Bootstrap random walks

Kais Hamza

Monash University

Joint work with Andrea Collevecchio & Meng Shi

The model setting

- Let $(\xi_n)_n$ be a sequence of independent identically distributed random variables such that

$$\mathbb{P}(\xi_i = -1) = \mathbb{P}(\xi_i = +1) = \frac{1}{2}.$$

- Let $\eta_k = \prod_{j=1}^{k} \xi_j$. Then $(\eta_n)_{n \geq 0} \stackrel{d}{=} (\xi_n)_{n \geq 0}$.
- The σ-algebras generated by these two sequences are the same, i.e.

$$\sigma(\xi_1, \cdots, \xi_n) = \sigma(\eta_1, \cdots, \eta_n).$$

- Let $X_n = \sum_{k=1}^{n} \xi_k$ and $Y_n = \sum_{k=1}^{n} \eta_k$, with $X_0 = 0,\ Y_0 = 0$.

Some basic observations

- $(X_n)_n$ and $(Y_n)_n$ are strongly dependent, yet $W_n = (X_n, Y_n)$ behaves like a 2-dimensional random walk. Let $W'_n = (X_n, Y'_n)$ with $(Y'_n)_n$ is an independent copy of $(X_n)_n$.

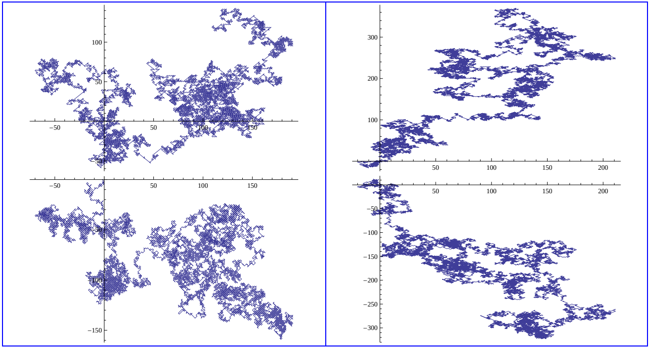

Figure : Two simulations of $(W_n)_n$ and $(W'_n)_n$

Some basic observations

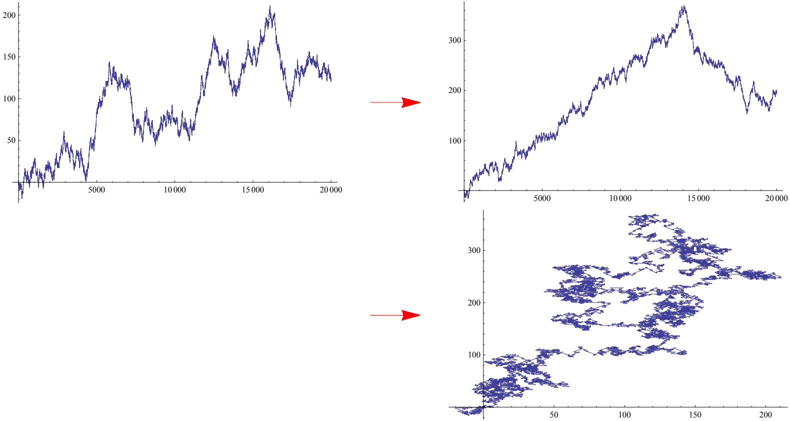

Figure : $(X_n)_n \xrightarrow{deter} (Y_n)_n \xrightarrow{deter} (W_n)_n$

Some basic observations

- Locally, $(W_n)_n$'s behaviour is different to that of a 2-dimensional random walk (with independent components).

