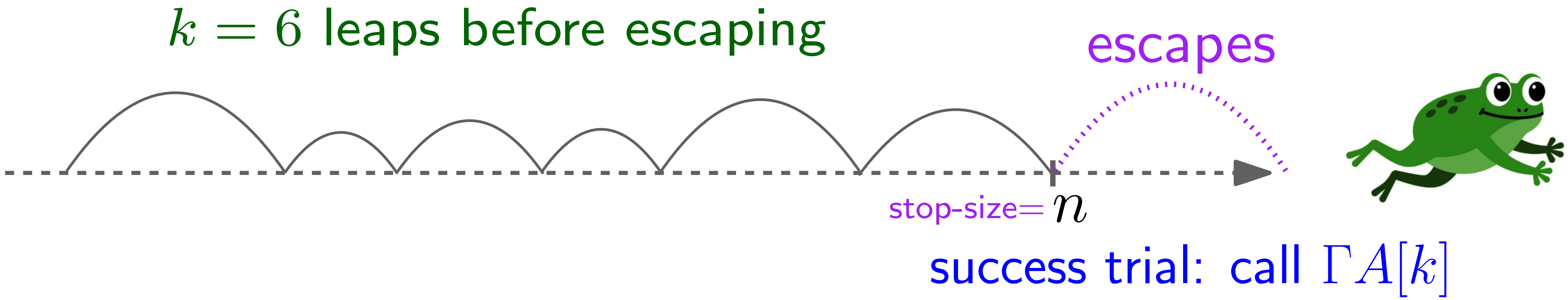
return $\alpha \circ (\beta_1, \ldots, \beta_k)$

Leap generator for supercritical compositions

[F, Pivoteau'26]

$\mathcal{C} = \mathcal{A} \circ \mathcal{B}$ $\qquad C(z) = A(B(z))$ $\qquad \rho$ such that $B(\rho) = \rho_A$

assumed: exact-size sampler $\Gamma A[k]$ for $\mathcal{A}$, Boltzmann sampler $\Gamma B(x)$ for $\mathcal{B}$



$\Gamma C[n]$: **Repeat**

draw $\gamma_1 \leftarrow \Gamma B(\rho)$, $\gamma_2 \leftarrow \Gamma B(\rho)$, $\ldots$

up to $|\gamma_1| + \cdots + |\gamma_{k+1}| > n$

if $|\gamma_1| + \cdots + |\gamma_k| = n$ (success)

$\alpha \leftarrow \Gamma A[k]$

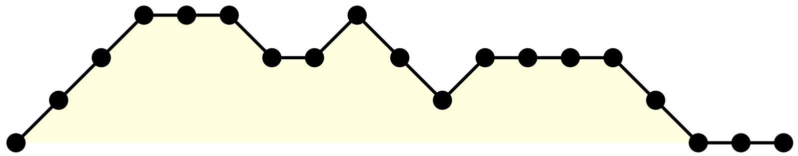
return $\alpha \circ (\beta_1, \ldots, \beta_k)$

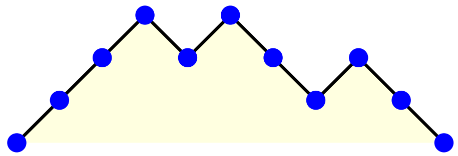
Complexity $O(n)$ if complexities of $\Gamma A[k]$ and $\Gamma B(x)$ are linear

Example on Motzkin paths

$\mathcal{C}$ = Motzkin paths $\mathcal{A}$ = Dyck paths $\mathcal{B} = \mathrm{Seq}_{\geq 1}$

(size=number of points = length $+1$)





Repeat:

generate $\ell_1 \leftarrow \mathrm{Geom}(1/3)$, $\ell_2 \leftarrow \mathrm{Geom}(1/3), \ldots$

up to $\ell_1 + \ldots + \ell_{k+1} > n$

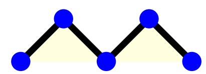
if $\ell_1 + \cdots + \ell_k = n$ and k is odd (success)

generate random Dyck path P of length $k-1$

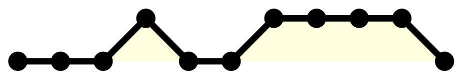
insert horizontal segment of length $\ell_i - 1$ in ith point of P

Ex: $(n = 11)$: $\ell_1 = 3,\ \ell_2 = 1,\ \ell_3 = 2,\ \ell_4 = 4,\ \ell_5 = 1$

Dyck path

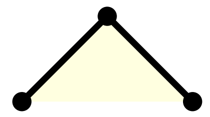


return



Induced distribution in small sizes

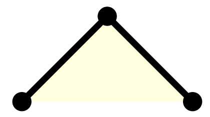
$n = 3$

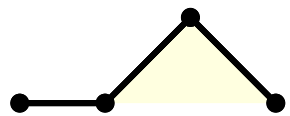
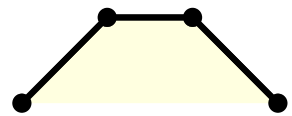
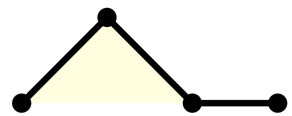


each trial	$\left(\frac{1}{3}\right)^2 \frac{2}{3} = \frac{2}{27}$	$\left(\frac{2}{3}\right)^3 = \frac{8}{27}$	$\Rightarrow \quad \mathbb{P}(\text{success}) = 10/27$
proba	$1/5$	$4/5$	

Induced distribution in small sizes

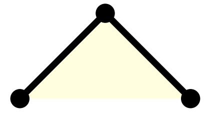
$n = 3$

			
each trial	$\left(\frac{1}{3}\right)^2 \frac{2}{3} = \frac{2}{27}$	$\left(\frac{2}{3}\right)^3 = \frac{8}{27}$	$\Rightarrow \quad \mathbb{P}(\text{success}) = 10/27$
proba	$1/5$	$4/5$	

size $n = 4$				
each trial	$\left(\frac{1}{3}\right)^3 \frac{2}{3} = \frac{2}{81}$	$\frac{1}{3}\left(\frac{2}{3}\right)^3 = \frac{8}{81}$	$\frac{1}{3}\left(\frac{2}{3}\right)^3 = \frac{8}{81}$	$\frac{1}{3}\left(\frac{2}{3}\right)^3 = \frac{8}{81}$
proba	$1/13$	$4/13$	$4/13$	$4/13$

Induced distribution in small sizes

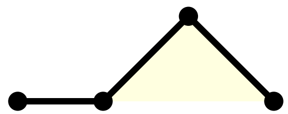
$n = 3$

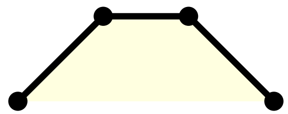


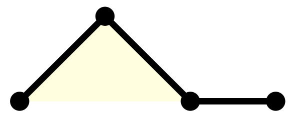
each trial	$\left(\frac{1}{3}\right)^2 \frac{2}{3} = \frac{2}{27}$	$\left(\frac{2}{3}\right)^3 = \frac{8}{27}$	$\Rightarrow$	$\mathbb{P}(\text{success}) = 10/27$
proba	$1/5$	$4/5$		

---

size $n = 4$

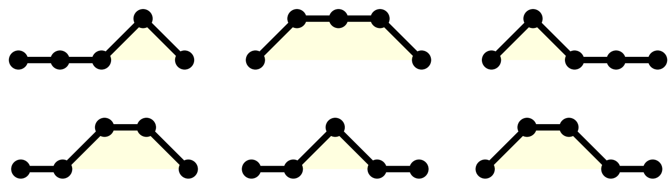


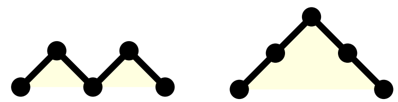




each trial	$\left(\frac{1}{3}\right)^3 \frac{2}{3} = \frac{2}{81}$	$\frac{1}{3}\left(\frac{2}{3}\right)^3 = \frac{8}{81}$	$\frac{1}{3}\left(\frac{2}{3}\right)^3 = \frac{8}{81}$	$\frac{1}{3}\left(\frac{2}{3}\right)^3 = \frac{8}{81}$
proba	$1/13$	$4/13$	$4/13$	$4/13$

$n = 5$





each trial	$\left(\frac{1}{3}\right)^4 \frac{2}{3} = \frac{2}{243}$	$\left(\frac{1}{3}\right)^2 \left(\frac{2}{3}\right)^3 = \frac{8}{243}$	$\left(\frac{2}{3}\right)^5 \frac{1}{2} = \frac{16}{243}$
proba	$1/41$	$4/41$	$8/41$

Distorsion from uniform distribution

Let π'_n be obtained distribution on $\mathcal{C}_n$ (leap distribution)

- at each trial, for $\gamma \in \mathcal{C}_n$ with core-size k

$$\mathbb{P}(\gamma) = \frac{\rho^{n_1}}{B(\rho)} \cdots \frac{\rho^{n_k}}{B(\rho)} \quad \frac{1}{|\mathcal{A}_k|} \quad = \frac{\rho^n}{a_k B(\rho)^k}$$

Distorsion from uniform distribution

Let π'_n be obtained distribution on $\mathcal{C}_n$ (leap distribution)

- at each trial, for $\gamma \in \mathcal{C}_n$ with core-size k

$$\mathbb{P}(\gamma) = \frac{\rho^{n_1}}{B(\rho)} \cdots \frac{\rho^{n_k}}{B(\rho)} \frac{1}{|\mathcal{A}_k|} = \frac{\rho^n}{a_k B(\rho)^k}$$

- $q_n :=$ success probability of trial $= \sum_{\gamma \in \mathcal{C}_n} \mathbb{P}(\gamma)$

$$\Rightarrow \boxed{\pi'_n(\gamma) = \frac{1}{q_n} \frac{\rho^n}{a_k B(\rho)^k} = \underbrace{\frac{1}{c_n}}_{\pi_n(\gamma)} \cdot \underbrace{\frac{1}{q_n} \frac{c_n \rho^n}{a_k B(\rho)^k}}_{\text{distorsion-factor } d_{n,k}}}$$

$\pi_n(\gamma)$

uniform distribution

distorsion-factor $d_{n,k}$

Distorsion from uniform distribution

Let π'_n be obtained distribution on $\mathcal{C}_n$ (leap distribution)

- at each trial, for $\gamma \in \mathcal{C}_n$ with core-size k

$$\mathbb{P}(\gamma) = \frac{\rho^{n_1}}{B(\rho)} \cdots \frac{\rho^{n_k}}{B(\rho)} \quad \frac{1}{|\mathcal{A}_k|} \quad = \frac{\rho^n}{a_k B(\rho)^k}$$

- $q_n :=$ success probability of trial $= \sum_{\gamma \in \mathcal{C}_n} \mathbb{P}(\gamma)$

$\sim \frac{1}{\mu}$, with $\mu = \mathbb{E}(\text{size of } \Gamma B(\rho))$

$$\Rightarrow \quad \boxed{\pi'_n(\gamma) = \frac{1}{q_n} \frac{\rho^n}{a_k B(\rho)^k} \quad = \underbrace{\frac{1}{c_n}}_{\substack{\pi_n(\gamma) \\ \text{uniform distribution}}} \cdot \underbrace{\frac{1}{q_n} \quad \frac{c_n \rho^n}{a_k B(\rho)^k}}_{\text{distorsion-factor } d_{n,k}}}$$

Distorsion from uniform distribution

Let π'_n be obtained distribution on $\mathcal{C}_n$ (leap distribution)

- at each trial, for $\gamma \in \mathcal{C}_n$ with core-size k

$$\mathbb{P}(\gamma) = \frac{\rho^{n_1}}{B(\rho)} \cdots \frac{\rho^{n_k}}{B(\rho)} \frac{1}{|\mathcal{A}_k|} = \frac{\rho^n}{a_k B(\rho)^k}$$

- $q_n :=$ success probability of trial $= \sum_{\gamma \in \mathcal{C}_n} \mathbb{P}(\gamma)$

$\sim \frac{1}{\mu}$, with $\mu = \mathbb{E}(\text{size of } \Gamma B(\rho))$

$$\Rightarrow \boxed{\pi'_n(\gamma) = \frac{1}{q_n} \frac{\rho^n}{a_k B(\rho)^k} = \frac{1}{c_n} \cdot \frac{1}{q_n} \frac{c_n \rho^n}{a_k B(\rho)^k}}$$

$\pi_n(\gamma)$
uniform distribution

distorsion-factor $d_{n,k}$

- **Asymptotics** of $d_{n,k}$

if $A(y) \sim \kappa(1 - y/\rho_A)^{-s}$ then $C(z) \sim \kappa'(1 - z/\rho)^{-s}$ $\kappa/\kappa' = \mu^s$

$$d_{n,k} \sim \mu \cdot \frac{\kappa' \, n^{s-1}}{\kappa \, k^{s-1}} \sim \left(\mu \, k/n\right)^{1-s} \quad \text{for } k \text{ large}$$

Asymptotic uniformity

- For two probability distributions π, π' on E

$$d_{\mathrm{TV}}(\pi, \pi') = \frac{1}{2} \sum_{x \in E} |\pi(x) - \pi'(x)|$$

Asymptotic uniformity

- For two probability distributions π, π' on E

$$d_{\mathrm{TV}}(\pi, \pi') = \tfrac{1}{2} \sum_{x \in E} |\pi(x) - \pi'(x)|$$

- Given $\mathcal{C}_{n,k} = \{\gamma \in \mathcal{C}_n \text{ of core-size } k\}$ $\qquad c_{n,k} = \mathrm{Card}(\mathcal{C}_{n,k})$

$$\boxed{d_{\mathrm{TV}}(\pi_n, \pi'_n) = \frac{1}{2} \sum_k \frac{c_{n,k}}{c_n} \left| 1 - d_{n,k} \right|}$$

Asymptotic uniformity

- For two probability distributions π, π' on E

$$d_{\mathrm{TV}}(\pi, \pi') = \tfrac{1}{2} \sum_{x \in E} |\pi(x) - \pi'(x)|$$

- Given $\mathcal{C}_{n,k} = \{\gamma \in \mathcal{C}_n \text{ of core-size } k\}$ $\qquad c_{n,k} = \mathrm{Card}(\mathcal{C}_{n,k})$

$$\boxed{d_{\mathrm{TV}}(\pi_n, \pi'_n) = \frac{1}{2} \sum_k \frac{c_{n,k}}{c_n} \Big| 1 - d_{n,k} \Big|}$$

$\mu = \mathbb{E}(\text{size of } \Gamma B(\rho))$
$\sigma^2 = \mathbb{V}(\text{size of } \Gamma B(\rho))$

core-size distribution $\frac{c_{n,k}}{c_n} \sim$ Gaussian$(an, b^2 n)$ $\qquad a = \frac{1}{\mu} \qquad b = \frac{\sigma}{\mu^{3/2}}$

Asymptotic uniformity

- For two probability distributions π, π' on E

$$d_{\mathrm{TV}}(\pi, \pi') = \tfrac{1}{2} \sum_{x \in E} |\pi(x) - \pi'(x)|$$

- Given $\mathcal{C}_{n,k} = \{\gamma \in \mathcal{C}_n \text{ of core-size } k\}$ $\qquad c_{n,k} = \mathrm{Card}(\mathcal{C}_{n,k})$

$$\boxed{d_{\mathrm{TV}}(\pi_n, \pi'_n) = \frac{1}{2} \sum_k \frac{c_{n,k}}{c_n} \left| 1 - d_{n,k} \right|}$$

$\mu = \mathbb{E}(\text{size of } \Gamma B(\rho))$

$\sigma^2 = \mathbb{V}(\text{size of } \Gamma B(\rho))$

core-size distribution $\frac{c_{n,k}}{c_n} \sim$ Gaussian$(an, b^2 n)$ $\qquad a = \frac{1}{\mu} \qquad b = \frac{\sigma}{\mu^{3/2}}$

$$k = an + x\, b\sqrt{n} \longrightarrow \frac{c_{n,k}}{c_n} \sim \frac{1}{b\sqrt{2\pi\, n}} e^{-x^2/2}$$

Asymptotic uniformity

- For two probability distributions π, π' on E
$$d_{\mathrm{TV}}(\pi, \pi') = \tfrac{1}{2} \textstyle\sum_{x \in E} |\pi(x) - \pi'(x)|$$

- Given $\mathcal{C}_{n,k} = \{\gamma \in \mathcal{C}_n \text{ of core-size } k\}$ $\qquad c_{n,k} = \mathrm{Card}(\mathcal{C}_{n,k})$

$$\boxed{d_{\mathrm{TV}}(\pi_n, \pi'_n) = \frac{1}{2} \sum_k \frac{c_{n,k}}{c_n} \Big| 1 - d_{n,k} \Big|}$$

$\mu = \mathbb{E}(\text{size of } \Gamma B(\rho))$
$\sigma^2 = \mathbb{V}(\text{size of } \Gamma B(\rho))$

core-size distribution $\frac{c_{n,k}}{c_n} \sim$ Gaussian$(an, b^2 n)$ $\qquad a = \frac{1}{\mu} \qquad b = \frac{\sigma}{\mu^{3/2}}$

$$k = an + x\,b\sqrt{n}$$

$$\frac{c_{n,k}}{c_n} \sim \frac{1}{b\sqrt{2\pi\, n}} e^{-x^2/2}$$

$$d_{n,k} \sim (\mu\, k/n)^{1-s} \sim 1 + \frac{b}{a}(1-s)\frac{x}{\sqrt{n}}$$

Asymptotic uniformity

- For two probability distributions π, π' on E

$$d_{\mathrm{TV}}(\pi, \pi') = \tfrac{1}{2} \textstyle\sum_{x \in E} |\pi(x) - \pi'(x)|$$

- Given $\mathcal{C}_{n,k} = \{\gamma \in \mathcal{C}_n \text{ of core-size } k\}$ $\qquad c_{n,k} = \mathrm{Card}(\mathcal{C}_{n,k})$

$$\boxed{d_{\mathrm{TV}}(\pi_n, \pi'_n) = \frac{1}{2} \sum_k \frac{c_{n,k}}{c_n} \Big| 1 - d_{n,k} \Big|}$$

$$\mu = \mathbb{E}(\text{size of } \Gamma B(\rho))$$
$$\sigma^2 = \mathbb{V}(\text{size of } \Gamma B(\rho))$$

core-size distribution $\frac{c_{n,k}}{c_n} \sim \text{Gaussian}(an, b^2 n)$ $\qquad a = \frac{1}{\mu} \qquad b = \frac{\sigma}{\mu^{3/2}}$

$$k = an + x\, b\sqrt{n}$$

$$\frac{c_{n,k}}{c_n} \sim \frac{1}{b\sqrt{2\pi\, n}} e^{-x^2/2}$$

$$d_{n,k} \sim (\mu\, k/n)^{1-s} \sim 1 + \frac{b}{a}(1-s)\frac{x}{\sqrt{n}}$$

$$\Rightarrow \boxed{d_{\mathrm{TV}}(\pi_n, \pi'_n) \sim \frac{c}{\sqrt{n}}}$$

$$c = \frac{b}{a}|1-s|\frac{1}{2\sqrt{2\pi}} \int_x e^{-x^2/2}|x|\,\mathrm{d}x$$
$$= \frac{\sigma\,|1-s|}{\sqrt{2\pi\mu}}$$

First order correction of distorsion

$$k = \frac{n}{\mu} + x\,\frac{\sigma}{\mu^{3/2}}\sqrt{n} \qquad d_{n,k} \sim 1 + \frac{\sigma}{\sqrt{\mu}}(1-s)\frac{x}{\sqrt{n}} =: f_n(x)$$

First order correction of distorsion

$$k = \frac{n}{\mu} + x\,\frac{\sigma}{\mu^{3/2}}\sqrt{n} \qquad d_{n,k} \sim 1 + \frac{\sigma}{\sqrt{\mu}}(1-s)\frac{x}{\sqrt{n}} =: f_n(x)$$

$\Gamma\hat{C}^{\,\mathrm{rej}}[n]:$ **Repeat**

$\gamma \leftarrow \Gamma\widehat{C}[n]$

$k \leftarrow$ core-size(γ); $\quad x \leftarrow \frac{k-n/\mu}{\sigma\mu^{-3/2}\sqrt{n}}$

$p_{\,\mathrm{keep}} \leftarrow \min(1, \frac{1}{2f_n(x)})$

if $\mathrm{Bern}(p_{\,\mathrm{keep}})$ return γ

First order correction of distorsion

$$k = \frac{n}{\mu} + x\,\frac{\sigma}{\mu^{3/2}}\sqrt{n} \qquad d_{n,k} \sim 1 + \frac{\sigma}{\sqrt{\mu}}(1-s)\frac{x}{\sqrt{n}} =: f_n(x)$$

$\Gamma\hat{C}^{\,\mathrm{rej}}[n]$: **Repeat**

$\gamma \leftarrow \Gamma\widehat{C}[n]$

$k \leftarrow$ core-size(γ); $x \leftarrow \frac{k - n/\mu}{\sigma\mu^{-3/2}\sqrt{n}}$

$p_{\,\mathrm{keep}} \leftarrow \min(1, \frac{1}{2 f_n(x)})$

if $\mathrm{Bern}(p_{\,\mathrm{keep}})$ return γ

- Complexity still $O(n)$ (expected number of trials ~ 2)

First order correction of distorsion

$$k = \frac{n}{\mu} + x\,\frac{\sigma}{\mu^{3/2}}\sqrt{n} \qquad d_{n,k} \sim 1 + \frac{\sigma}{\sqrt{\mu}}(1-s)\,\frac{x}{\sqrt{n}} =: f_n(x)$$

$\Gamma\hat{C}^{\,\mathrm{rej}}[n]$: **Repeat**

$\gamma \leftarrow \Gamma\widehat{C}[n]$

$k \leftarrow$ core-size(γ); $x \leftarrow \frac{k-n/\mu}{\sigma\mu^{-3/2}\sqrt{n}}$

$p_{\,\mathrm{keep}} \leftarrow \min(1, \frac{1}{2f_n(x)})$

if $\mathrm{Bern}(p_{\,\mathrm{keep}})$ return γ

- Complexity still $O(n)$ (expected number of trials ~ 2)
- Objects $\gamma \in \mathcal{C}_{n,k}$ satisfy $\pi_n^{\mathrm{rej}}(\gamma) = \frac{1}{c_n} d_{n,k}^{\mathrm{rej}}$ $\qquad d_{n,k}^{\mathrm{rej}} \sim 1 + (a_0 + a_2 x^2)\frac{1}{n}$

First order correction of distorsion

$$k = \frac{n}{\mu} + x\,\frac{\sigma}{\mu^{3/2}}\sqrt{n} \qquad d_{n,k} \sim 1 + \frac{\sigma}{\sqrt{\mu}}(1-s)\,\frac{x}{\sqrt{n}} =: f_n(x)$$

$\Gamma\hat{C}^{\,\mathrm{rej}}[n]$: **Repeat**

$\gamma \leftarrow \Gamma\widehat{C}[n]$

$k \leftarrow$ core-size(γ); $\quad x \leftarrow \frac{k-n/\mu}{\sigma\mu^{-3/2}\sqrt{n}}$

$p_{\,\mathrm{keep}} \leftarrow \min(1, \frac{1}{2f_n(x)})$

if $\mathrm{Bern}(p_{\,\mathrm{keep}})$ return γ

- Complexity still $O(n)$ (expected number of trials ~ 2)
- Objects $\gamma \in \mathcal{C}_{n,k}$ satisfy $\pi_n^{\mathrm{rej}}(\gamma) = \frac{1}{c_n} d_{n,k}^{\mathrm{rej}}$ $\qquad d_{n,k}^{\mathrm{rej}} \sim 1 + (a_0 + a_2x^2)\frac{1}{n}$

$$d_{\mathrm{TV}}(\pi_n, \pi_n^{\mathrm{rej}}) = \frac{1}{2}\sum_k \frac{c_{n,k}}{c_n}\left|1 - d_{n,k}^{\mathrm{rej}}\right| \sim \frac{c'}{n}$$

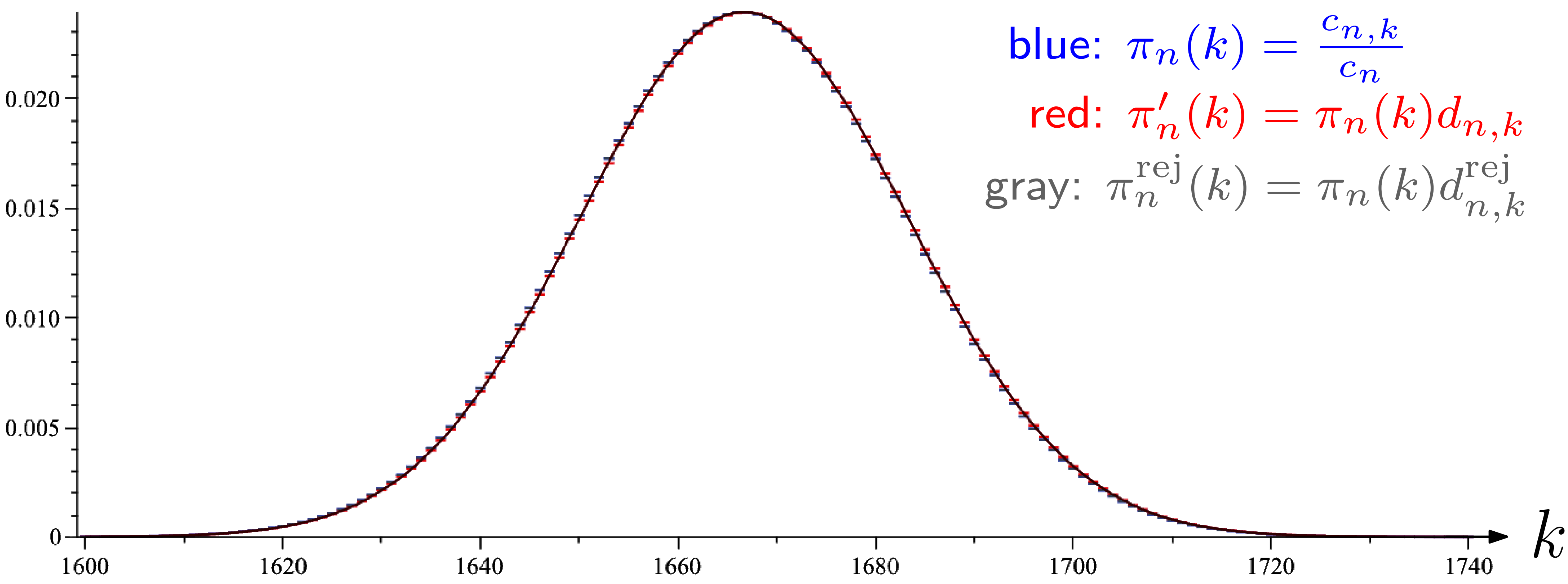
for c' explicit

Plots of core-size distributions

For Motzkin paths of size $n = 5000$

plots of core-size distribution under uniform (blue), leap (red), leap-reject (gray)

superimposed with the Gaussian limit curve (rescaled for size n)