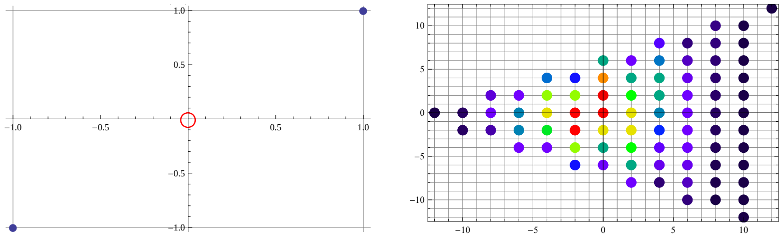

Figure : Possible positions for W_1 and W_{12}

Some basic observations

- $(X_n)_n$ and $(Y_n)_n$ are both simple symmetric random walks.
- If $\mathbb{P}(\xi_n = 1) = p \neq 1/2$, then
  - $\eta_n \overset{d}{\neq} \xi_n$:
  $$\mathbb{P}(\eta_n = 1) = \frac{1}{2}(1 + (2p-1)^n);$$
  - $\eta_1, \ldots, \eta_n$ are not independent:
  $$\begin{aligned} &\mathbb{P}(\eta_n = \varepsilon_n | \eta_n = \varepsilon_1, \ldots, \eta_{n-1} = \varepsilon_{n-1}) \\ &= \mathbb{P}(\eta_n = \varepsilon_n | \eta_{n-1} = \varepsilon_{n-1}) = \left(\frac{p}{1-p}\right)^{(1+\varepsilon_{n-1}\varepsilon_n)/2}. \end{aligned}$$
- $\mathbb{E}[X_n Y_n] = \sum_{k,l=1}^{n} \mathbb{E}[\xi_k \eta_l] = 1 + \sum_{(k,l)\neq(1,1)} \mathbb{E}[\xi_k \eta_l] = 1.$

The Markov property

- W_n is a time-inhomogeneous Markov process:

$$\begin{cases} X_{n+1} &= X_n + \xi_{n+1} \\ Y_{n+1} &= Y_n + (-1)^{\frac{n-X_n}{2}} \xi_{n+1}. \end{cases}$$

- But, with $K_n = n \mod 4$, (X_n, Y_n, K_n) is a time-homogeneous Markov process:

$$\begin{cases} X_{n+1} &= X_n + \xi_{n+1} \\ Y_{n+1} &= Y_n + (-1)^{\frac{K_n-X_n}{2}} \xi_{n+1} \\ K_{n+1} &= K_n + 1 \mod 4 \end{cases}$$

The probability of return to $(0,0)$ and recurrence

Proposition

$$\mathbb{P}\left(W_{4n} = (0,0)\right) = \binom{2n-1}{n}\binom{2n}{n}\left(\frac{1}{2}\right)^{4n} \sim \frac{1}{4\pi n}$$

and

$$\mathbb{P}\left(W_{4n+2} = (0,0)\right) = \binom{2n+1}{n+1}\binom{2n}{n}\left(\frac{1}{2}\right)^{4n+2} \sim \frac{1}{4\pi n}.$$

It follows that

$$\sum_{n=0}^{\infty} \mathbb{P}\left(W_{4n} = (0,0)\right) = +\infty.$$

Theorem

$W_n = (X_n, Y_n)$ is recurrent.

The transition probabilities

Proposition

If $\frac{1}{2}(n-k)$ is even,

$$\mathbb{P}\left(X_n = k, Y_n = l\right) = \binom{\frac{n+l}{2}}{\frac{n+k+2l}{4}}\binom{\frac{n-l-2}{2}}{\frac{n+k-2l}{4}}\left(\frac{1}{2}\right)^n$$

If $\frac{1}{2}(n-k)$ is odd,

$$\mathbb{P}\left(X_n = k, Y_n = l\right) = \binom{\frac{n+l}{2}}{\frac{n+k+2l}{4}}\binom{\frac{n-l-2}{2}}{\frac{n+k-2l}{4}}\left(\frac{1}{2}\right)^n$$

The three-dimensional process

- Consider the three-dimensional random walk (X_n, Y_n, Z_n), also denoted W_n, where

$$Z_n = \sum_{k=1}^{n} \zeta_k, \ \ n \geq 1 \text{ and } Z_0 = 0,$$

- $\zeta_k = \prod_{j=1}^{k} \eta_j = \prod_{j=1}^{k} \prod_{i=1}^{j} \xi_i$:

$$\zeta_1 = \xi_1 \qquad \zeta_2 = \xi_1 \xi_2 \qquad \zeta_3 = \xi_1 \xi_2 \xi_3$$

$$\zeta_4 = \xi_1 \xi_2 \xi_3 \xi_4 \qquad \zeta_5 = \xi_1 \xi_2 \xi_3 \xi_4 \xi_5 \qquad \zeta_6 = \xi_1 \xi_2 \xi_3 \xi_4 \xi_5 \xi_6$$