gray and blue are almost indistinguishable

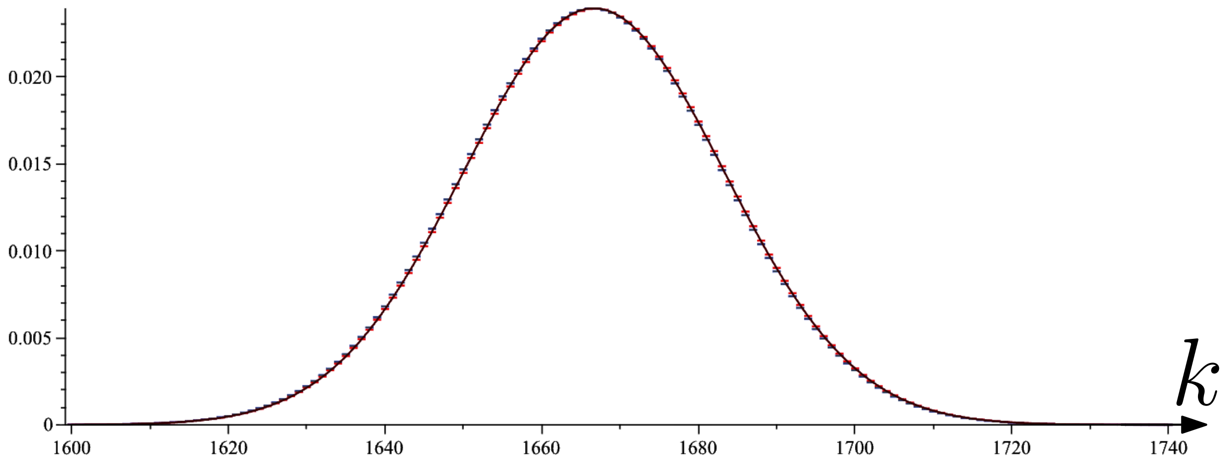
Plots of core-size distributions

core-size distribution under

uniform: $\pi_n(k) = \frac{c_{n,k}}{c_n}$

leap: $\pi'_n(k) = \pi_n(k) d_{n,k}$

leap-reject: $\pi_n^{\mathrm{rej}}(k) = \pi_n(k) d_{n,k}^{\mathrm{rej}}$



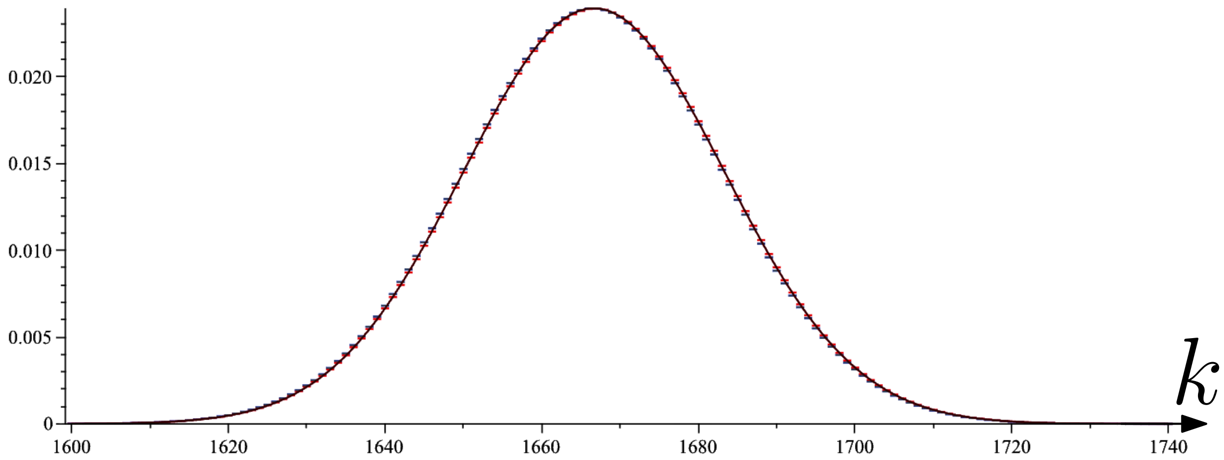
Plots of core-size distributions

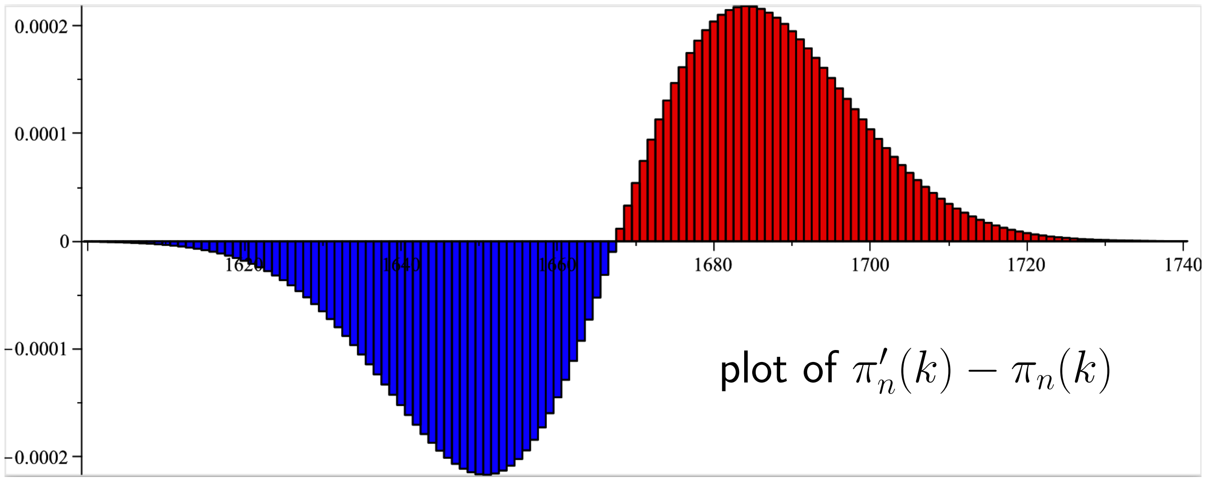
core-size distribution under

uniform: $\pi_n(k) = \frac{c_{n,k}}{c_n}$

leap: $\pi'_n(k) = \pi_n(k) d_{n,k}$

leap-reject: $\pi_n^{\mathrm{rej}}(k) = \pi_n(k) d_{n,k}^{\mathrm{rej}}$





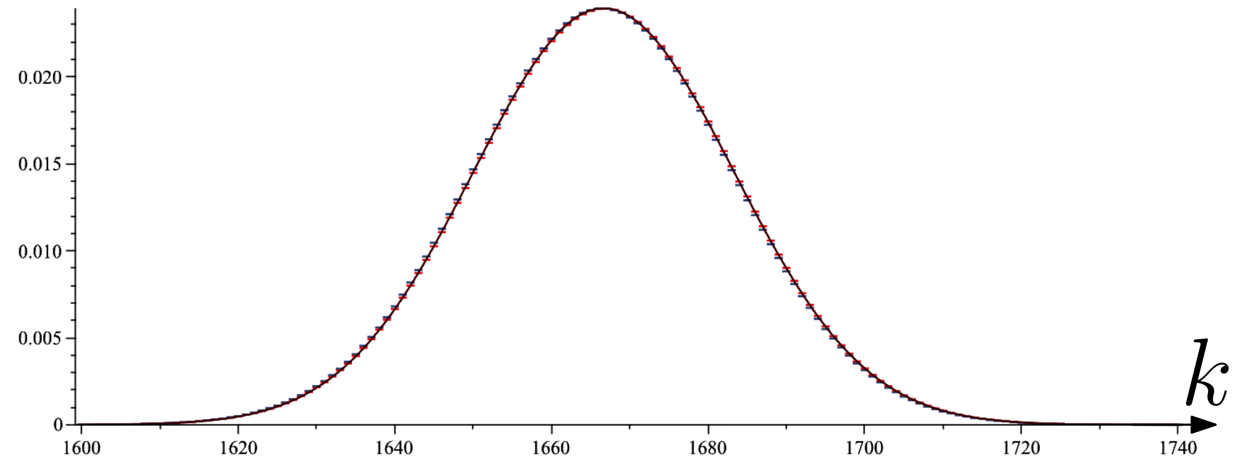
Plots of core-size distributions

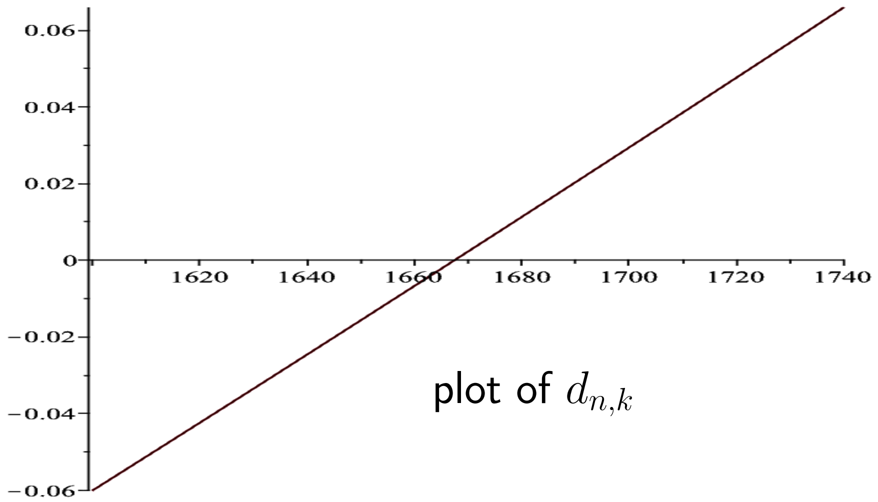
core-size distribution under

uniform: $\pi_n(k) = \frac{c_{n,k}}{c_n}$

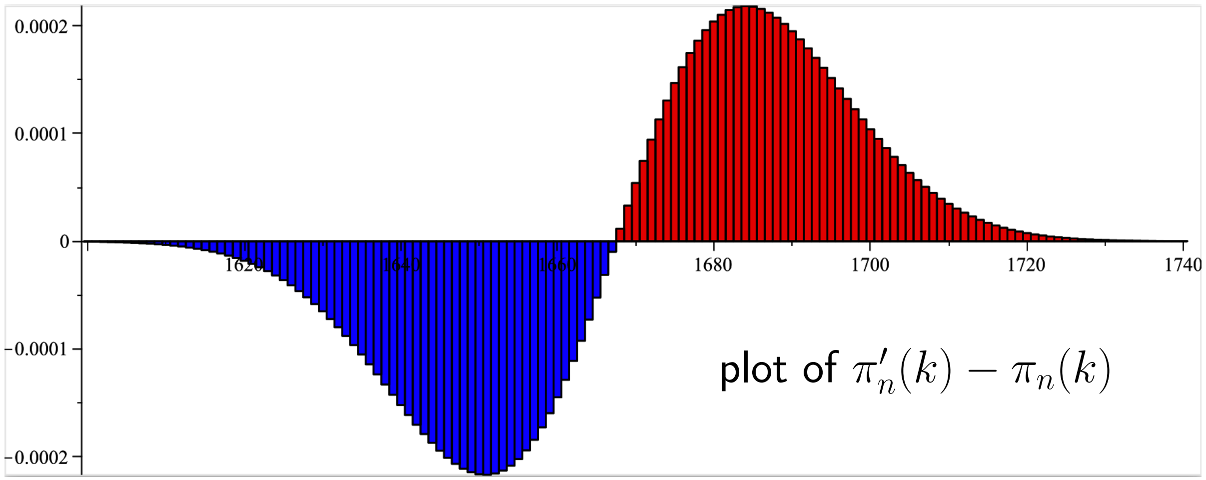
leap: $\pi'_n(k) = \pi_n(k) d_{n,k}$

leap-reject: $\pi_n^{\mathrm{rej}}(k) = \pi_n(k) d_{n,k}^{\mathrm{rej}}$





plot of $d_{n,k}$



plot of $\pi'_n(k) - \pi_n(k)$

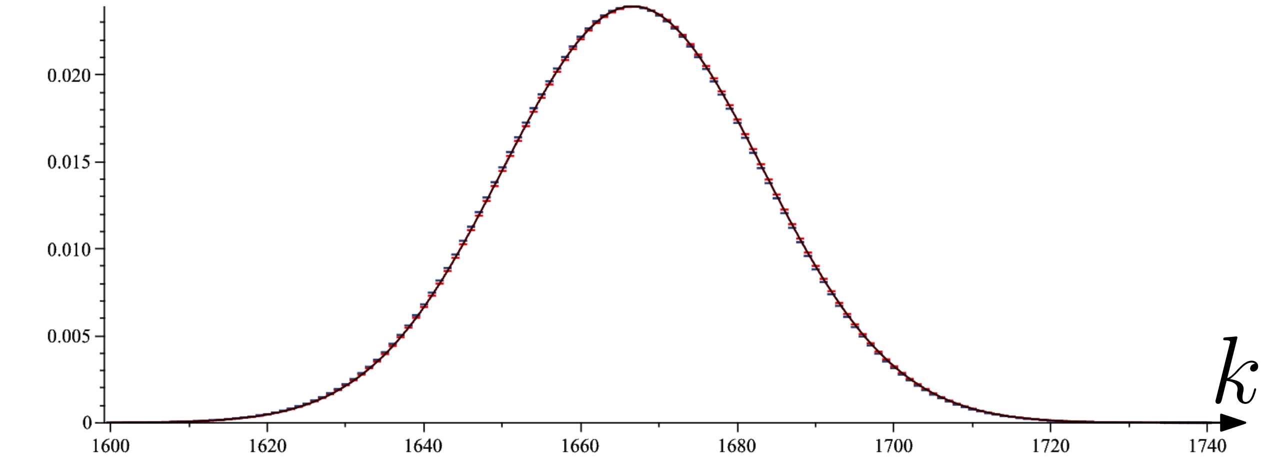
Plots of core-size distributions

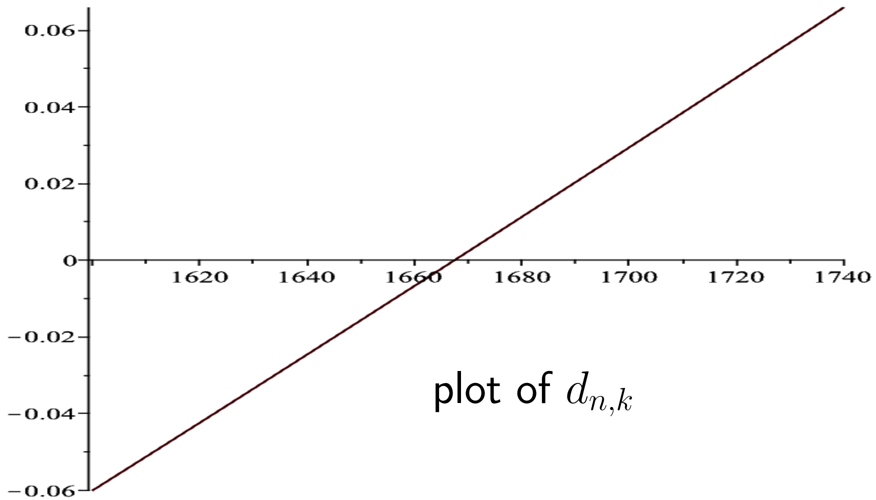
core-size distribution under

uniform: $\pi_n(k) = \frac{c_{n,k}}{c_n}$

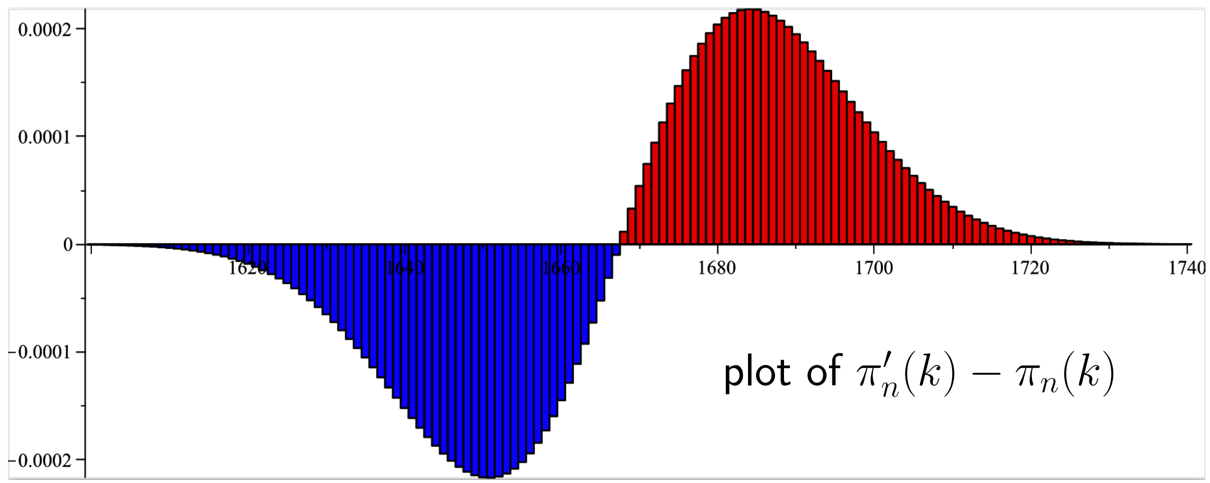
leap: $\pi'_n(k) = \pi_n(k) d_{n,k}$

leap-reject: $\pi_n^{\mathrm{rej}}(k) = \pi_n(k) d_{n,k}^{\mathrm{rej}}$

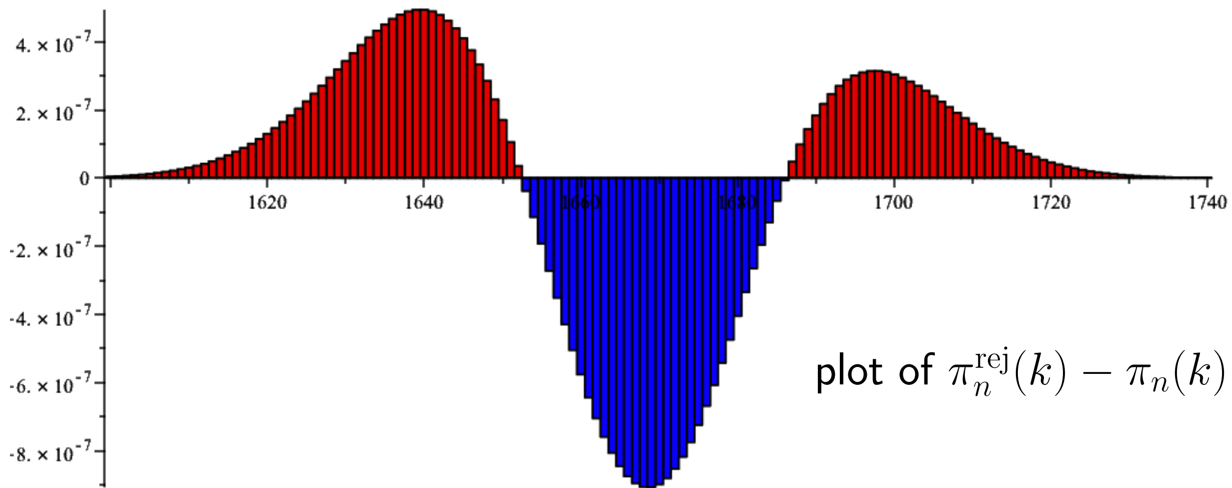




plot of $d_{n,k}$



plot of $\pi'_n(k) - \pi_n(k)$



plot of $\pi_n^{\mathrm{rej}}(k) - \pi_n(k)$

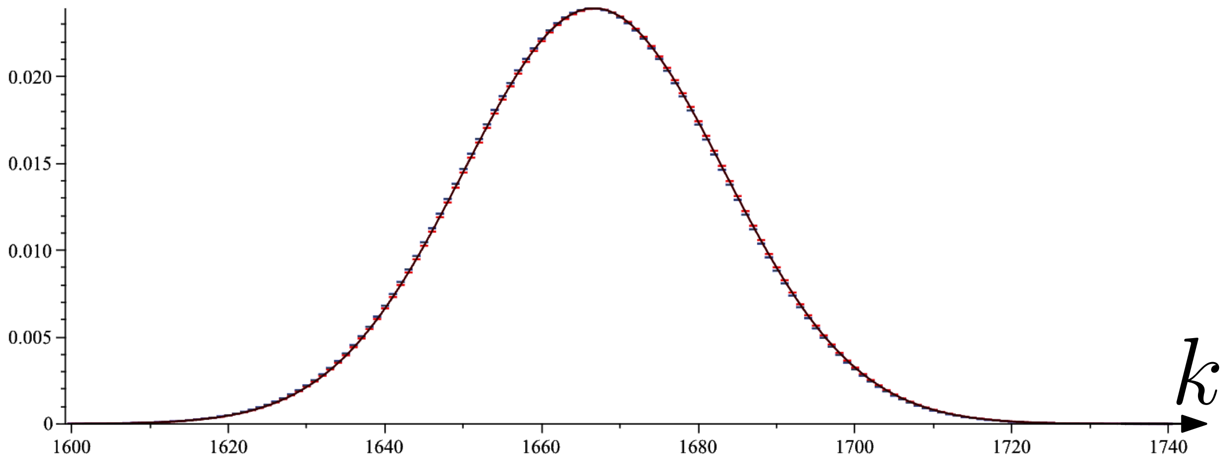
Plots of core-size distributions

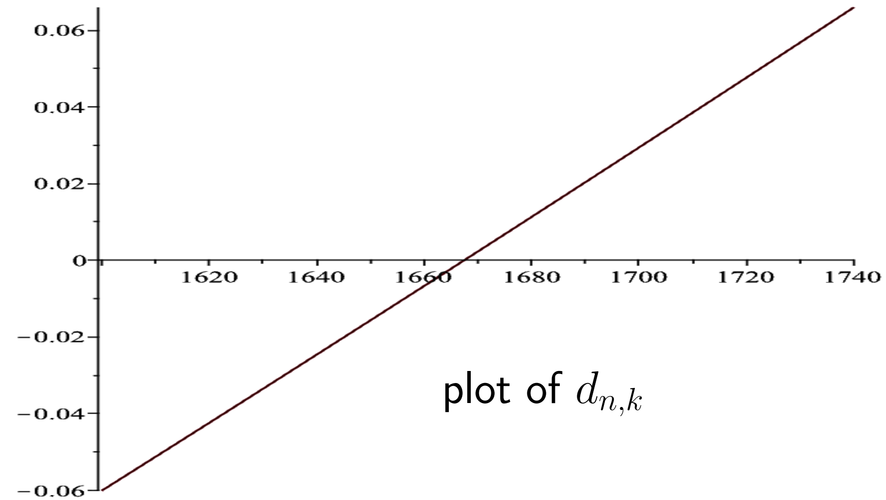
core-size distribution under

uniform: $\pi_n(k) = \dfrac{c_{n,k}}{c_n}$

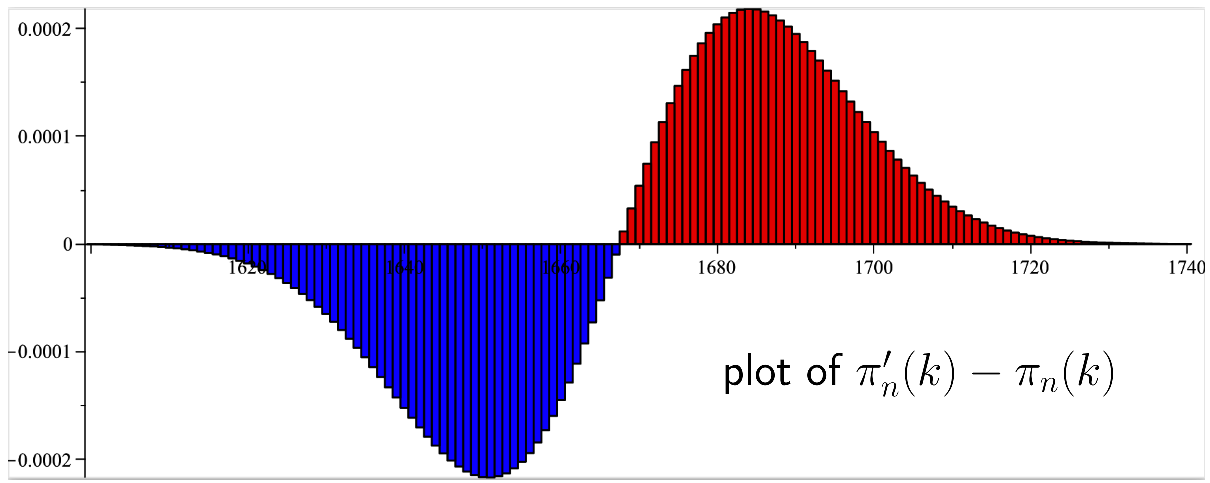
leap: $\pi'_n(k) = \pi_n(k) d_{n,k}$

leap-reject: $\pi^{\mathrm{rej}}_n(k) = \pi_n(k) d^{\mathrm{rej}}_{n,k}$

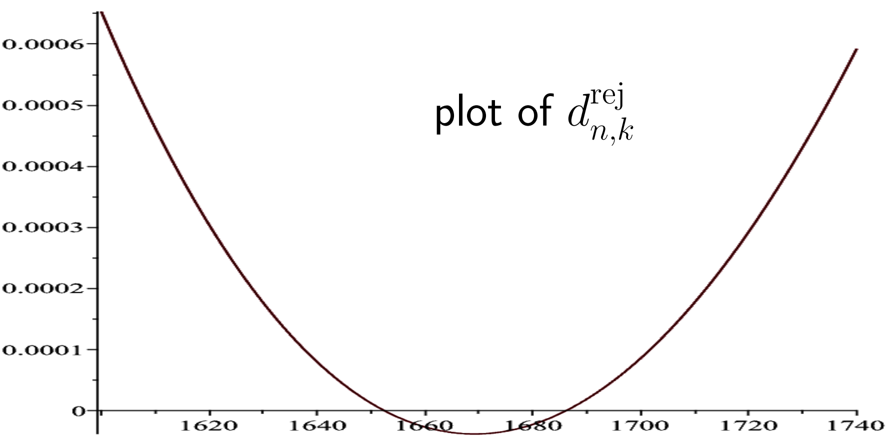




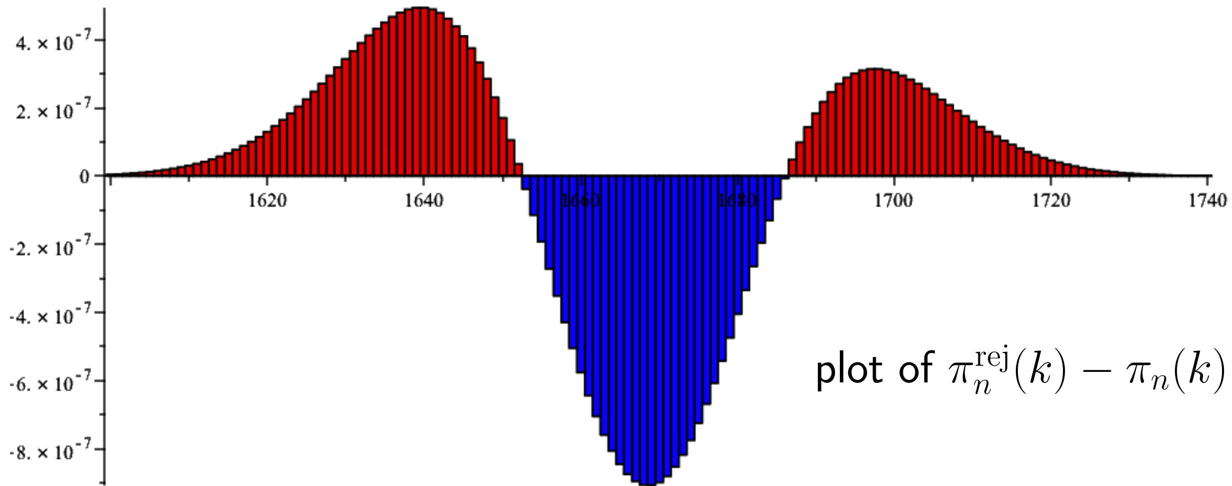
plot of $d_{n,k}$



plot of $\pi'_n(k) - \pi_n(k)$

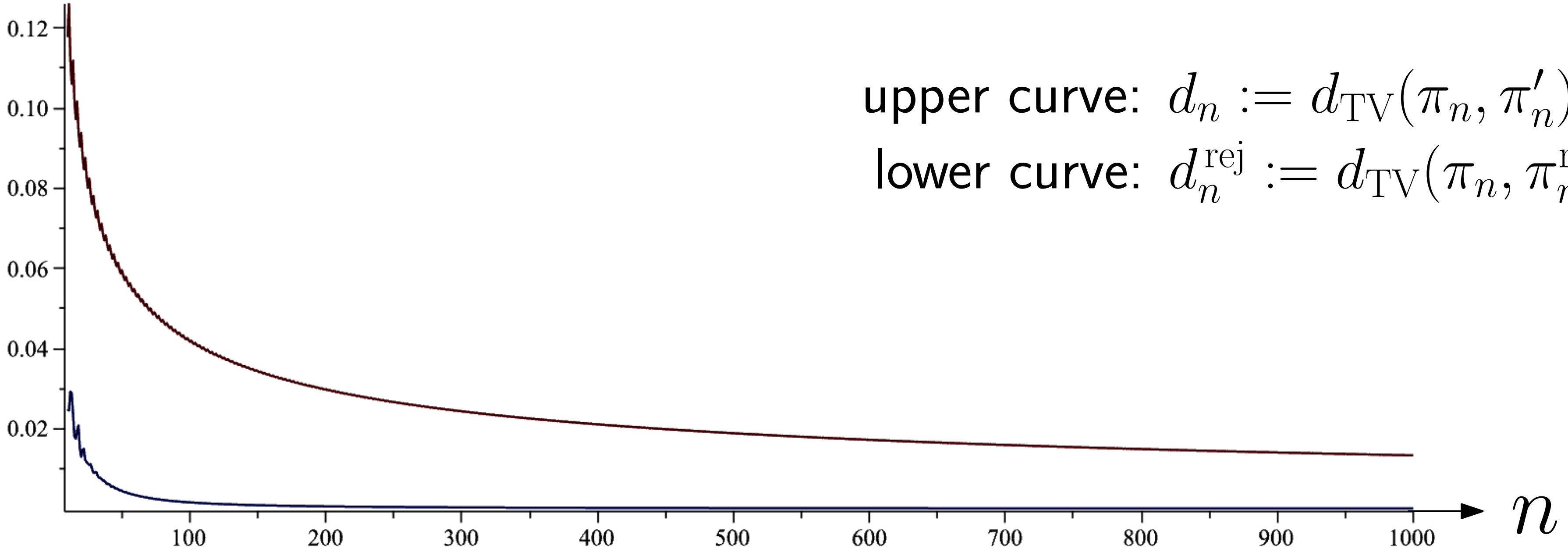


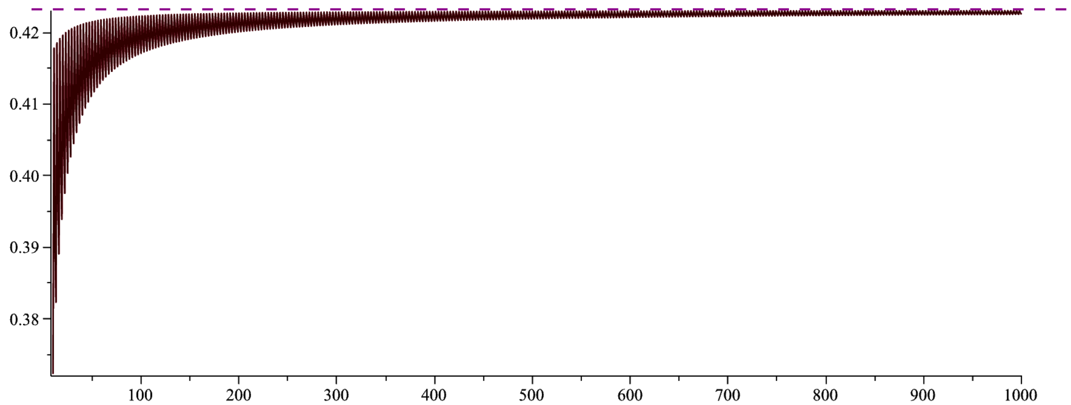
plot of $d^{\mathrm{rej}}_{n,k}$



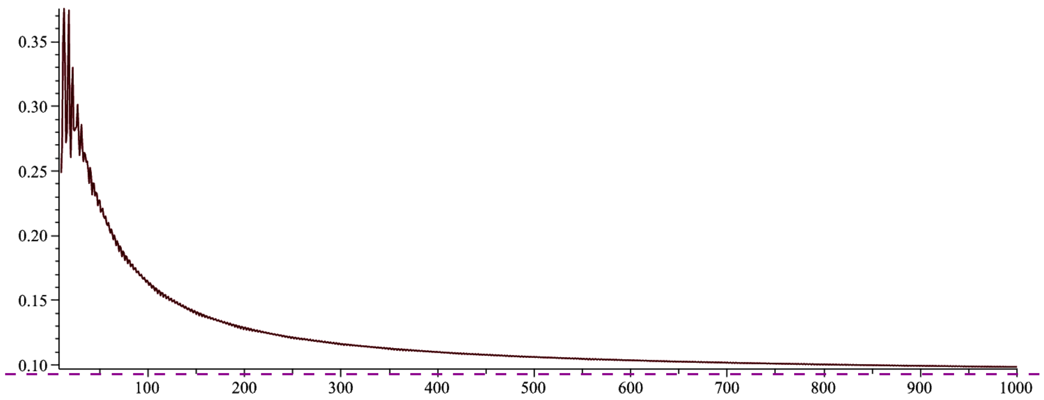
plot of $\pi^{\mathrm{rej}}_n(k) - \pi_n(k)$

Plots of total variation distances



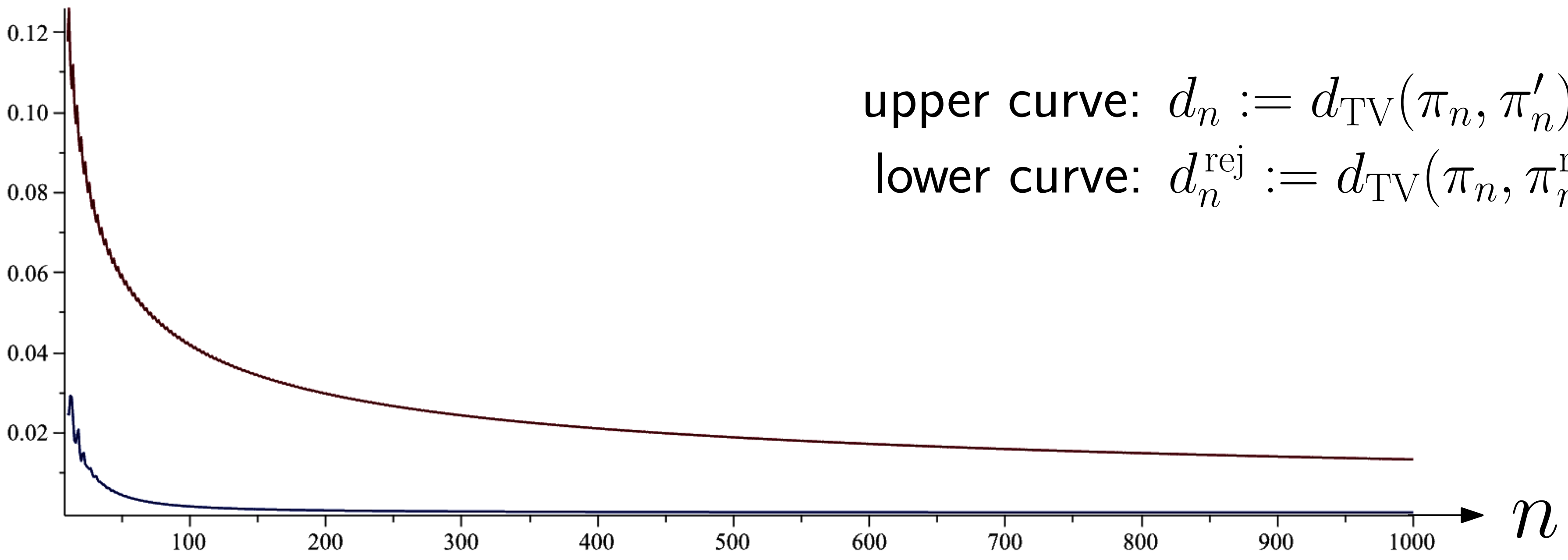


plot of $n^{1/2} \cdot d_n$
(predicted limit ≈ 0.42314)

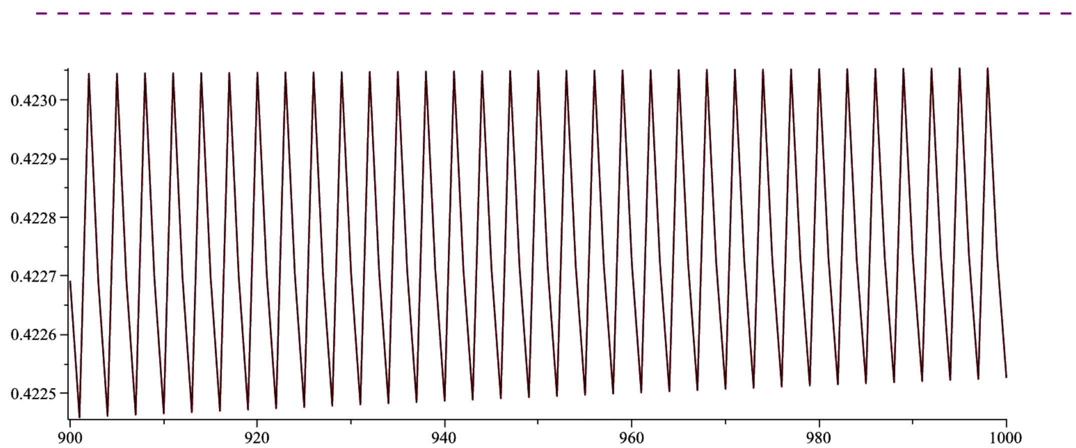


plot of $n \cdot d_n^{\mathrm{rej}}$
(predicted limit ≈ 0.09074)

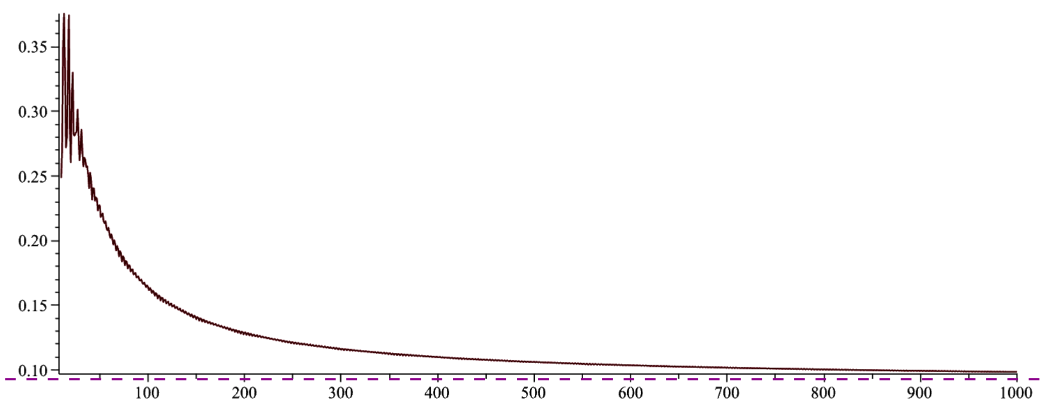
Plots of total variation distances



upper curve: $d_n := d_{\mathrm{TV}}(\pi_n, \pi'_n)$
lower curve: $d_n^{\mathrm{rej}} := d_{\mathrm{TV}}(\pi_n, \pi_n^{\mathrm{rej}})$



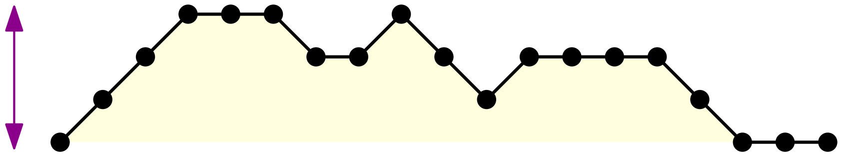
plot of $n^{1/2} \cdot d_n$
(predicted limit ≈ 0.42314)



plot of $n \cdot d_n^{\mathrm{rej}}$
(predicted limit ≈ 0.09074)

Plots of height distributions

height $= 3$

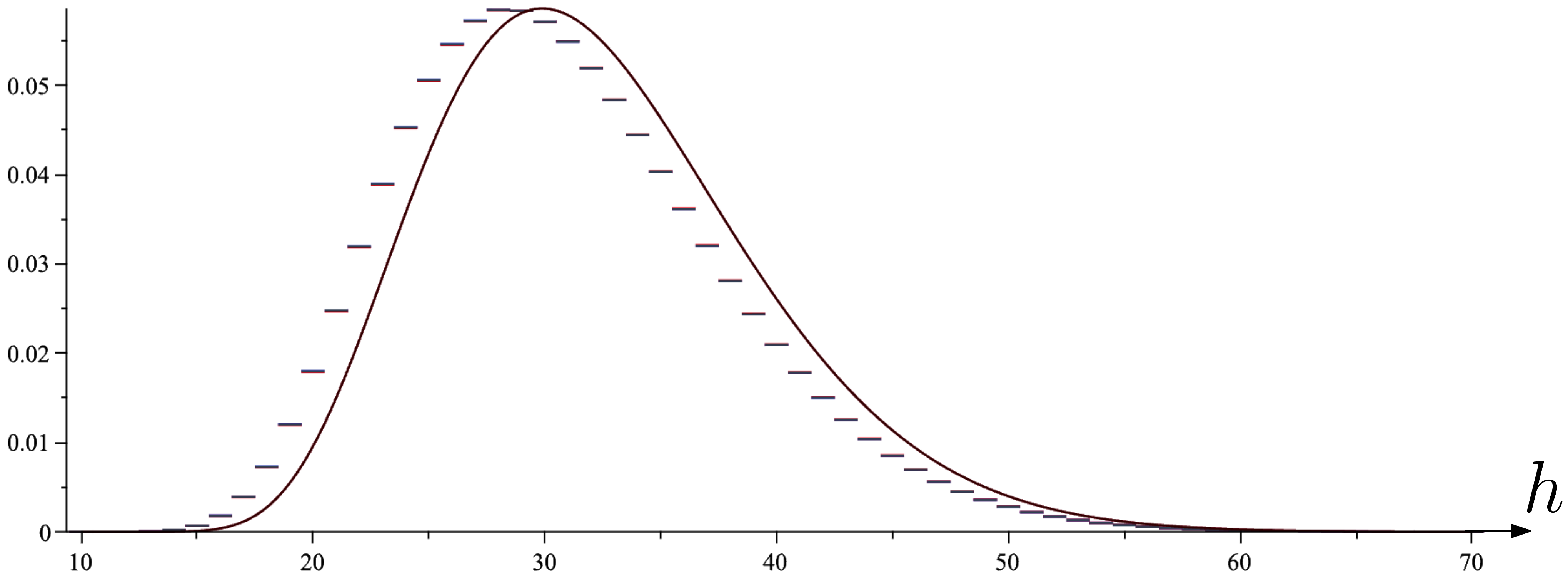


plots of distribution for the height h under

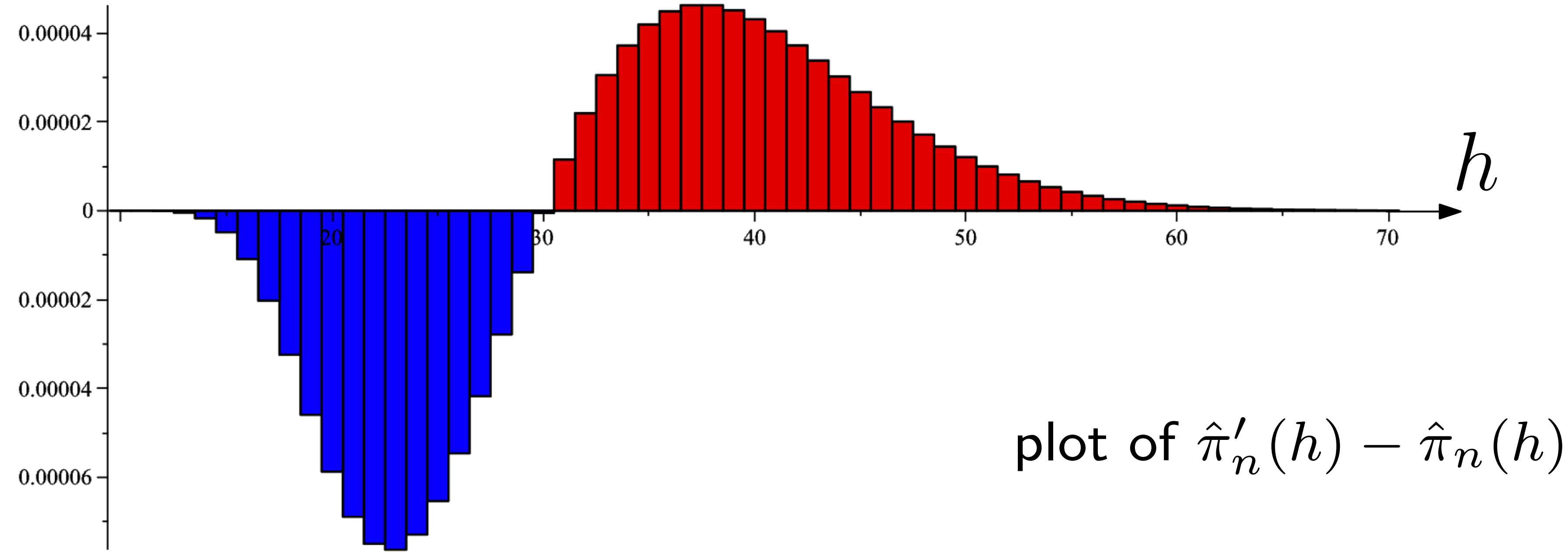
in size $\boxed{n = 1000}$

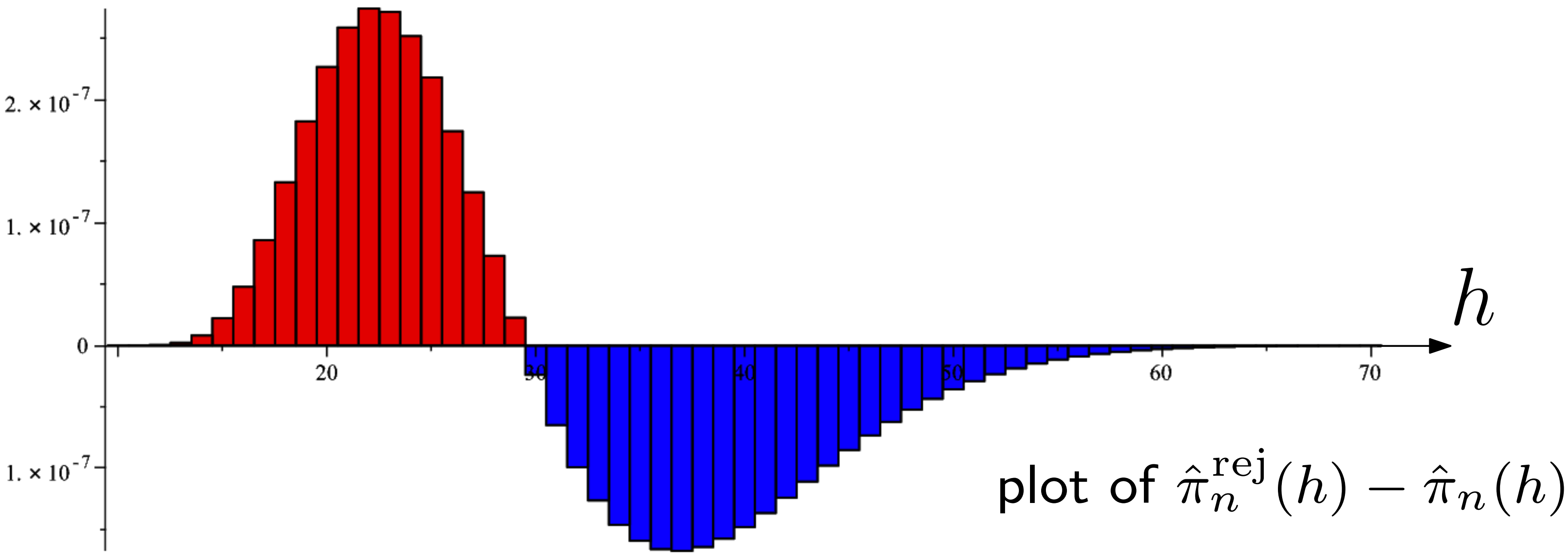
uniform: $\hat{\pi}_n(h)$
leap: $\hat{\pi}'_n(h)$
leap-reject: $\hat{\pi}^{\mathrm{rej}}_n(h)$

superimposed with the ($\sqrt{n}$-rescaled) limit curve



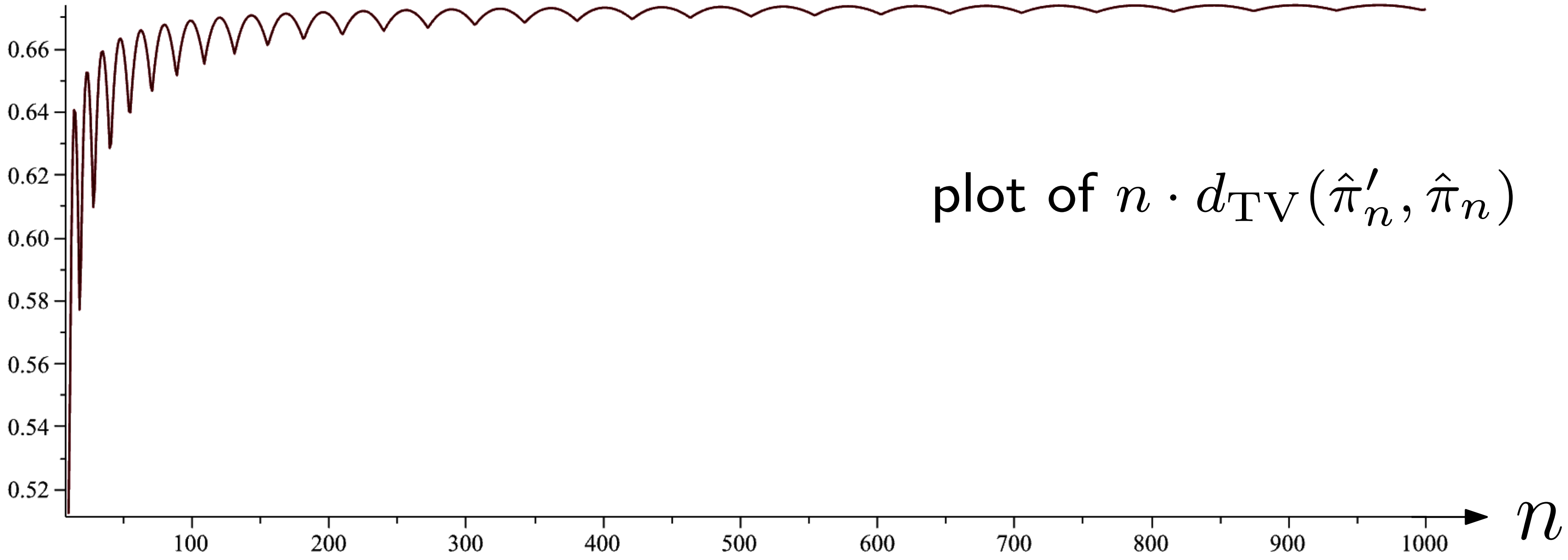
Plots of height distributions

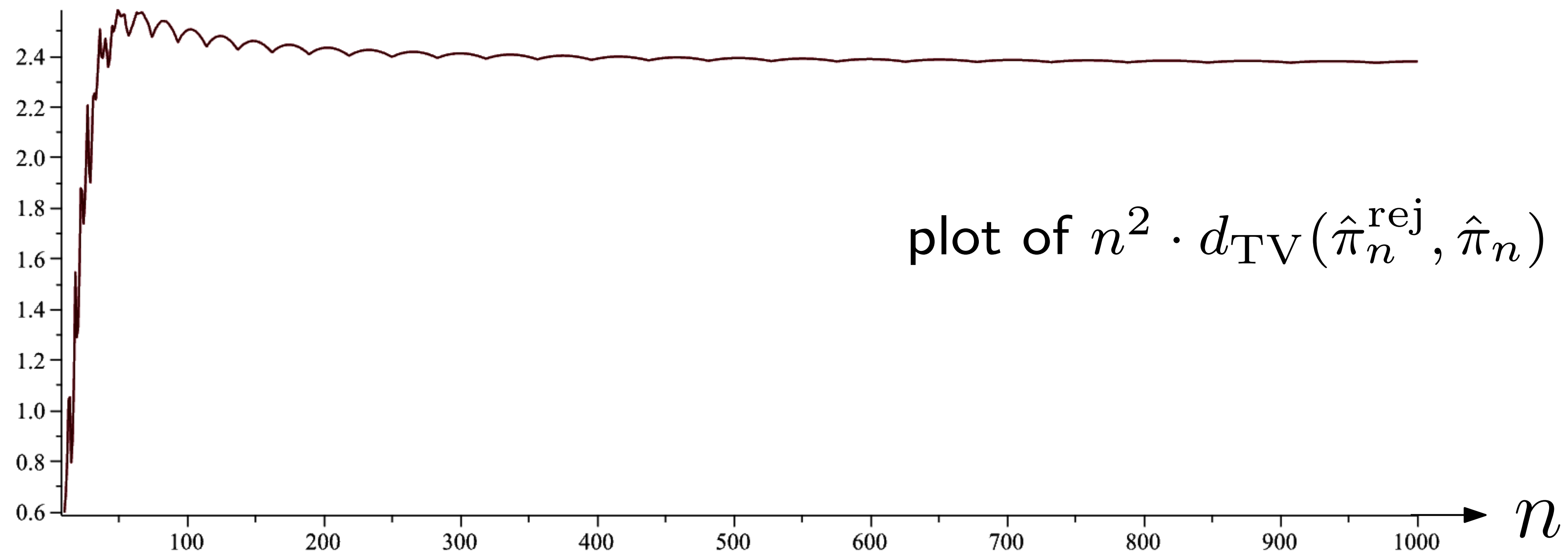




Plots of total-variation distance for height

Coupling argument $\Rightarrow d_{\mathrm{TV}}(\hat{\pi}'_n, \hat{\pi}_n) = O(1/n)$ and $d_{\mathrm{TV}}(\hat{\pi}^{\mathrm{rej}}_n, \hat{\pi}_n) = O(1/n^2)$





Burnside's lemma

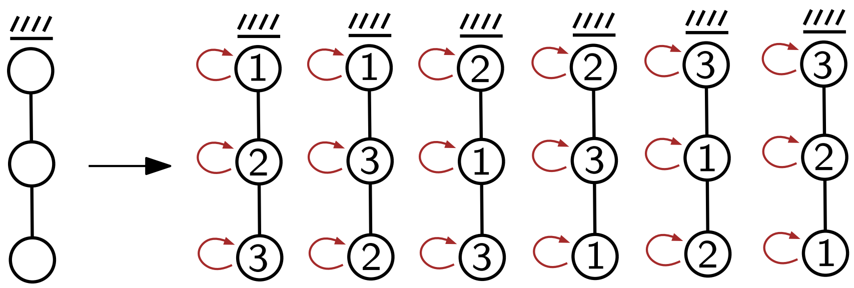
$\mathcal{L}_n$ labeled class (e.g. Cayley trees)

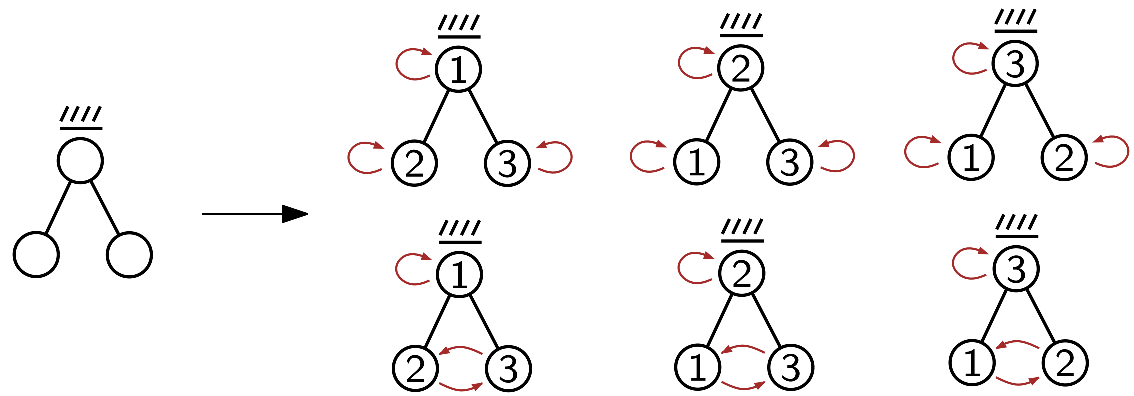
$$\mathcal{U}_n = \mathcal{L}_n / \mathfrak{S}_n \xrightarrow[\text{Burnside}]{\times\, n!} \mathcal{S}_n = \mathrm{Sym}(\mathcal{L})_n$$

unlabeled class (e.g. Pólya trees)

symmetry class

$$= \{(\sigma, \gamma) \in \mathfrak{S}_n \times \mathcal{C}_n, \quad \sigma \cdot \gamma = \gamma\}$$





Burnside's lemma

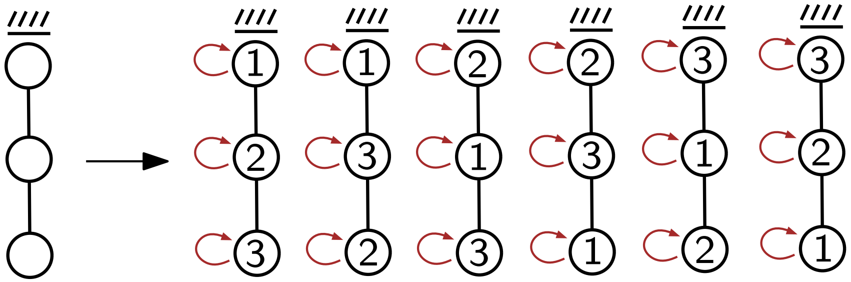
$\mathcal{L}_n$ labeled class (e.g. Cayley trees)

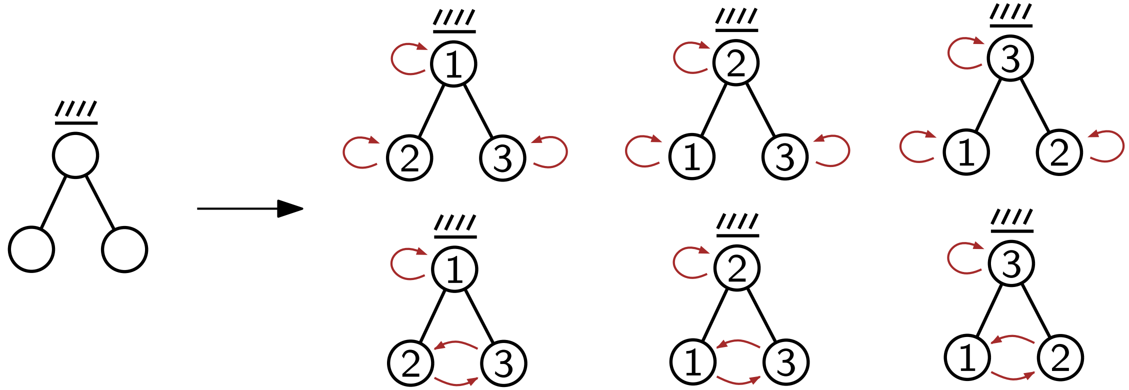
$\mathcal{U}_n = \mathcal{L}_n / \mathfrak{S}_n \xrightarrow[\text{Burnside}]{\times\, n!} \mathcal{S}_n = \mathrm{Sym}(\mathcal{L})_n$ symmetry class

unlabeled class

(e.g. Pólya trees)

$= \{(\sigma, \gamma) \in \mathfrak{S}_n \times \mathcal{C}_n, \quad \sigma \cdot \gamma = \gamma\}$





$\Rightarrow$ uniform (resp. asympt. uniform) sampling in $\mathcal{U}_n \Leftarrow$ in $\mathcal{S}_n$

Burnside's lemma

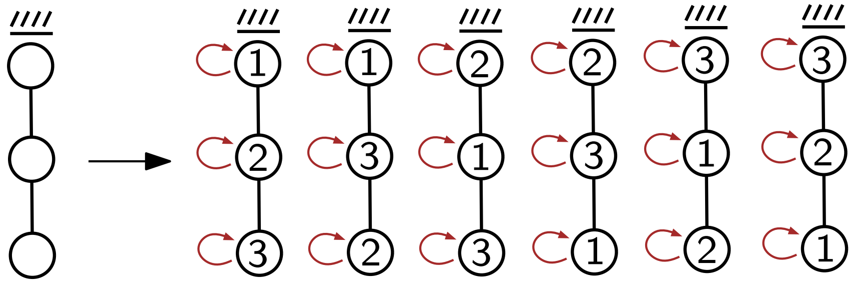
$\mathcal{L}_n$ labeled class (e.g. Cayley trees)

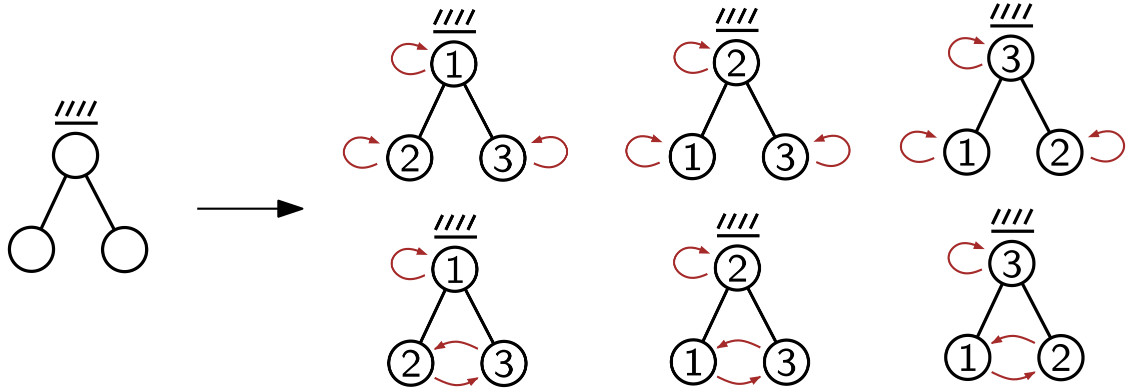
$\mathcal{U}_n = \mathcal{L}_n/\mathfrak{S}_n \xrightarrow[\text{Burnside}]{\times\, n!} \mathcal{S}_n = \mathrm{Sym}(\mathcal{L})_n$ symmetry class

unlabeled class

(e.g. Pólya trees)

$= \{(\sigma, \gamma) \in \mathfrak{S}_n \times \mathcal{C}_n, \quad \sigma \cdot \gamma = \gamma\}$





$\Rightarrow$ uniform (resp. asympt. uniform) sampling in $\mathcal{U}_n \Leftarrow$ in $\mathcal{S}_n$

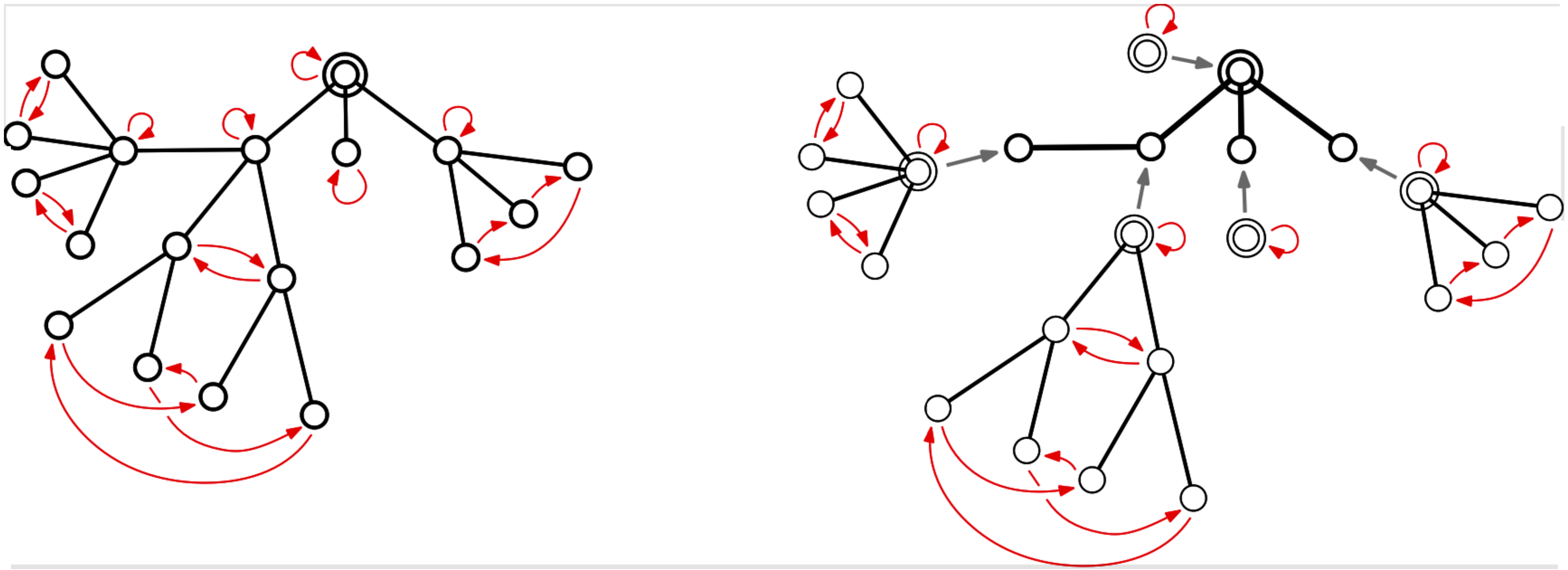
Rk: recent Markov chain generator for Pólya trees [Bartholdi, Diaconis'24]

Composition scheme for Pólya trees

For $\mathcal{A}$ = Cayley trees let $\mathcal{C} = \mathrm{Sym}(\mathcal{A})$

Composition scheme: $\boxed{\mathcal{C} = \mathcal{A} \circ \mathcal{B}}$ where $\mathcal{B} \subset \mathcal{C}$: only the root is fixed

(used to prove convergence of random Pólya trees to CRT [Panagiotou,Stufler'18])

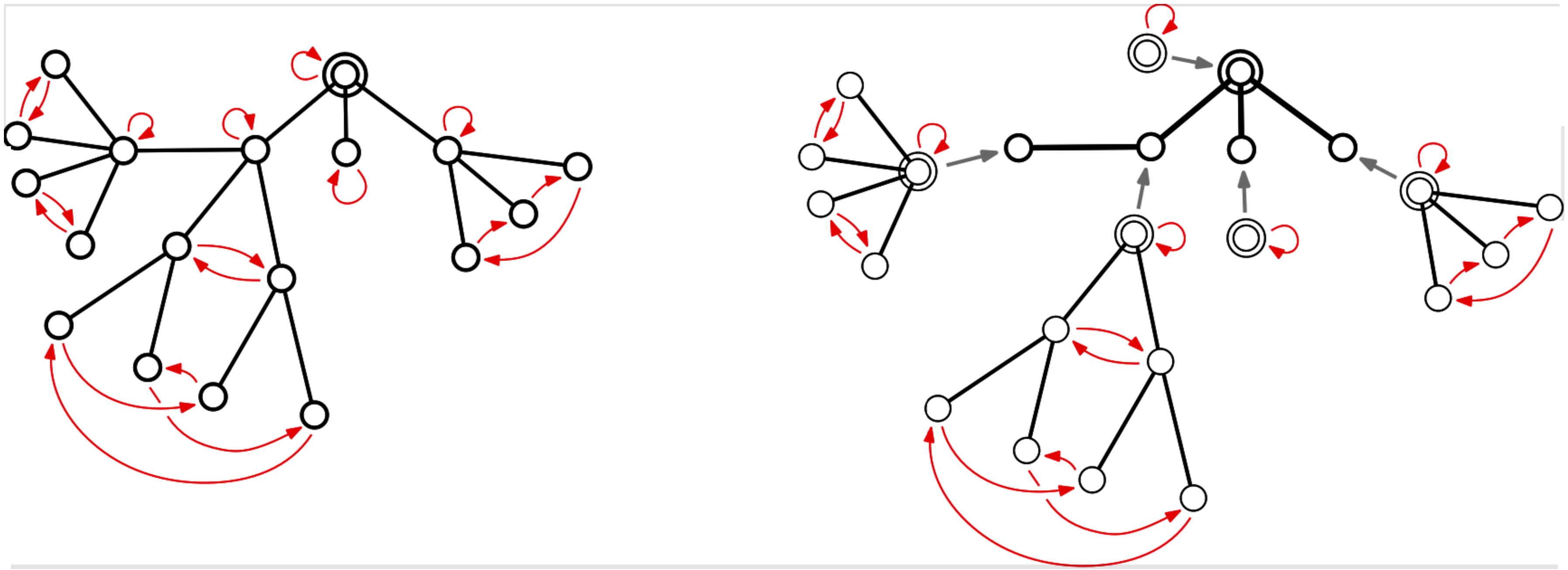


Composition scheme for Pólya trees

For $\mathcal{A} =$ Cayley trees   let $\mathcal{C} = \mathrm{Sym}(\mathcal{A})$

Composition scheme: $\boxed{\mathcal{C} = \mathcal{A} \circ \mathcal{B}}$ where $\mathcal{B} \subset \mathcal{C}$: only the root is fixed

(used to prove convergence of random Pólya trees to CRT [Panagiotou,Stufler'18])



generating functions

Cayley trees: $A(y) = y \exp(A(y))$

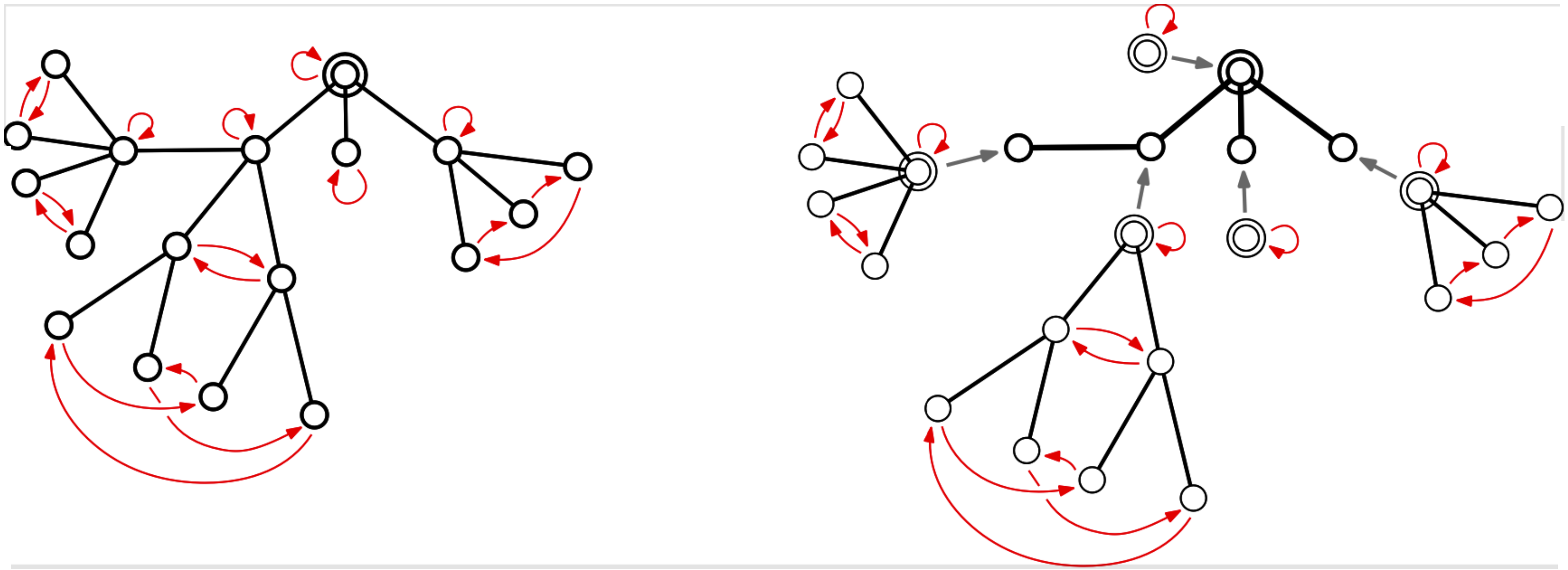
Pólya trees: $C(z) = z \exp\left(\sum_{i \geq 1} \frac{1}{i} C(z^i)\right)$

Composition scheme for Pólya trees

For $\mathcal{A}$ = Cayley trees let $\mathcal{C} = \mathrm{Sym}(\mathcal{A})$

Composition scheme: $\boxed{\mathcal{C} = \mathcal{A} \circ \mathcal{B}}$ where $\mathcal{B} \subset \mathcal{C}$: only the root is fixed

(used to prove convergence of random Pólya trees to CRT [Panagiotou,Stufler'18])



generating functions

Cayley trees: $A(y) = y\exp(A(y))$

Pólya trees: $C(z) = z\exp\left(\sum_{i\geq 1}\frac{1}{i}C(z^i)\right) = B(z)\exp(C(z))$

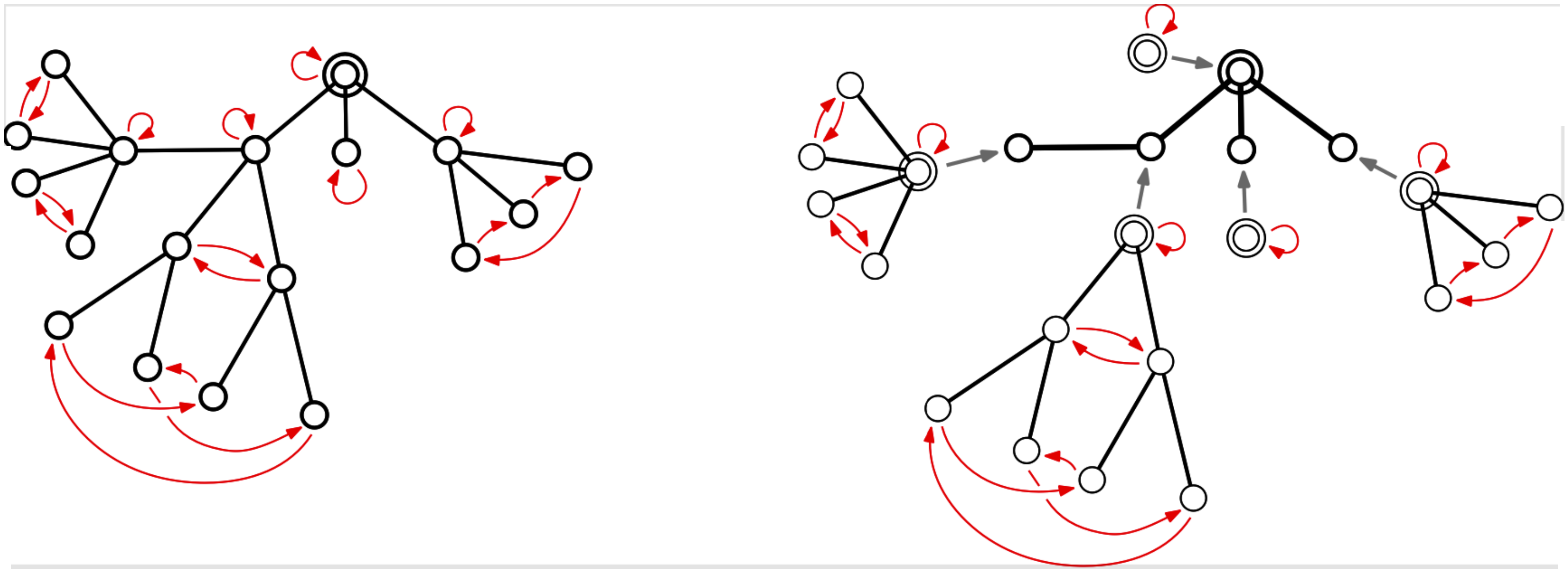
where $B(z) = z\exp\left(\sum_{i\geq 2}\frac{1}{i}C(z^i)\right)$

Composition scheme for Pólya trees

For $\mathcal{A}$ = Cayley trees let $\mathcal{C} = \mathrm{Sym}(\mathcal{A})$

Composition scheme: $\boxed{\mathcal{C} = \mathcal{A} \circ \mathcal{B}}$ where $\mathcal{B} \subset \mathcal{C}$: only the root is fixed

(used to prove convergence of random Pólya trees to CRT [Panagiotou,Stufler'18])



generating functions

Cayley trees: $A(y) = y \exp(A(y))$

Pólya trees: $C(z) = z \exp\left(\sum_{i\geq 1} \frac{1}{i} C(z^i)\right) = B(z) \exp(C(z))$

where $B(z) = z \exp\left(\sum_{i\geq 2} \frac{1}{i} C(z^i)\right)$

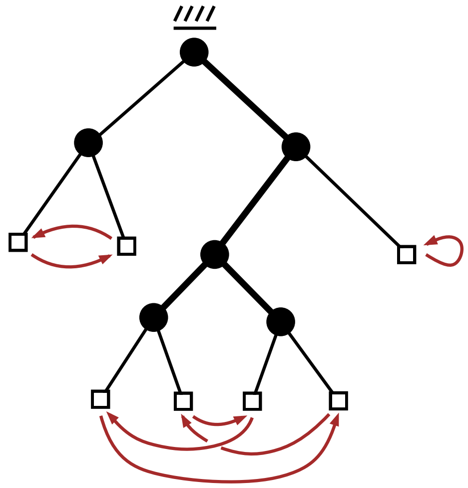
$\Rightarrow$ $\boxed{C(z) = A(y) \quad \text{where } y = B(z)}$

Composition scheme for non-plane binary trees

size = number of leaves

$\mathcal{A}$ = non-plane binary trees, labeled at leaves $\qquad \mathcal{C} = \mathrm{Sym}(\mathcal{A})$

Composition scheme: $\boxed{\mathcal{C} = \mathcal{A} \circ \mathcal{B}}$ where $\mathcal{B} \subset \mathcal{C}$: only the root is fixed

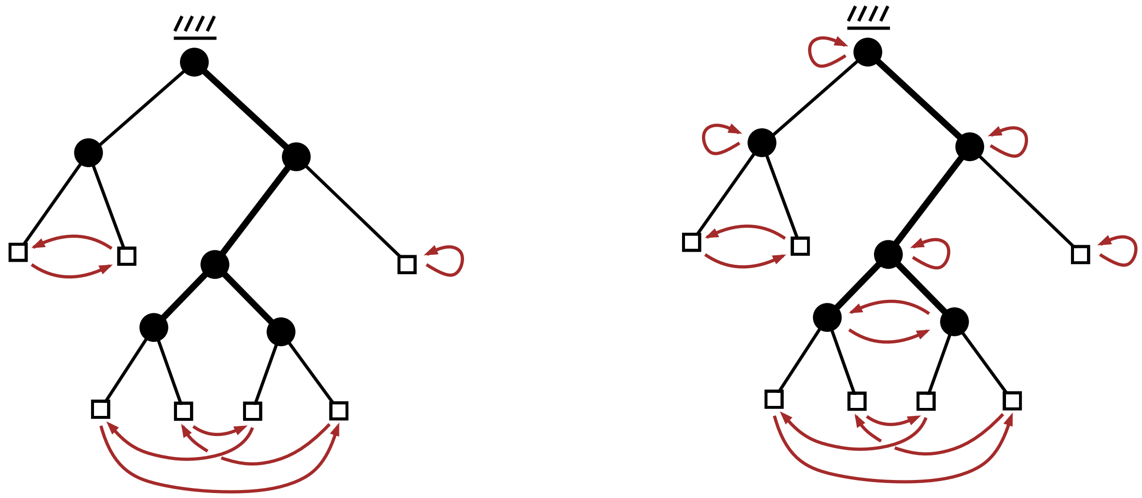


Composition scheme for non-plane binary trees

size = number of leaves

$\mathcal{A}$ = non-plane binary trees, labeled at leaves $\qquad \mathcal{C} = \mathrm{Sym}(\mathcal{A})$

Composition scheme: $\boxed{\mathcal{C} = \mathcal{A} \circ \mathcal{B}}$ where $\mathcal{B} \subset \mathcal{C}$: only the root is fixed

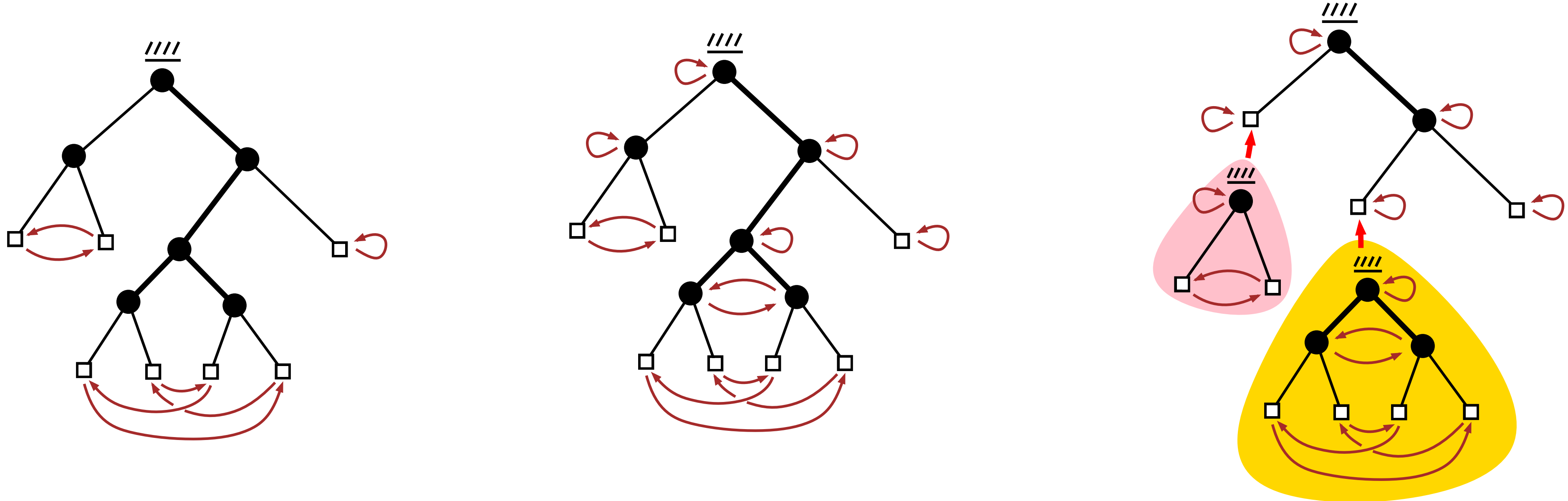


Composition scheme for non-plane binary trees

size = number of leaves

$\mathcal{A}$ = non-plane binary trees, labeled at leaves $\qquad \mathcal{C} = \mathrm{Sym}(\mathcal{A})$

Composition scheme: $\boxed{\mathcal{C} = \mathcal{A} \circ \mathcal{B}}$ where $\mathcal{B} \subset \mathcal{C}$: only the root is fixed

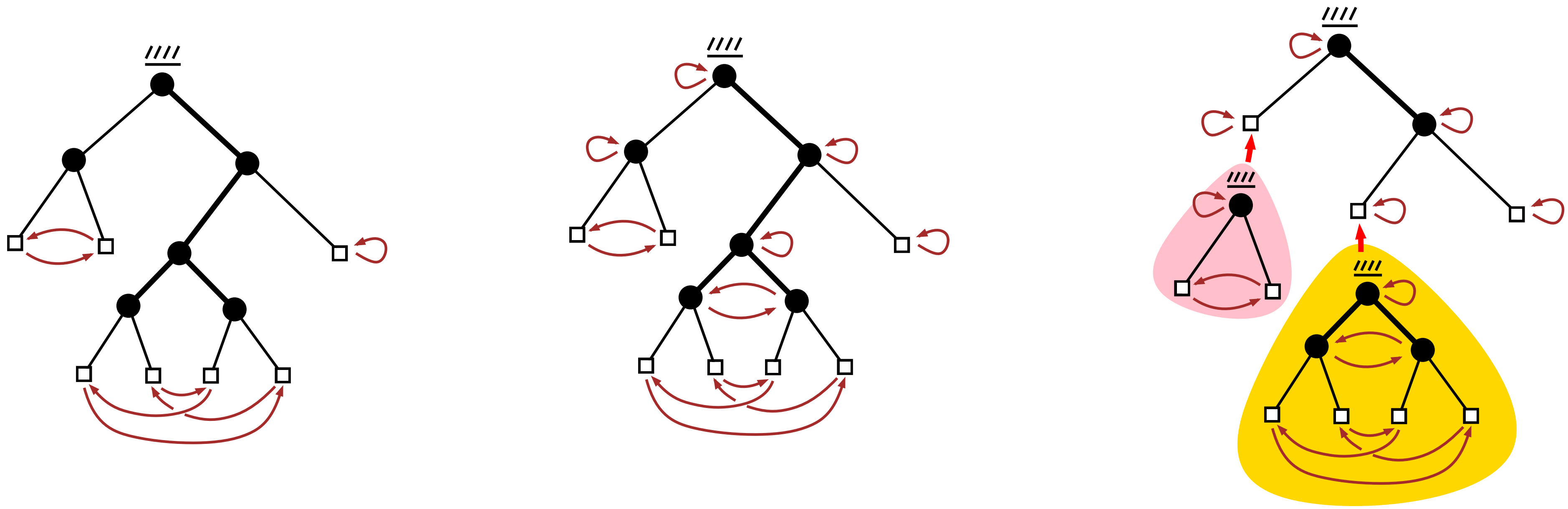


Composition scheme for non-plane binary trees

size = number of leaves

$\mathcal{A}$ = non-plane binary trees, labeled at leaves $\qquad\qquad \mathcal{C} = \mathrm{Sym}(\mathcal{A})$

Composition scheme: $\boxed{\mathcal{C} = \mathcal{A} \circ \mathcal{B}}$ where $\mathcal{B} \subset \mathcal{C}$: only the root is fixed



generating functions

Labeled trees: $A(y) = y + \frac{1}{2}A(y)^2$

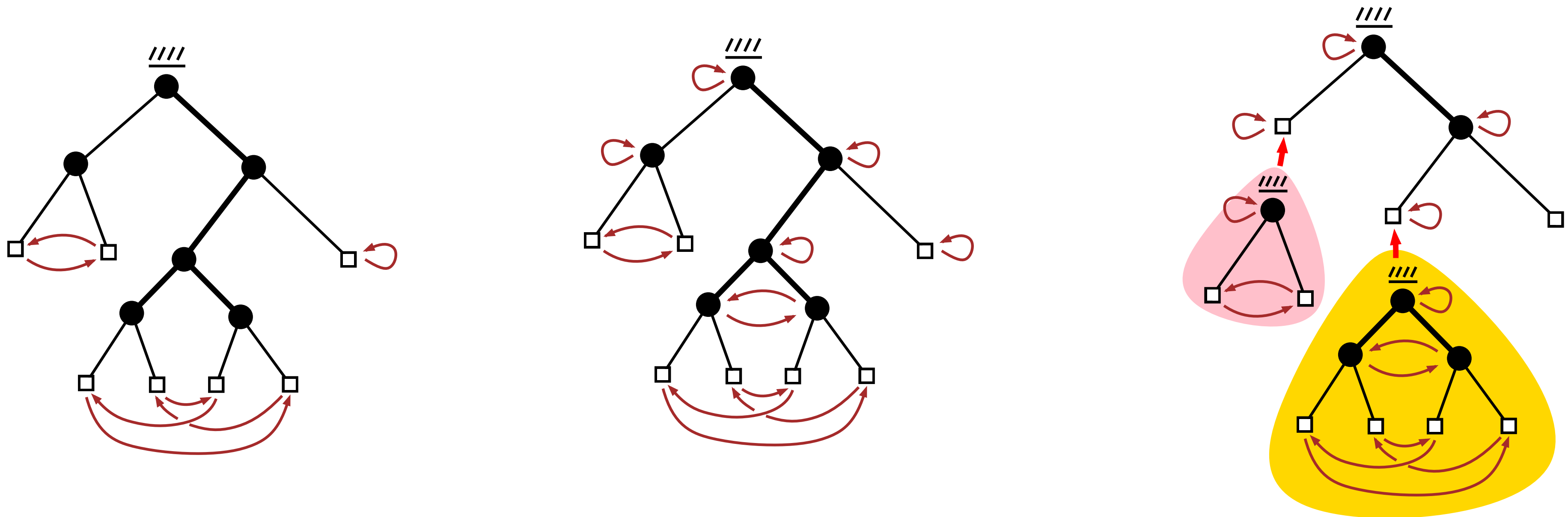
Unlabeled trees: $C(z) = z + \frac{1}{2}C(z)^2 + \frac{1}{2}C(z^2)$

Composition scheme for non-plane binary trees

size = number of leaves

$\mathcal{A}$ = non-plane binary trees, labeled at leaves $\qquad \mathcal{C} = \mathrm{Sym}(\mathcal{A})$

Composition scheme: $\boxed{\mathcal{C} = \mathcal{A} \circ \mathcal{B}}$ where $\mathcal{B} \subset \mathcal{C}$: only the root is fixed



generating functions

Labeled trees: $A(y) = y + \frac{1}{2}A(y)^2$

Unlabeled trees: $C(z) = z + \frac{1}{2}C(z)^2 + \frac{1}{2}C(z^2) \; = B(z) + \frac{1}{2}C(z)^2$

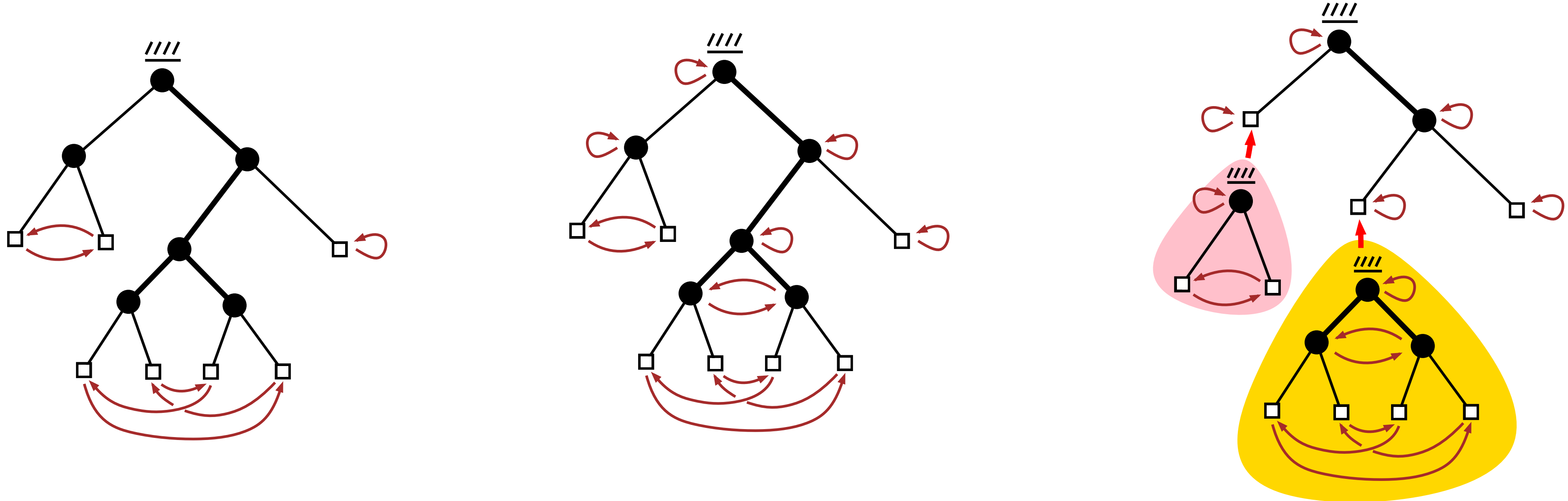
where $B(z) = z + \frac{1}{2}C(z^2)$

Composition scheme for non-plane binary trees

size = number of leaves

$\mathcal{A}$ = non-plane binary trees, labeled at leaves $\qquad \mathcal{C} = \mathrm{Sym}(\mathcal{A})$

Composition scheme: $\boxed{\mathcal{C} = \mathcal{A} \circ \mathcal{B}}$ where $\mathcal{B} \subset \mathcal{C}$: only the root is fixed



generating functions

Labeled trees: $A(y) = y + \frac{1}{2}A(y)^2$

Unlabeled trees: $C(z) = z + \frac{1}{2}C(z)^2 + \frac{1}{2}C(z^2) \;\; = B(z) + \frac{1}{2}C(z)^2$

where $B(z) = z + \frac{1}{2}C(z^2)$

Rk: extends to $\mathcal{U} = \mathcal{Z} + \mathrm{Cyc}_k(\mathcal{U})$ (k-ary mobiles)

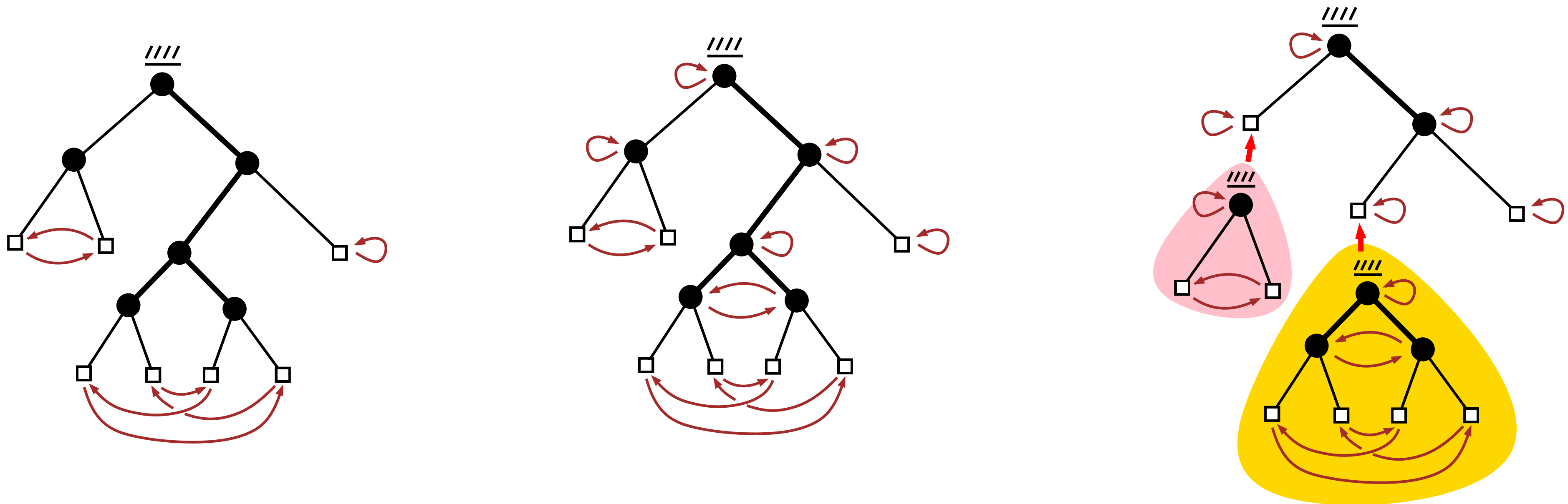
but not to $\mathcal{U} = \mathcal{Z} + \mathrm{MSet}_k(\mathcal{U})$ (k-ary trees)

Composition scheme for non-plane binary trees

size = number of leaves

$\mathcal{A}$ = non-plane binary trees, labeled at leaves $\qquad \mathcal{C} = \mathrm{Sym}(\mathcal{A})$

Composition scheme: $\boxed{\mathcal{C} = \mathcal{A} \circ \mathcal{B}}$ where $\mathcal{B} \subset \mathcal{C}$: only the root is fixed



generating functions

Labeled trees: $A(y) = y + \frac{1}{2}A(y)^2$

Unlabeled trees: $C(z) = z + \frac{1}{2}C(z)^2 + \frac{1}{2}C(z^2) \; = B(z) + \frac{1}{2}C(z)^2$

where $B(z) = z + \frac{1}{2}C(z^2)$

Rk: extends to $\mathcal{U} = \mathcal{Z} + \mathrm{Cyc}_k(\mathcal{U})$ (k-ary mobiles)

but not to $\mathcal{U} = \mathcal{Z} + \mathrm{MSet}_k(\mathcal{U})$ (k-ary trees) $C(z) = z + \frac{1}{6}C(z)^3 + \frac{1}{2}C(z^2)C(z) + \frac{1}{3}C(z^3)$

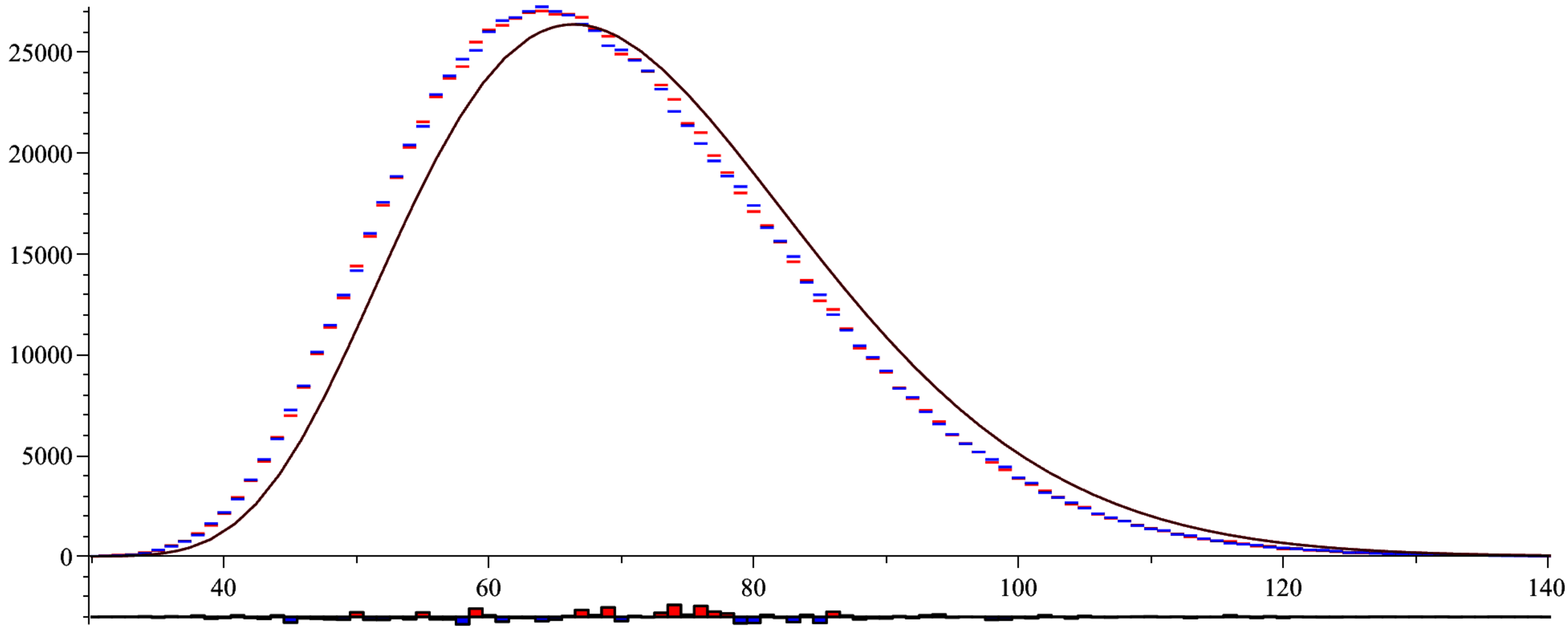
Histograms for the height

sample-size: $N = 10^6$
tree-size: $n = 10^3$

histograms for the height of random **Pólya trees**

in blue: under the uniform distribution (recursive method)
in red: under the leap distribution

superimposed with the ($\sqrt{n}$-rescaled) limit curve



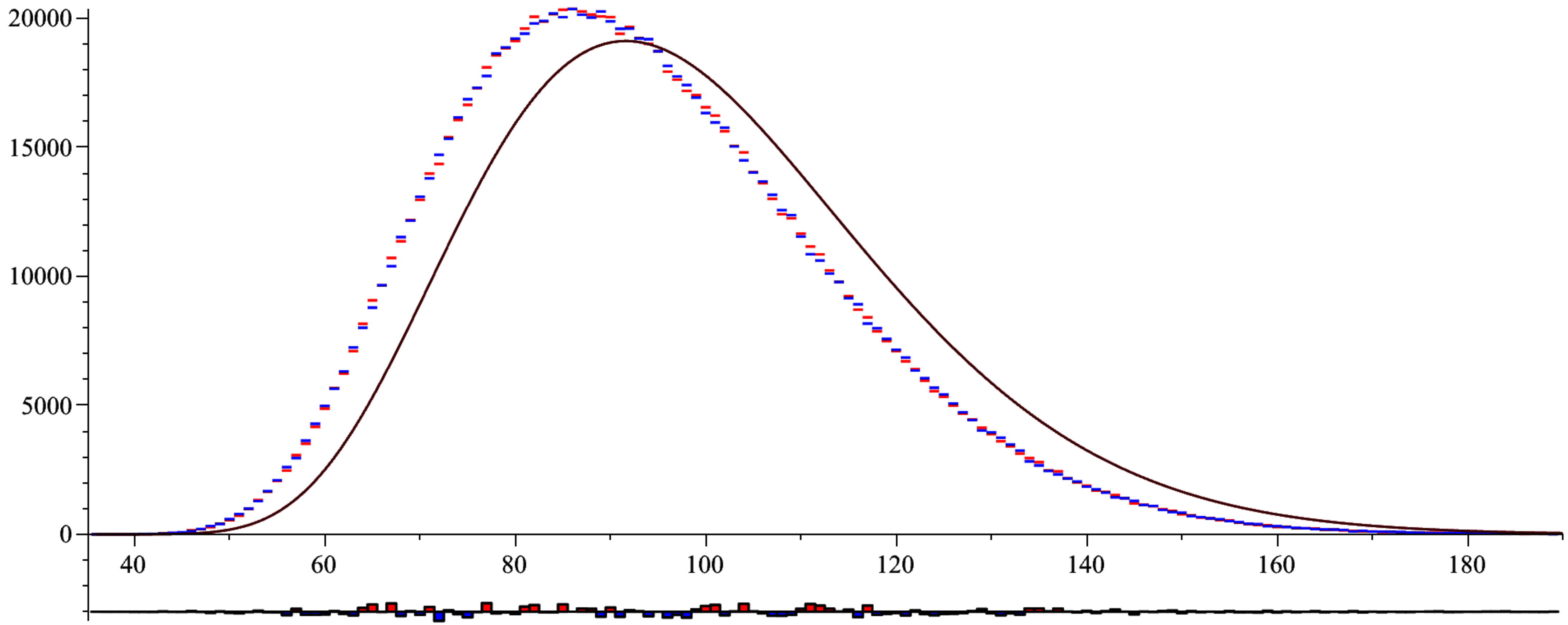
Histograms for the height

sample-size: $N = 10^6$
tree-size: $n = 10^3$

histograms for the height of non-plane **binary trees**

in blue: under the uniform distribution (recursive method)
in red: under the leap distribution

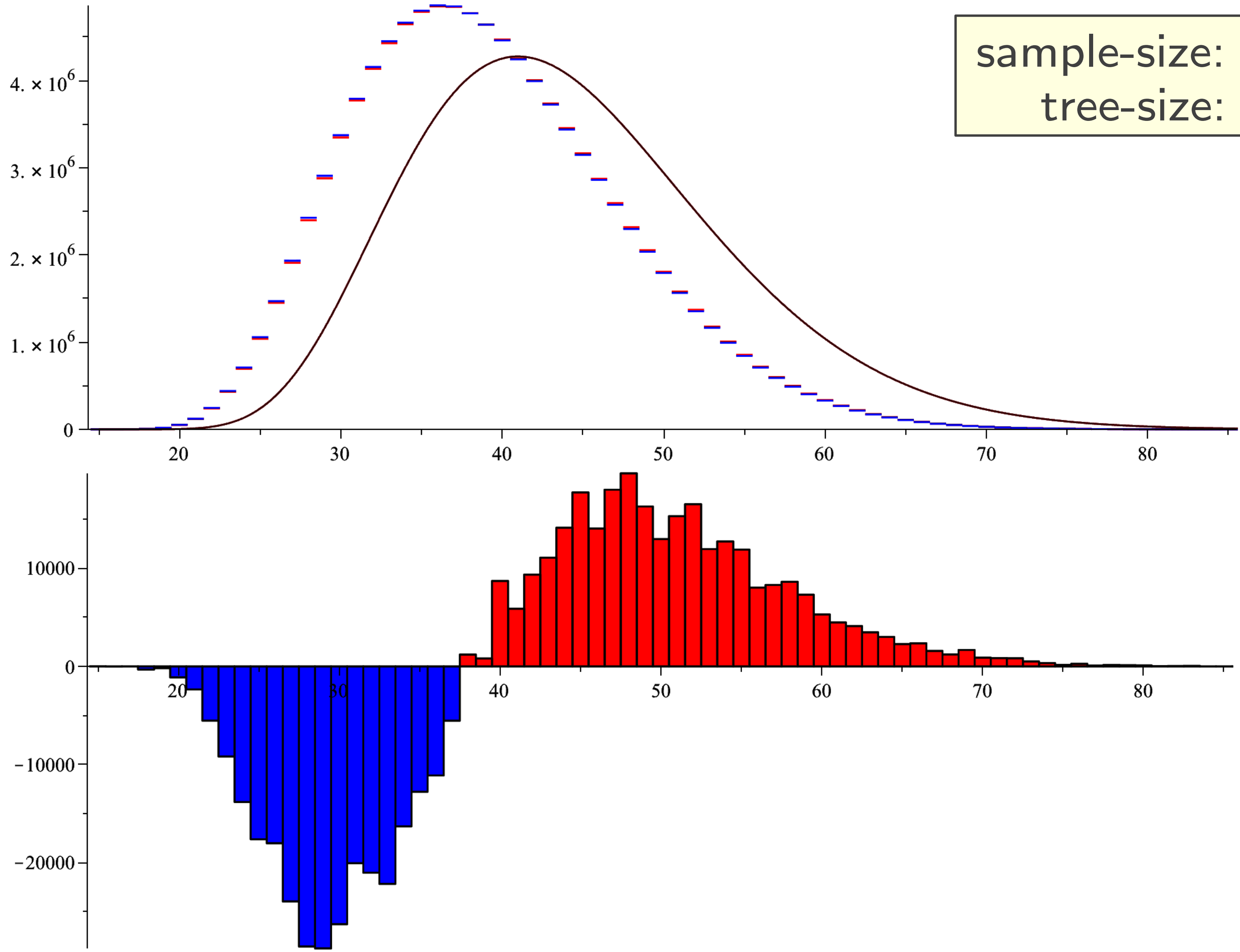
superimposed with the ($\sqrt{n}$-rescaled) limit curve



Histograms for the height

non-plane **binary trees**

sample-size: $N = 10^8$
tree-size: $n = 200$



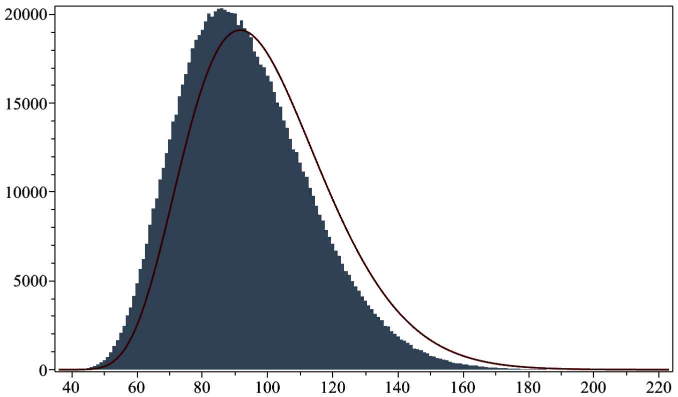
Histograms for the height

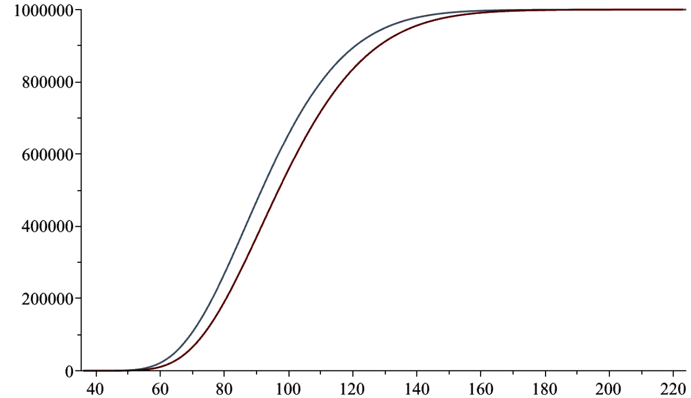
non-plane **binary trees**

leap distribution, sample-size$= 10^6$ in each case

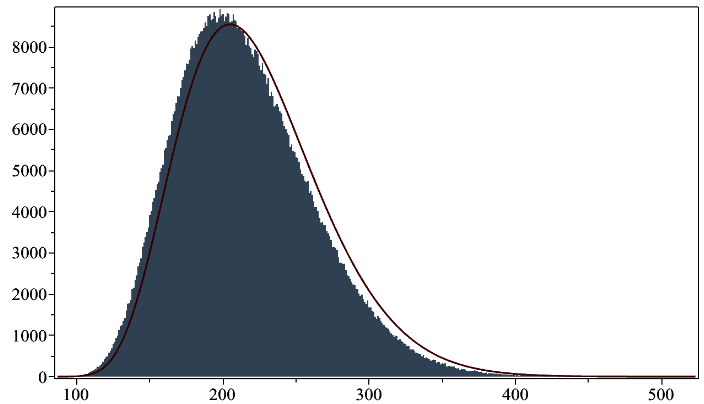
cumulative histograms

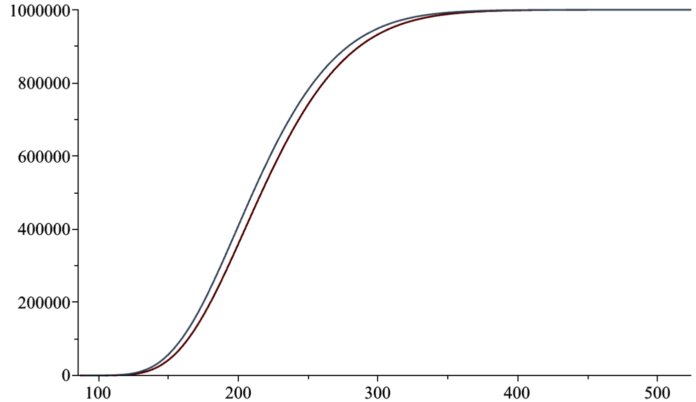
$n = 1000$



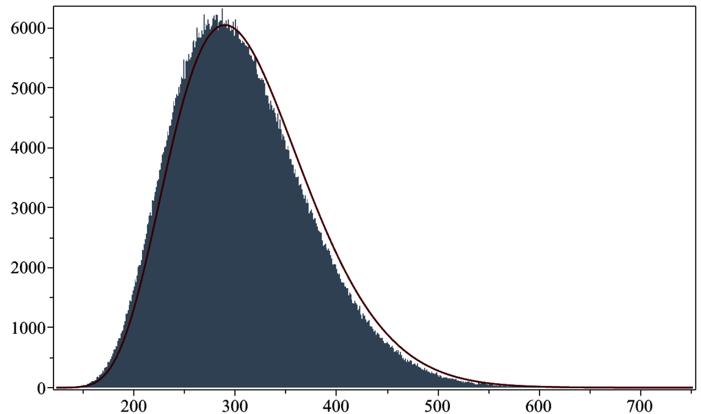


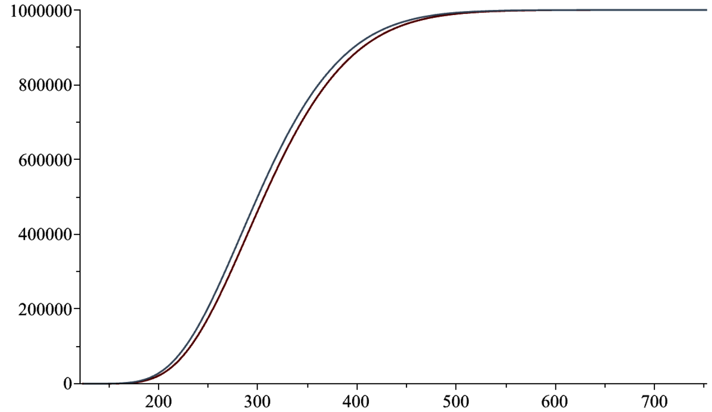
$n = 5000$





$n = 10\,000$





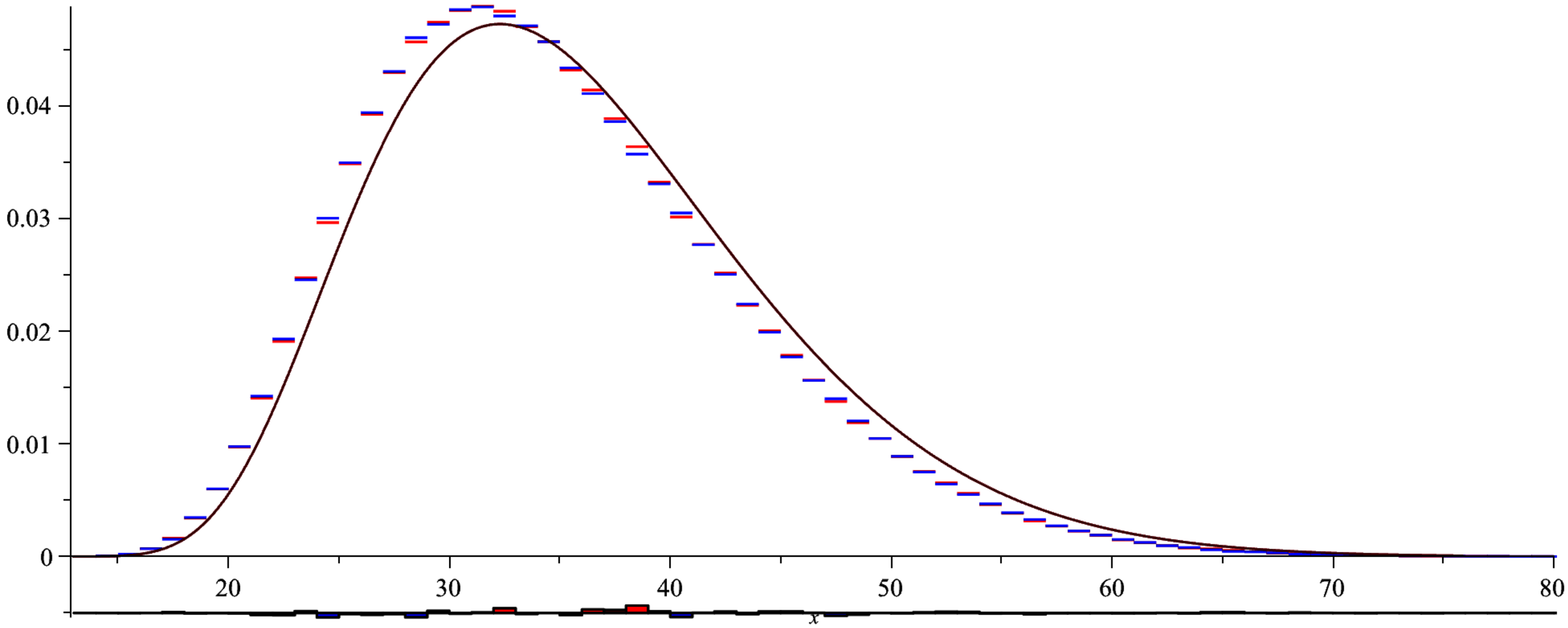
Histograms for the path-length

histograms for the path-length of random **Pólya trees**

sample-size: $N = 10^6$
tree-size: $n = 10^3$

in blue: under the uniform distribution (recursive method)
in red: under the leap distribution

superimposed with the ($\sqrt{n}$-rescaled) limit curve



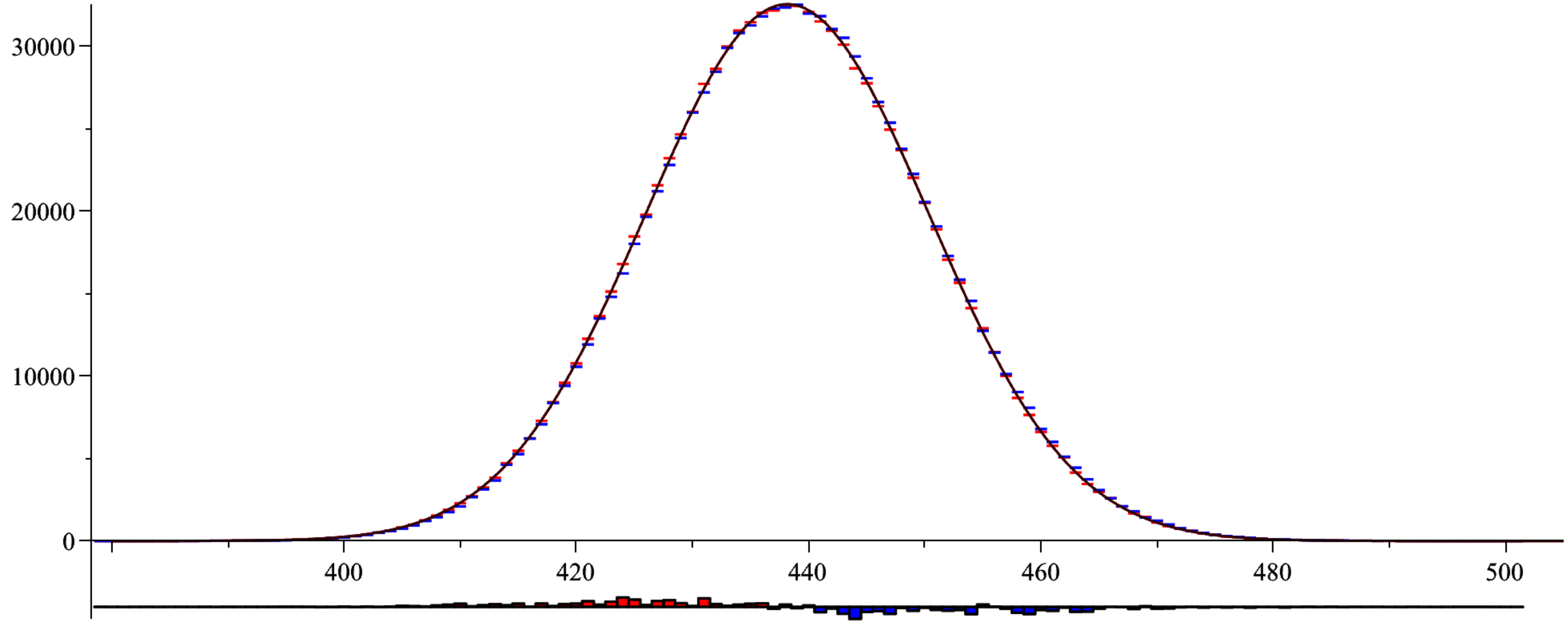
Histograms for local parameters

sample-size: $N = 10^6$
tree-size: $n = 10^3$

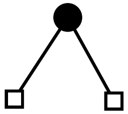
histograms for # leaves in random **Pólya trees**

in blue: under the uniform distribution (recursive method)
in red: under the leap distribution

superimposed with the ($\sqrt{n}$-rescaled) Gaussian limit curve



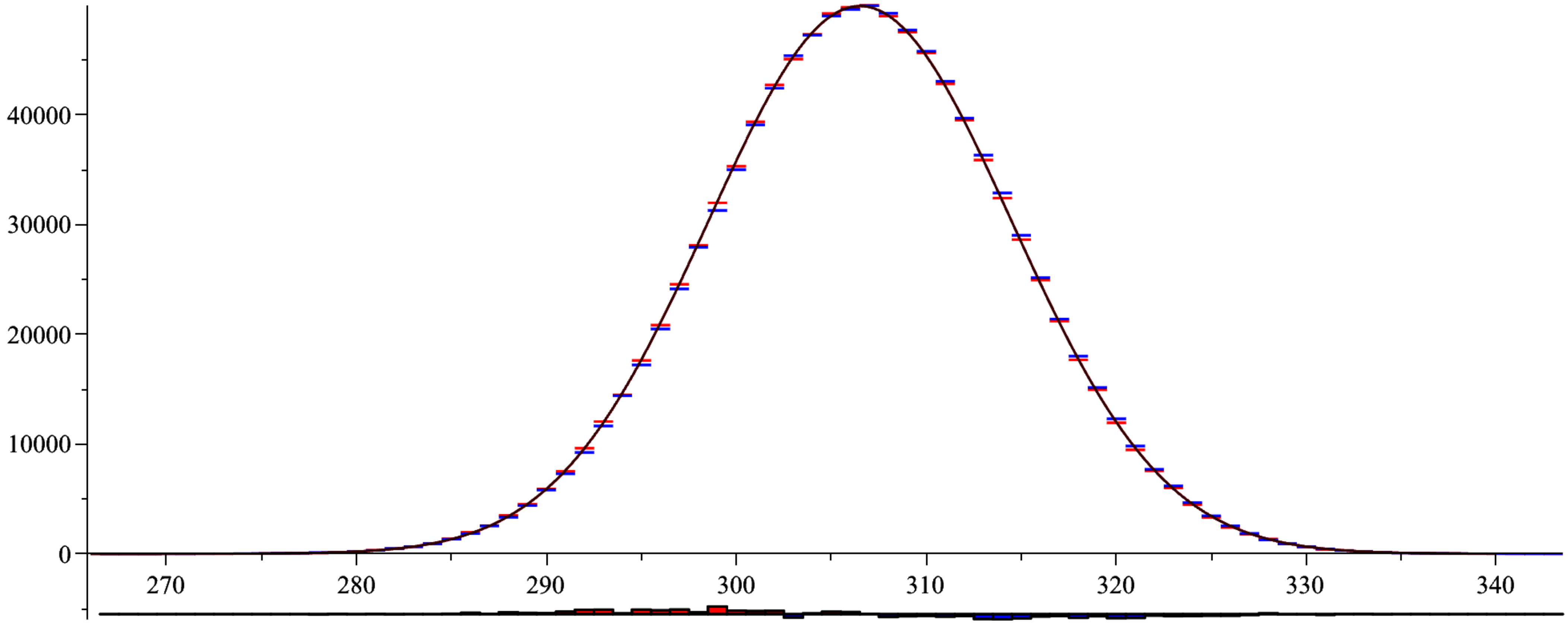
Histograms for local parameters

\# cherries  in random non-plane **binary trees**

sample-size: $N = 10^6$
tree-size: $n = 10^3$

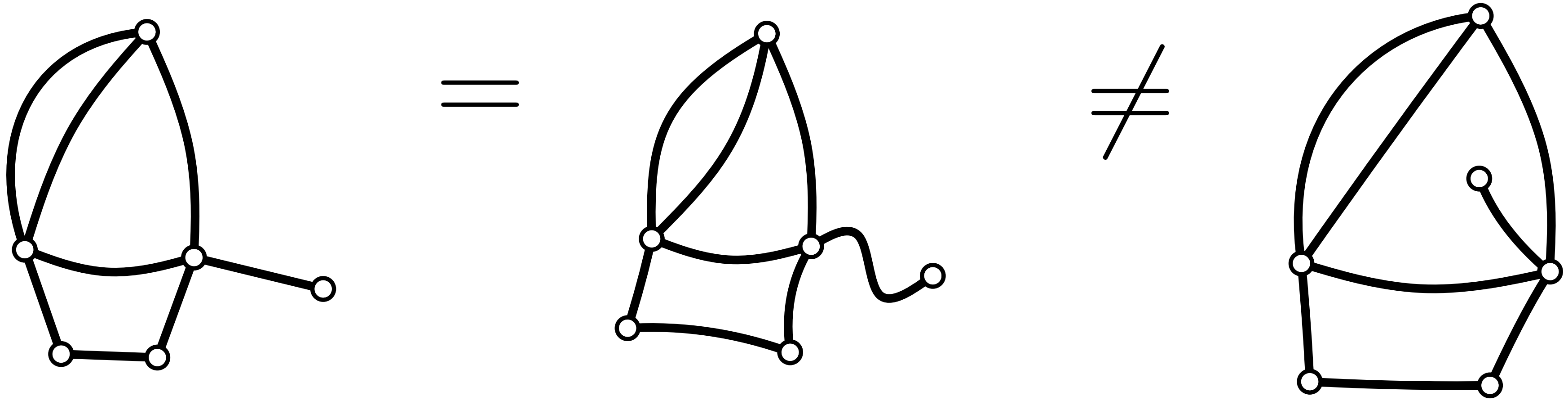
in blue: under the uniform distribution (recursive method)
in red: under the leap distribution

superimposed with the ($\sqrt{n}$-rescaled) Gaussian limit curve



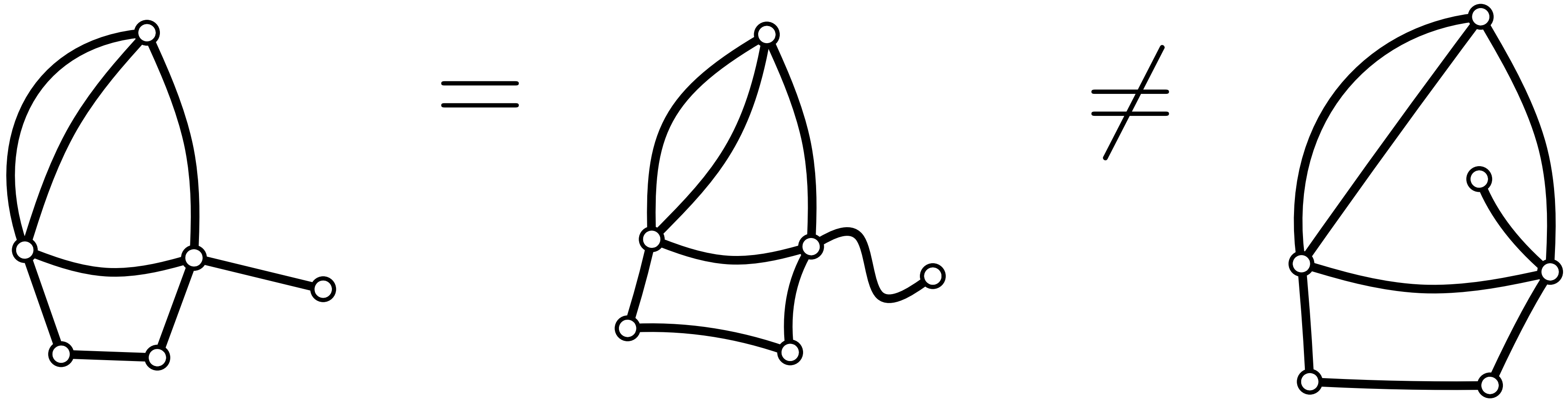
Planar maps

Planar map = connected multigraph embedded in the plane up to isotopy

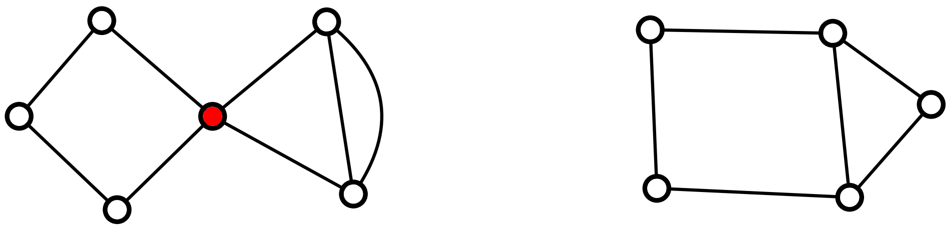


Planar maps

Planar map = connected multigraph embedded in the plane up to isotopy

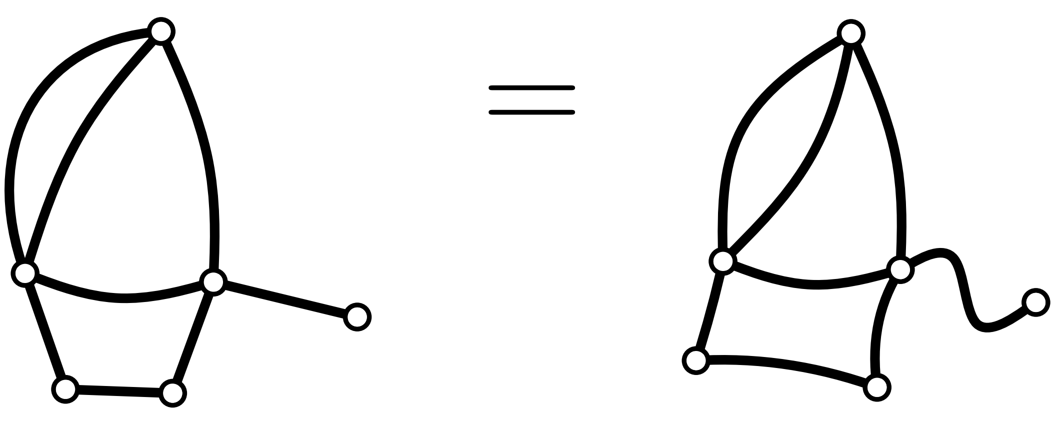


2-connected: no separating vertex

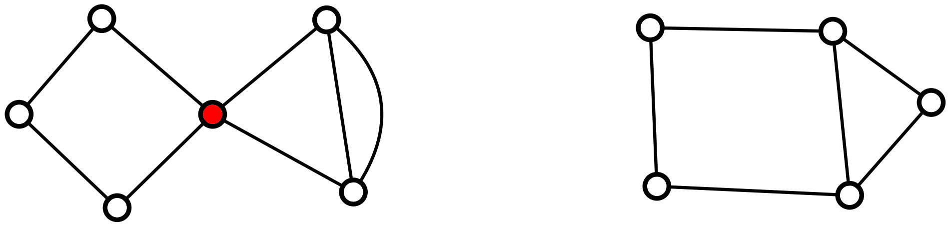


Planar maps

Planar map = connected multigraph embedded in the plane up to isotopy

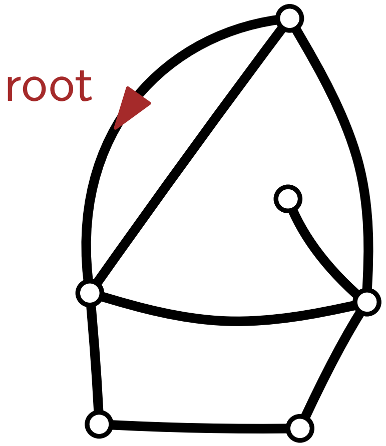


2-connected: no separating vertex



not 2-connected 2-connected

rooted planar map: has a marked
oriented edge with outer face on its right



Composition schemes for planar maps

[Tutte'63, Banderier-Flajolet-Schaeffer-Soria'02]

Several examples of composition schemes relating planar map families

$$\mathcal{C} = \mathcal{A} \circ \mathcal{B}$$

lower connectivity

higher connectivity

Composition schemes for planar maps

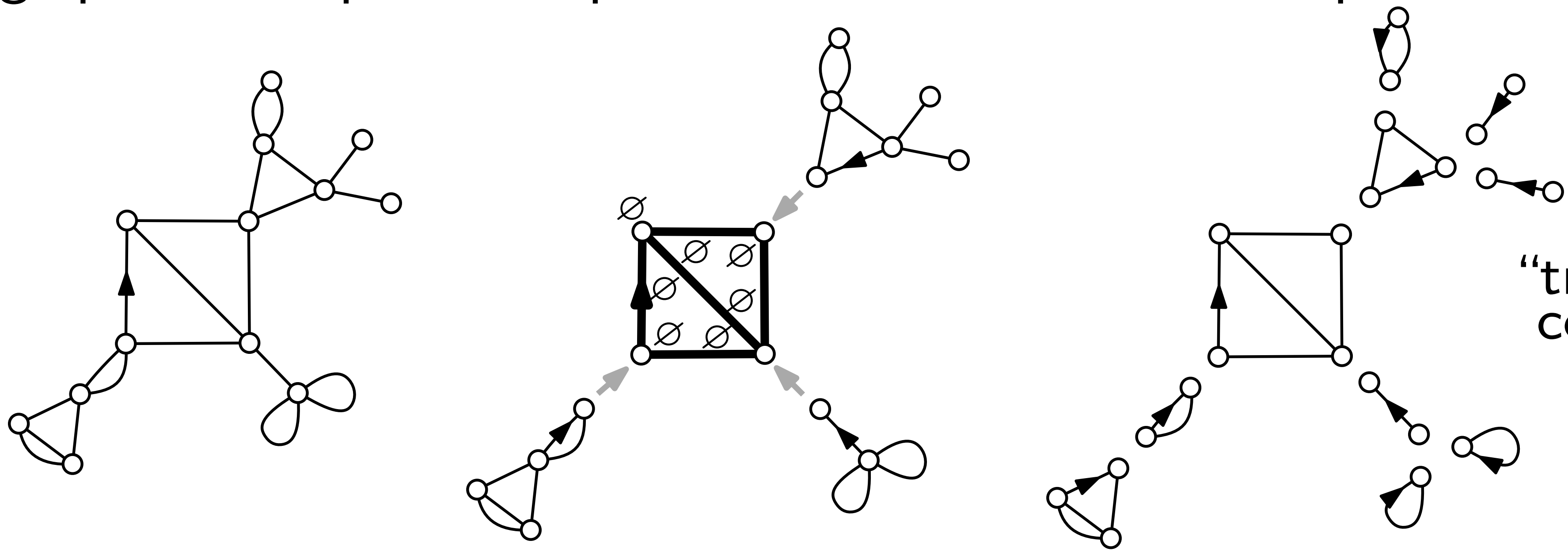
[Tutte'63, Banderier-Flajolet-Schaeffer-Soria'02]

Several examples of composition schemes relating planar map families

$$\mathcal{C} = \mathcal{A} \circ \mathcal{B}$$

lower connectivity   higher connectivity

e.g. planar maps decomposed into 2-connected components



"tree" of 2-c components

$\mathcal{C}$ = planar maps $\mathcal{A}$ = 2-connected planar maps $\mathcal{B} = \mathcal{Z} \times (1 + \mathcal{C})^2$

Composition schemes for planar maps

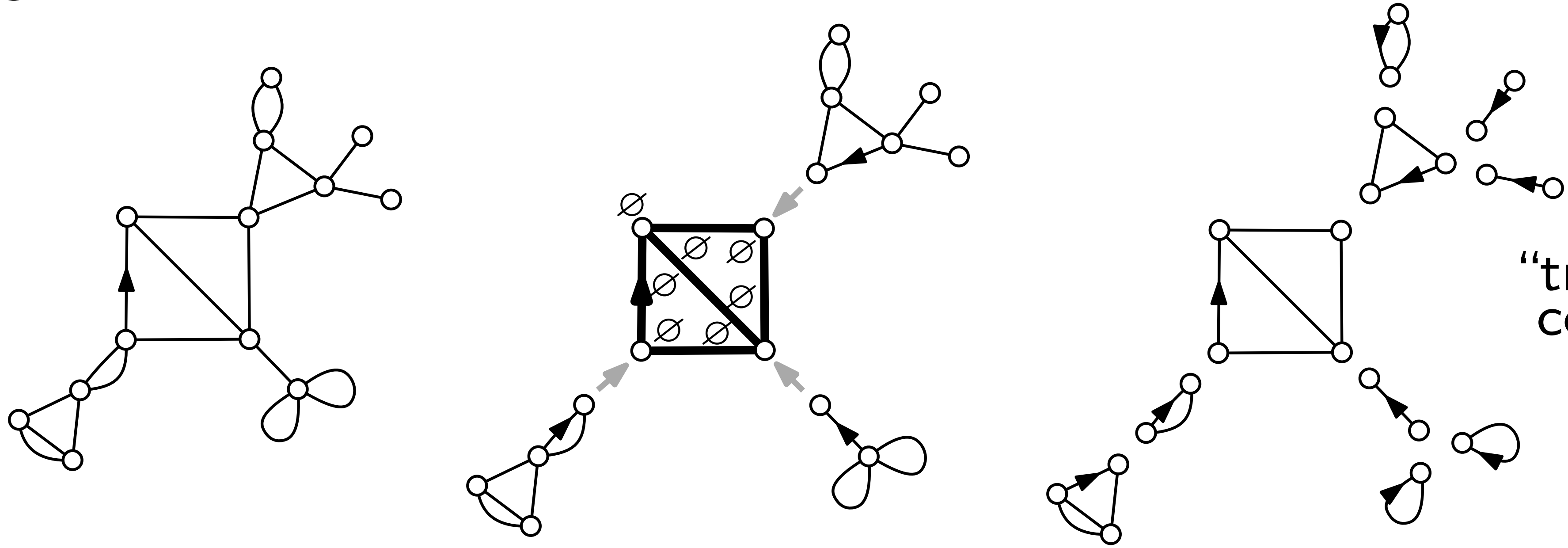
[Tutte'63, Banderier-Flajolet-Schaeffer-Soria'02]

Several examples of composition schemes relating planar map families

$$\mathcal{C} = \mathcal{A} \circ \mathcal{B}$$

lower connectivity   higher connectivity

e.g. planar maps decomposed into 2-connected components



$\mathcal{C}$ = planar maps $\mathcal{A}$ = 2-connected planar maps $\mathcal{B} = \mathcal{Z} \times (1+\mathcal{C})^2$

these schemes are **critical** $\boxed{B(\rho_B) = \rho_A}$ singular type $(1-z/\rho)^{3/2}$

Composition schemes for planar maps

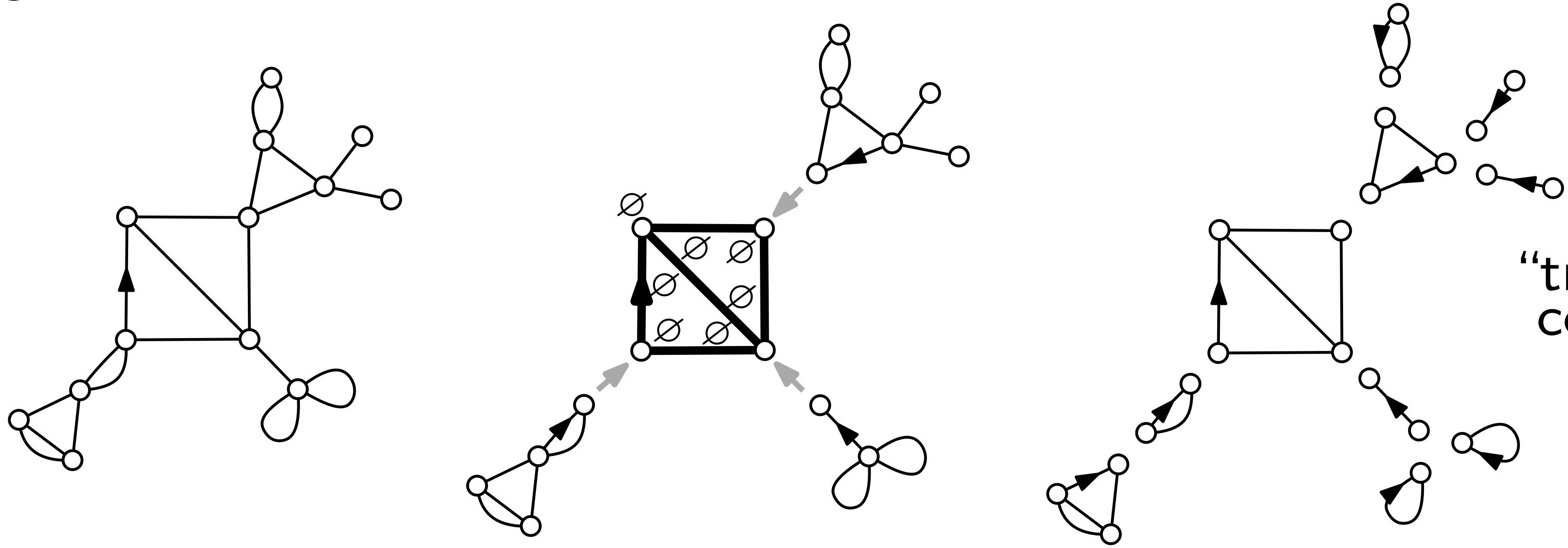
[Tutte'63, Banderier-Flajolet-Schaeffer-Soria'02]

Several examples of composition schemes relating planar map families

$$\mathcal{C} = \mathcal{A} \circ \mathcal{B}$$

lower connectivity higher connectivity

e.g. planar maps decomposed into 2-connected components



"tree" of 2-c components

$\mathcal{C}$ = planar maps $\mathcal{A}$ = 2-connected planar maps $\mathcal{B} = \mathcal{Z} \times (1+\mathcal{C})^2$

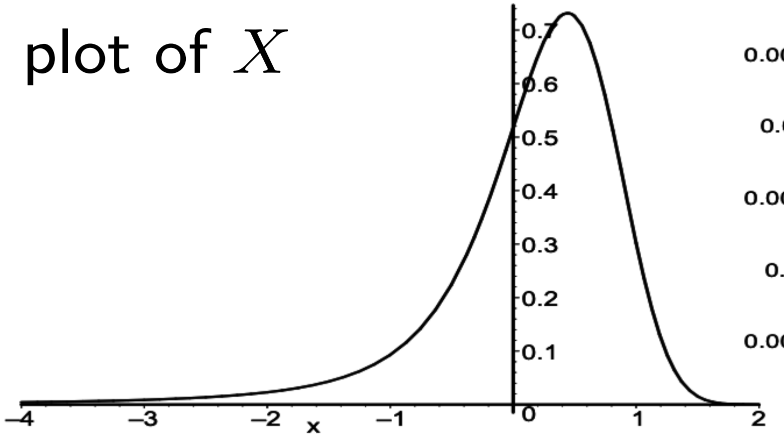
these schemes are **critical** $\boxed{B(\rho_B) = \rho_A}$ singular type $(1 - z/\rho)^{3/2}$

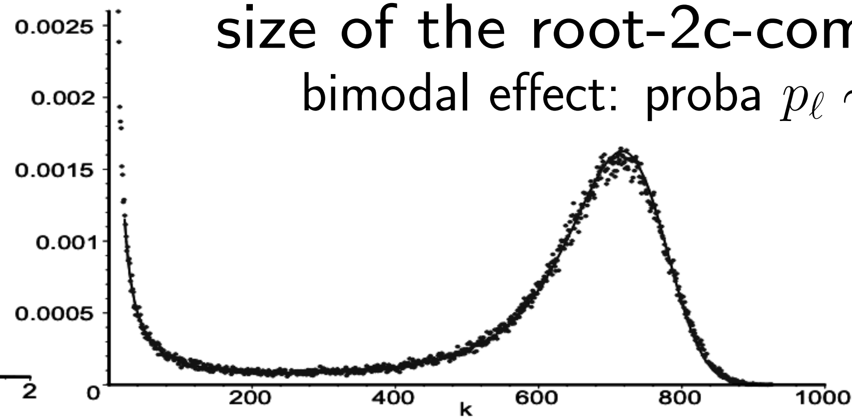
Rk: $\mu := \mathbb{E}(\text{size } \Gamma B(\rho_B))$ is still finite

Core-size and leap generator

- giant 2c-component, of size $\sim \frac{1}{\mu}n + n^{2/3}X$ [Banderier et al'02]

  X Airy distribution of map-type



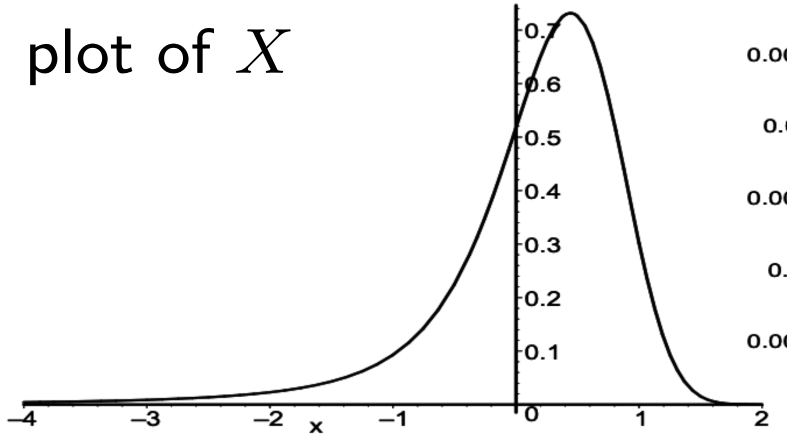


Core-size and leap generator

- giant 2c-component, of size $\sim \frac{1}{\mu}n + n^{2/3}X$ [Banderier et al'02]

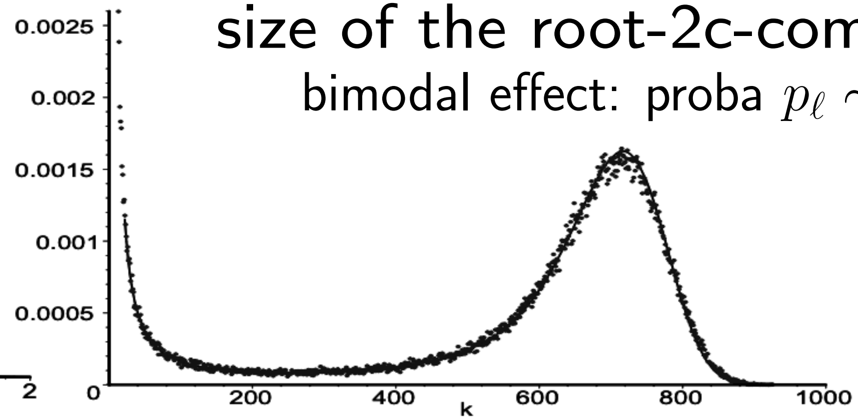
  X Airy distribution of map-type

plot of X



size of the root-2c-component

bimodal effect: proba $p_\ell \sim 1/\mu$ to be giant



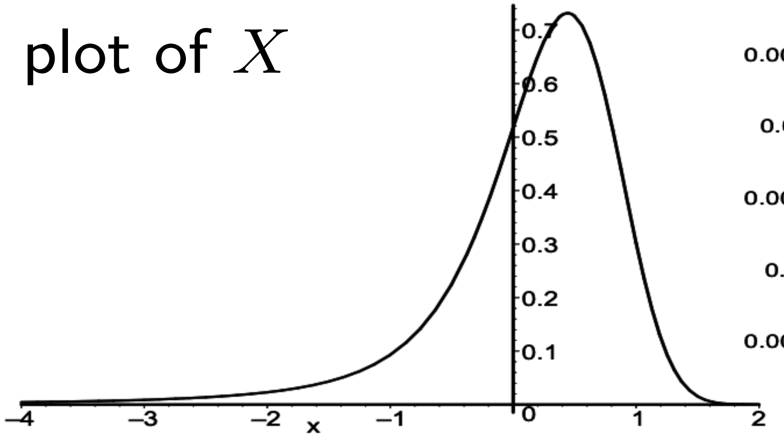
- leap generator $\Gamma\hat{C}[n]$ proceeds the same as in supercritical schemes
  $+$ reroot at random corner (if not, the root-2c-component is a.s. giant)

Core-size and leap generator

- giant 2c-component, of size $\sim \frac{1}{\mu}n + n^{2/3}X$ [Banderier et al'02]

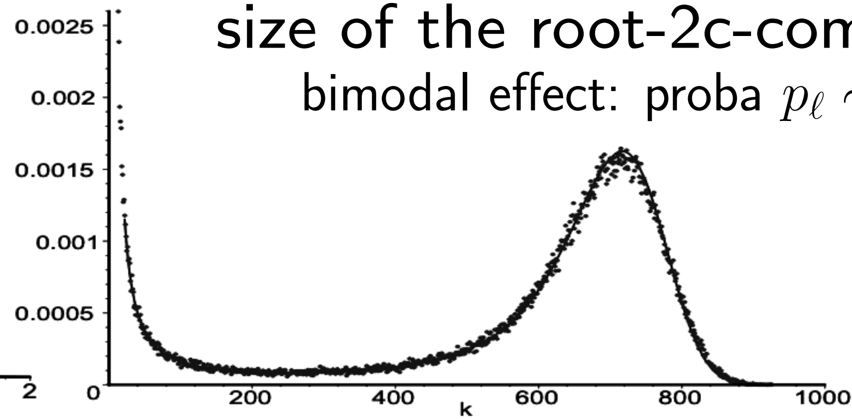
  X Airy distribution of map-type

plot of X



size of the root-2c-component

bimodal effect: proba $p_\ell \sim 1/\mu$ to be giant



- leap generator $\Gamma\hat{C}[n]$ proceeds the same as in supercritical schemes
  $+$ reroot at random corner (if not, the root-2c-component is a.s. giant)

$$d_{\mathrm{TV}}(\pi_n, \pi'_n) \sim \frac{c}{n^{1/3}}$$

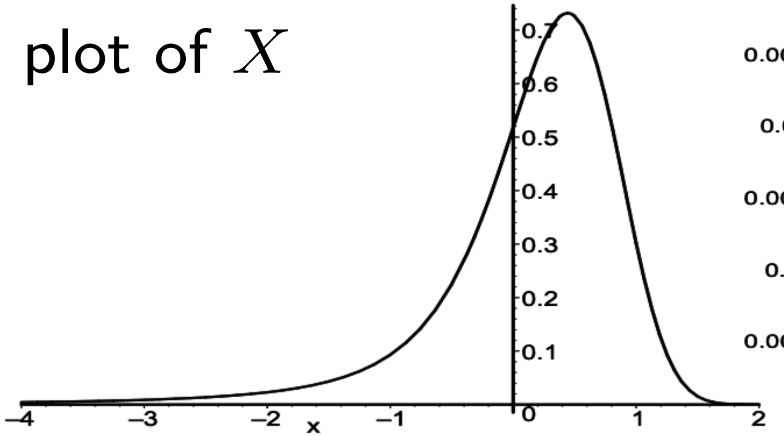
expression of $c \propto \int_{\mathbb{R}} X(x)|x|\mathrm{d}x$

Core-size and leap generator

- giant 2c-component, of size $\sim \frac{1}{\mu} n + n^{2/3} X$ [Banderier et al'02]

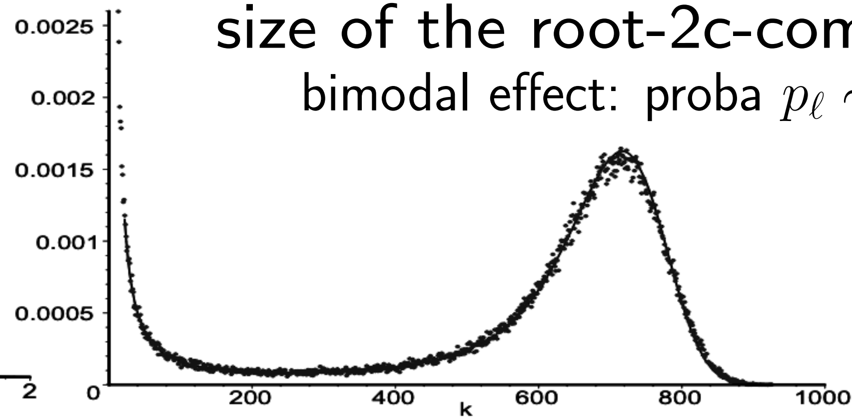
  X Airy distribution of map-type

plot of X



size of the root-2c-component

bimodal effect: proba $p_\ell \sim 1/\mu$ to be giant



- leap generator $\Gamma\hat{C}[n]$ proceeds the same as in supercritical schemes
  $+$ reroot at random corner (if not, the root-2c-component is a.s. giant)

$$d_{\mathrm{TV}}(\pi_n, \pi'_n) \sim \frac{c}{n^{1/3}}$$

expression of $c \propto \int_{\mathbb{R}} X(x)|x|\mathrm{d}x$

applies to block-weighted random maps [Stufler'20], [Fleurat-Salvy'23]

$\&$ to cubic planar graphs

Core extraction generators

core-size study of [Banderier et al'02] motivated by

analysis of core-extraction random generators

[Schaeffer'99]

$$\mathcal{C} = \mathcal{A} \circ \mathcal{B}$$

random generator $\Gamma C[n]$ $\Rightarrow$ random generator ΓA with size $\approx n/\mu$

(fluctuations $\Theta(n^{2/3})$)

exactly uniform, quasi exact-size

(fluctuations $\Theta(n^{1/2})$ in supercritical case)

Core extraction generators

core-size study of [Banderier et al'02] motivated by

analysis of core-extraction random generators

[Schaeffer'99]

$$\mathcal{C} = \mathcal{A} \circ \mathcal{B}$$

random generator $\Gamma C[n]$ $\Rightarrow$ random generator ΓA with size $\approx n/\mu$

(fluctuations $\Theta(n^{2/3})$)

exactly uniform, quasi exact-size

(fluctuations $\Theta(n^{1/2})$ in supercritical case)

---

leap generators

random generators $\Gamma A[k], \Gamma B(x)$ $\Rightarrow$ random generator $\Gamma \hat{C}[n]$

quasi uniform, exact-size

Extension of Devroye's method

work in progress with K. Panagiotou

Example on Cayley trees

Critical Galton-Watson trees for Cayley trees: $\mu = \mathrm{Pois}(1)$

[Duchon et al'04]: algo with $\Theta(\sqrt{n})$ trials, each of complexity $O(n)$

Repeat

draw $d_1 \leftarrow \mathrm{Pois}(1), \dots, d_n \leftarrow \mathrm{Pois}(1)$ independently

if $d_1 + \cdots + d_n = n - 1$ (proba $= \Theta(1/\sqrt{n})$) success

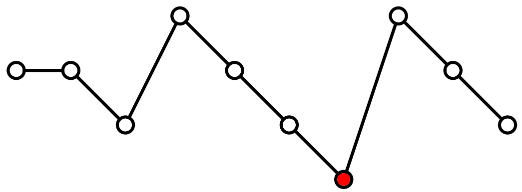
let $W =$ walk whose ith step is $d_i - 1$, for i from 1 to n

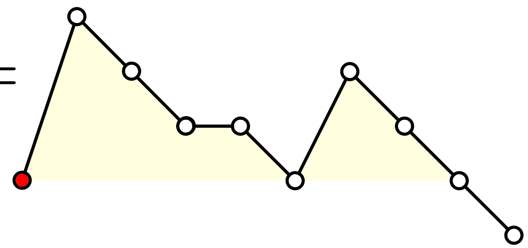
let $W_{\mathrm{pos}} =$ unique positive **cyclic shift** of W

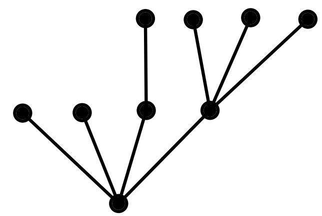
return tree (degree-)encoded by W_{pos}

Ex:

d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8	d_9
1	0	3	0	0	0	4	0	0

$W =$ 

$W_{\mathrm{pos}} =$ 

Tree: 

Example on Cayley trees

Critical Galton-Watson trees for Cayley trees: $\mu = \mathrm{Pois}(1)$

[Duchon et al'04]: algo with $\Theta(\sqrt{n})$ trials, each of complexity $O(n)$

Repeat

draw $d_1 \leftarrow \mathrm{Pois}(1), \ldots, d_n \leftarrow \mathrm{Pois}(1)$ independently

if $d_1 + \cdots + d_n = n-1$ (proba $= \Theta(1/\sqrt{n})$) success

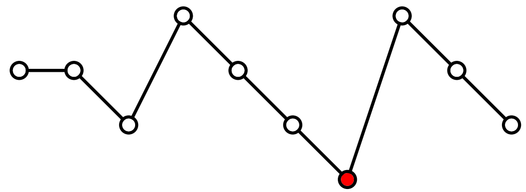
let W = walk whose ith step is $d_i - 1$, for i from 1 to n

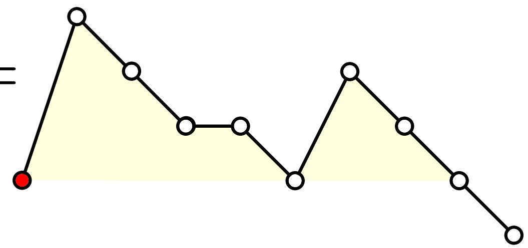
let W_{pos} = unique positive **cyclic shift** of W

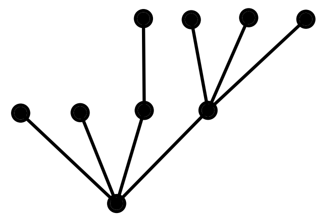
return tree (degree-)encoded by W_{pos}

Ex:

d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8	d_9
1	0	3	0	0	0	4	0	0

$W=$ 

$W_{\mathrm{pos}} =$ 

Tree: 

[Devroye'12]: at each trial draw only degree distribution $N_0, N_1, N_2, \ldots$

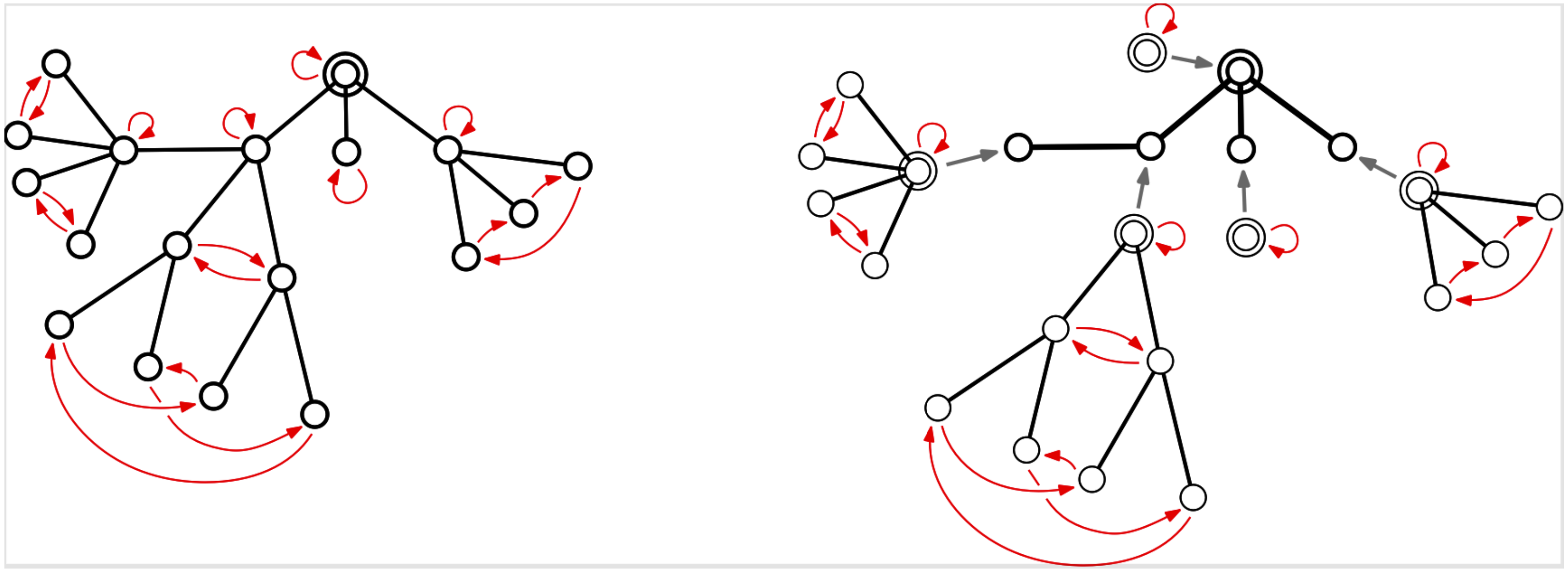
$N_0 \leftarrow \mathrm{Bin}(n, \mu(0)),\ N_1 \leftarrow \mathrm{Bin}(n - N_0, \mu(1)), \ldots$ (cost $\widetilde{O}(1)$)

success if $\sum_i iN_i = n-1$ then $(d_1, \ldots, d_n) \leftarrow \mathrm{shuffle}(0^{N_0} 1^{N_1} \cdots)$

Adapted leap generator for Pólya trees

For $\mathcal{A} =$ Cayley trees let $\mathcal{C} = \mathrm{Sym}(\mathcal{A})$

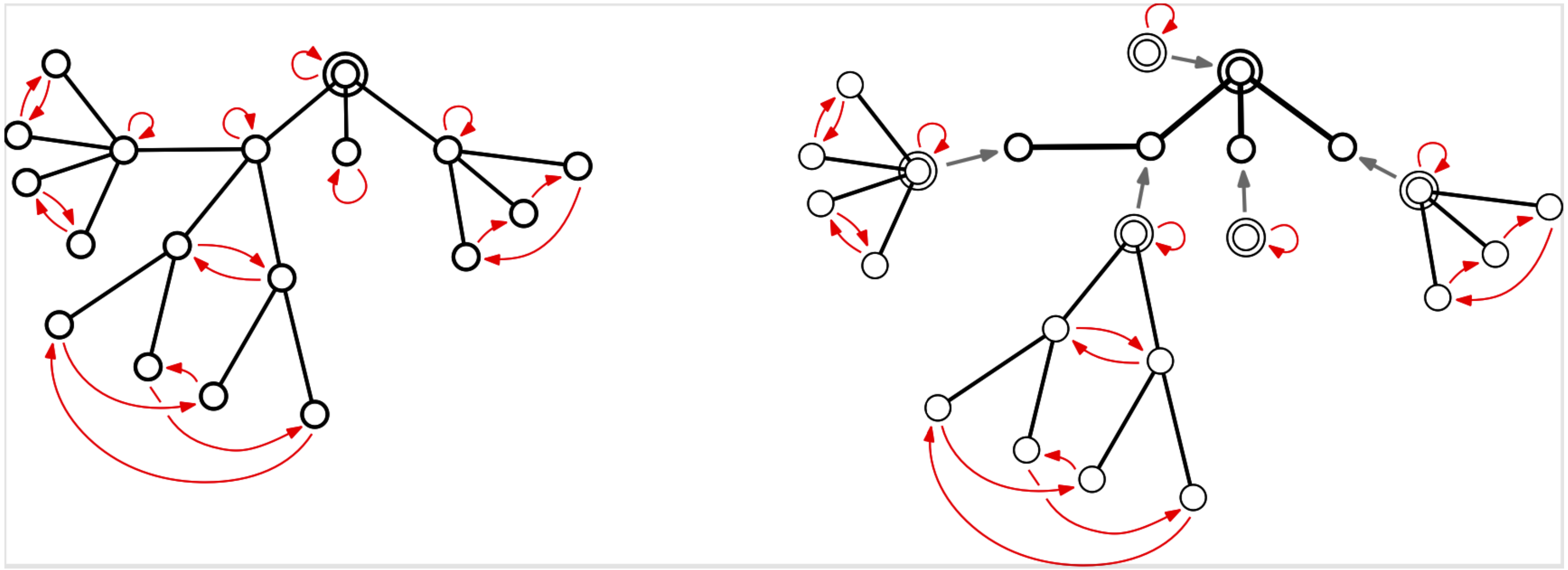
Composition scheme: $\boxed{\mathcal{C} = \mathcal{A} \circ \mathcal{B}}$ where $\mathcal{B} \subset \mathcal{C}$: only the root is fixed



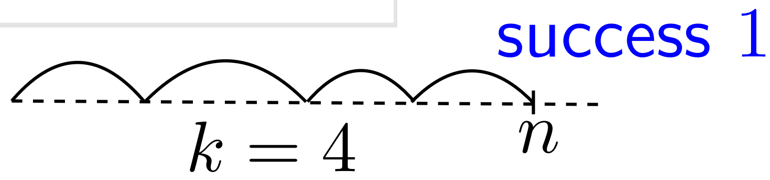
Adapted leap generator for Pólya trees

For $\mathcal{A}$ = Cayley trees let $\mathcal{C} = \mathrm{Sym}(\mathcal{A})$

Composition scheme: $\boxed{\mathcal{C} = \mathcal{A} \circ \mathcal{B}}$ where $\mathcal{B} \subset \mathcal{C}$: only the root is fixed



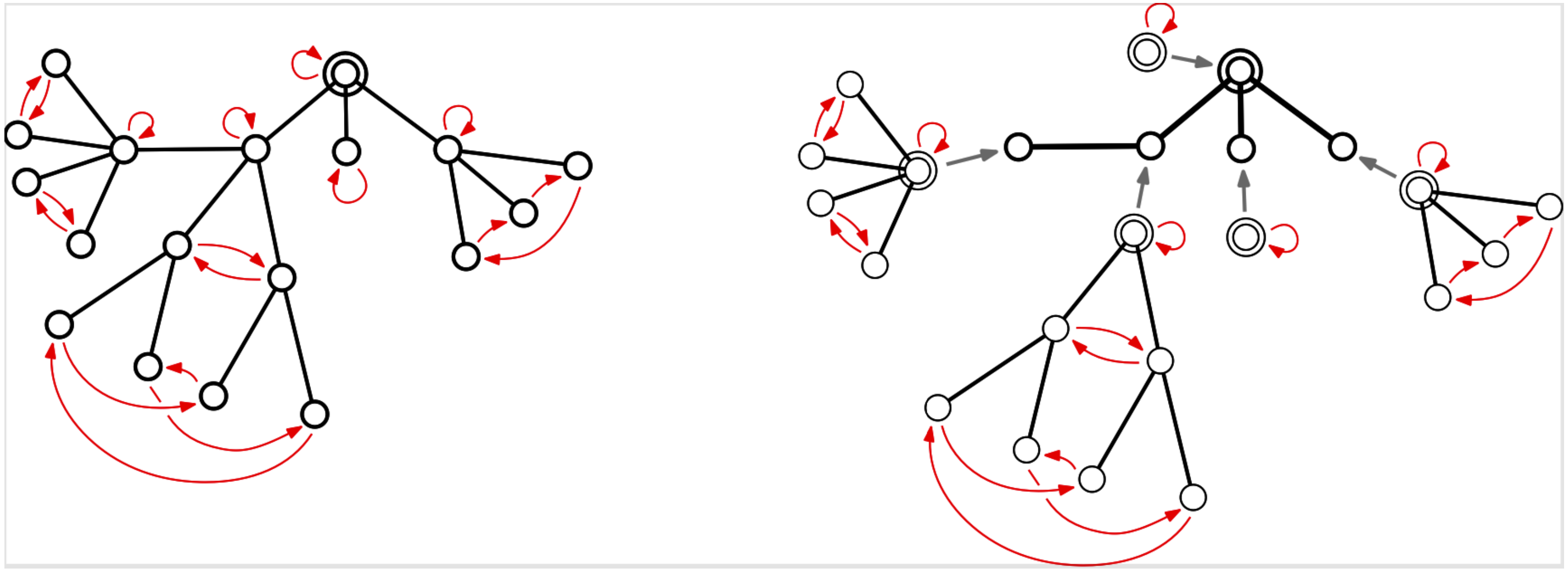
- use the same leap process of calls to $\Gamma B(\rho)$
  (determining if success 1 can be done in time $\widetilde{O}(1)$)



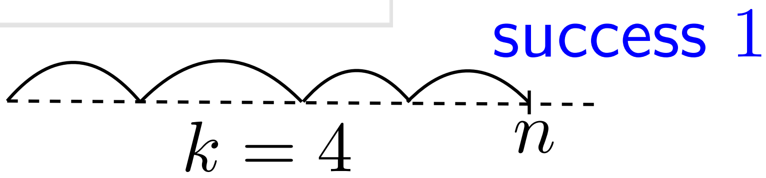
Adapted leap generator for Pólya trees

For $\mathcal{A}$ = Cayley trees   let $\mathcal{C} = \mathrm{Sym}(\mathcal{A})$

Composition scheme: $\boxed{\mathcal{C} = \mathcal{A} \circ \mathcal{B}}$ where $\mathcal{B} \subset \mathcal{C}$: only the root is fixed



- use the same leap process of calls to $\Gamma B(\rho)$
  (determining if success 1 can be done in time $\widetilde{O}(1)$)

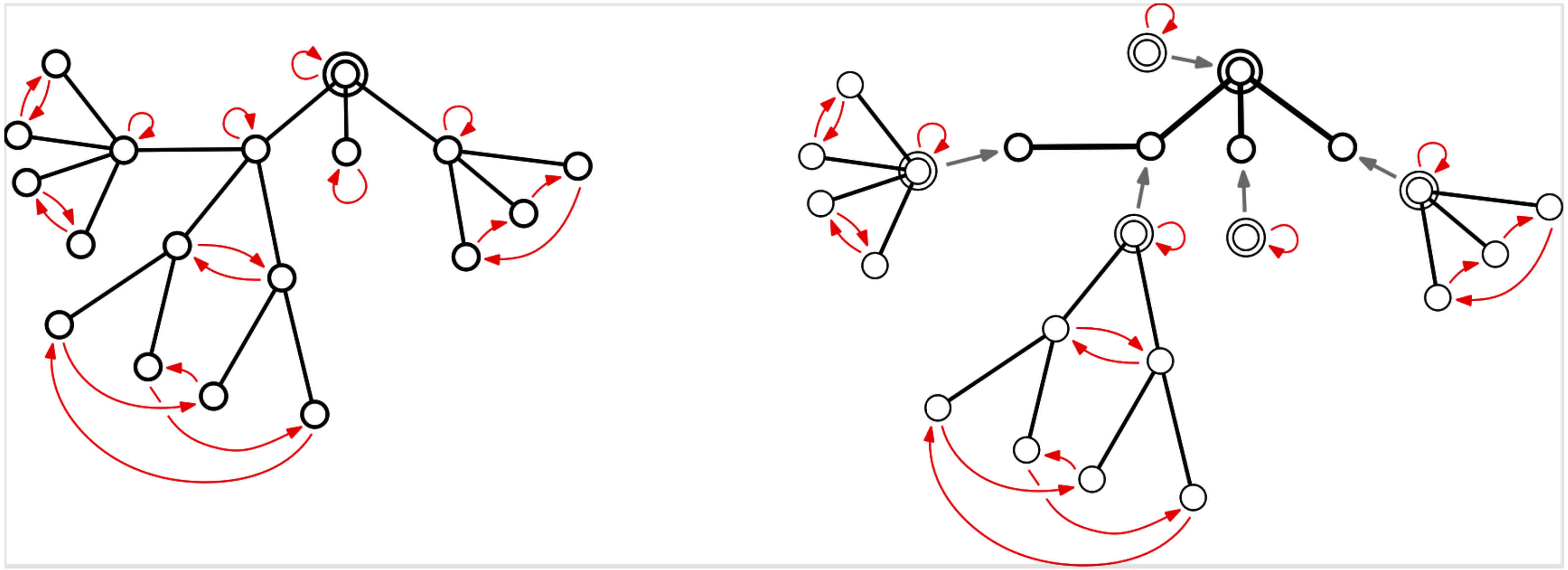


- to generate core (Cayley tree of size k)
  perform a **single trial** of Devroye's method
  if success (success 2), return composed tree
  otherwise restart the whole process

Adapted leap generator for Pólya trees

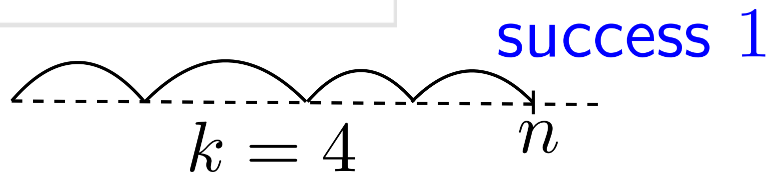
For $\mathcal{A} =$ Cayley trees   let $\mathcal{C} = \mathrm{Sym}(\mathcal{A})$

Composition scheme: $\boxed{\mathcal{C} = \mathcal{A} \circ \mathcal{B}}$ where $\mathcal{B} \subset \mathcal{C}$: only the root is fixed



- use the same leap process of calls to $\Gamma B(\rho)$

  (determining if success 1 can be done in time $\widetilde{O}(1)$)



- to generate core (Cayley tree of size k)

  perform a **single trial** of Devroye's method

  if success (success 2), return composed tree

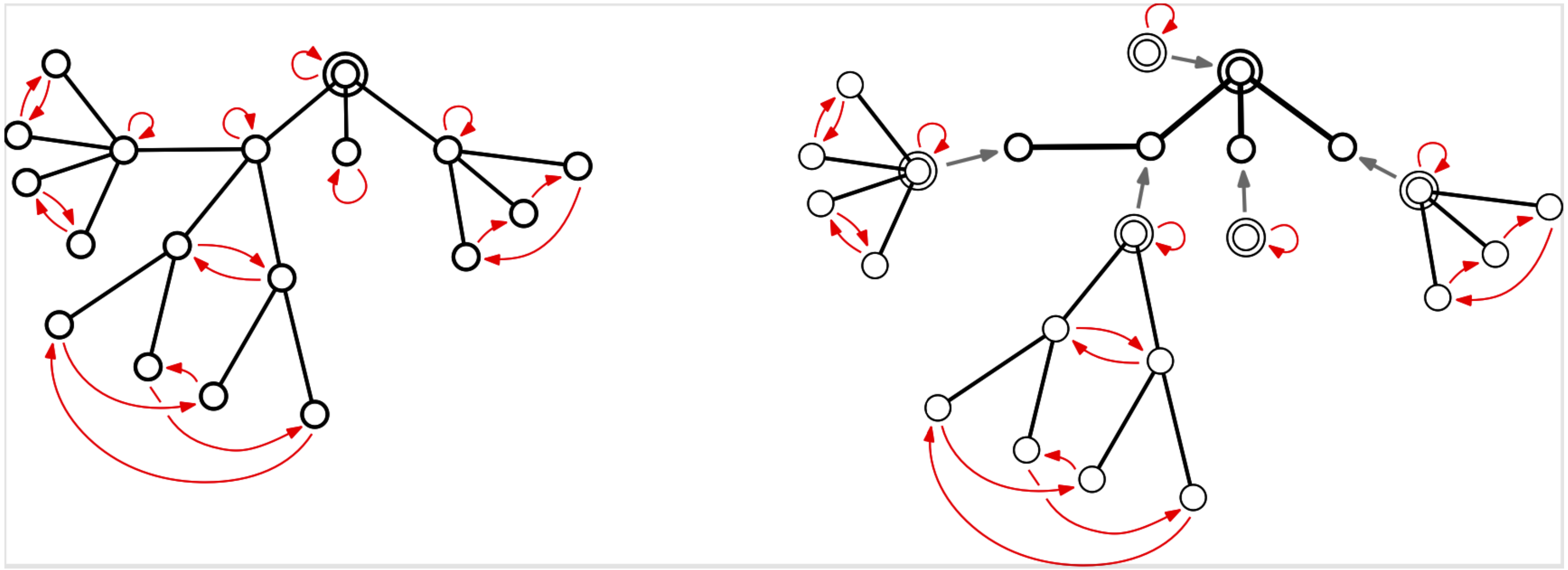
  otherwise restart the whole process

structure in $\mathcal{C}_{n,k}$ has proba $\propto k$

Adapted leap generator for Pólya trees

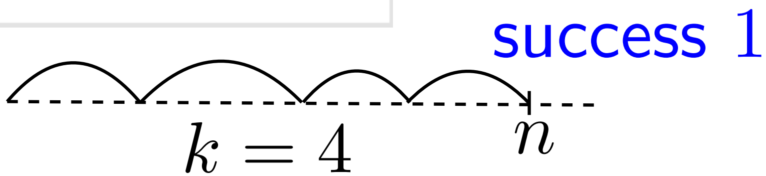
For $\mathcal{A}$ = Cayley trees   let $\mathcal{C} = \mathrm{Sym}(\mathcal{A})$

Composition scheme: $\boxed{\mathcal{C} = \mathcal{A} \circ \mathcal{B}}$ where $\mathcal{B} \subset \mathcal{C}$: only the root is fixed



- use the same leap process of calls to $\Gamma B(\rho)$

  (determining if success 1 can be done in time $\widetilde{O}(1)$)



- to generate core (Cayley tree of size k)

  perform a **single trial** of Devroye's method

  if success (success 2), return composed tree

  otherwise restart the whole process

---

structure in $\mathcal{C}_{n,k}$ has proba $\propto k$ (instead of $\propto \frac{k!}{k^{k-1}} e^k = \Theta(k^{3/2})$)

Adapted leap generator for Pólya trees

obtained sampler on $\mathcal{C}_n$ of complexity $O(n)$

for $\gamma \in \mathcal{C}_n, \quad \mathbb{P}(\gamma) \propto k$ (with k =core-size)

add rejection-step with $p_{\text{success}} = 1/k$ $\quad\Rightarrow$ number of trials $= \Theta(n)$

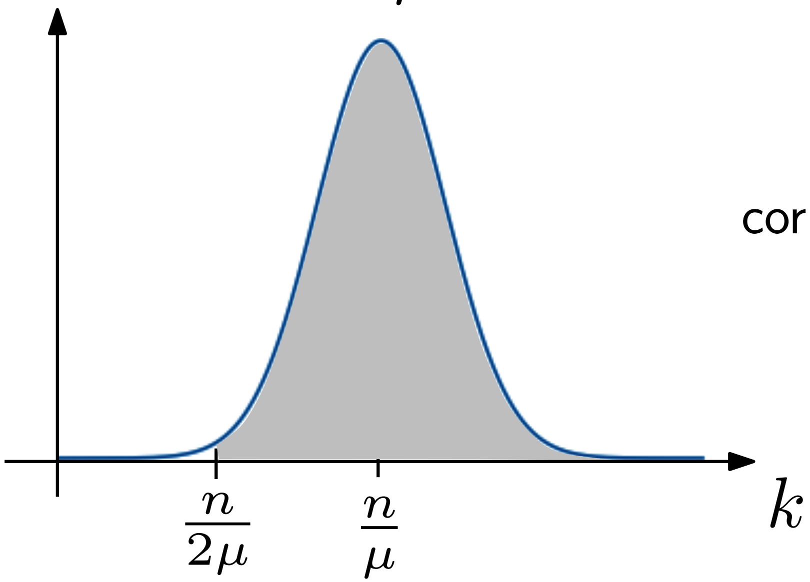
Adapted leap generator for Pólya trees

obtained sampler on $\mathcal{C}_n$ of complexity $O(n)$

for $\gamma \in \mathcal{C}_n$, $\mathbb{P}(\gamma) \propto k$ (with k =core-size)

add rejection-step with $p_{\text{success}} = 1/k$ $\Rightarrow$ number of trials $= \Theta(n)$

rather use $p_{\text{success}} = \min\Big(1, \dfrac{n}{2\mu}\dfrac{1}{k}\Big)$ $\Rightarrow$ number of trials $= \Theta(1)$



corrects bias in area of proba $1 - O(e^{-cn})$

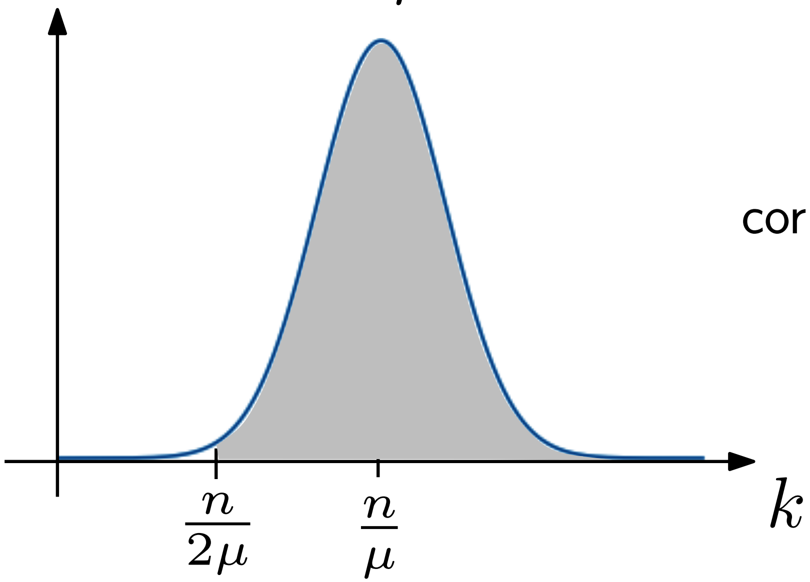
Adapted leap generator for Pólya trees

obtained sampler on $\mathcal{C}_n$ of complexity $O(n)$

for $\gamma \in \mathcal{C}_n$, $\mathbb{P}(\gamma) \propto k$ (with k =core-size)

add rejection-step with $p_{\text{success}} = 1/k$ $\Rightarrow$ number of trials $= \Theta(n)$

rather use $p_{\text{success}} = \min\Big(1, \frac{n}{2\mu}\frac{1}{k}\Big)$ $\Rightarrow$ number of trials $= \Theta(1)$



corrects bias in area of proba $1 - O(e^{-cn})$

Rk: method should adapt to Pólya trees with vertex-degrees in set Ω
(e.g. non-plane k-ary trees)