The probability of return to $(0,0,0)$ and recurrence

Proposition

1. For any $n \geq 2$,

$$\mathbb{P}\left(W_{4n}=0\right)=2^{-4n}\sum_{k=0}^{n-2}\binom{n-1}{k}\binom{n}{k+1}\binom{n}{k+1}\binom{n-1}{k+1}$$

and

$$\mathbb{P}\left(W_{4n+2}=0\right)=2^{-(4n+2)}\sum_{k=1}^{n}\binom{n+1}{k}\binom{n-1}{k-1}\binom{n}{k}\binom{n+1}{k+1}.$$

2. $\mathbb{P}(W_{2n}=0)=O(n^{\alpha-2})$, for any $\alpha \in (1/2,1)$.
3. $(W_n)_n$ is transient; it will visit the origin finitely often.

A functional central limit theorem

- Donsker: $\mathfrak{X}_n(t) = \frac{1}{\sqrt{n}} X_{[nt]}$ converges weakly to a Brownian motion $(t \in [0,1])$.
- The same is true for $\mathfrak{Y}_n(t) = \frac{1}{\sqrt{n}} Y_{[nt]}$ and $\mathfrak{Z}_n(t) = \frac{1}{\sqrt{n}} Z_{[nt]}$.
- The functional dependence between $\mathfrak{X}_n(t)$, $\mathfrak{Y}_n(t)$ and $\mathfrak{Z}_n(t)$ is complete lost at infinity.

Theorem

$\mathfrak{W}_n(t) = (\mathfrak{X}_n(t), \mathfrak{Y}_n(t), \mathfrak{Z}_n(t))$ converges weakly to a three-dimensional Brownian motion.

The model setting

- $\eta_{0,n} = \xi_n$ and $\eta_{1,n} = \eta_n = \prod_{k=1}^{n} \xi_k$.
- $\eta_{2,n} = \prod_{k=1}^{n} \eta_{1,k} = \prod_{k=1}^{n} \left(\prod_{j=1}^{k} \xi_j\right)$.
- $\eta_{3,n} = \prod_{k=1}^{n} \eta_{2,k} = \prod_{\ell=1}^{n} \left(\prod_{k=1}^{\ell} \left(\prod_{j=1}^{k} \xi_j\right)\right)$.

$$
\begin{array}{cccc}
\kappa = 0 & \kappa = 1 & \kappa = 2 & \kappa = 3 \\
\begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \\ \xi_5 \\ \xi_6 \end{pmatrix} &
\begin{pmatrix} \xi_1 \\ \xi_1\xi_2 \\ \xi_1\xi_2\xi_3 \\ \xi_1\xi_2\xi_3\xi_4 \\ \xi_1\xi_2\xi_3\xi_4\xi_5 \\ \xi_1\xi_2\xi_3\xi_4\xi_5\xi_6 \end{pmatrix} &
\begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_1\xi_3 \\ \xi_2\xi_4 \\ \xi_1\xi_3\xi_5 \\ \xi_2\xi_4\xi_6 \end{pmatrix} &
\begin{pmatrix} \xi_1 \\ \xi_1\xi_2 \\ \xi_2\xi_3 \\ \xi_3\xi_4 \\ \xi_1\xi_4\xi_5 \\ \xi_1\xi_2\xi_5\xi_6 \end{pmatrix}
\end{array}
$$

$\xi_n^2 = (\pm 1)^2 = 1$. Need to understand which ξ's are "switched on".

The model setting

- $\eta_{0,n} = \xi_n$ and $\eta_{1,n} = \eta_n = \prod_{\ell=1}^{n} \xi_\ell$.
- $\eta_{\kappa+1,n} = \prod_{\ell=1}^{n} \eta_{\kappa,\ell}$.
- $\eta_{\kappa,n} = \prod_{\ell=1}^{n} \xi_{n-\ell+1}^{\nu_{\kappa,\ell}}$, where $\nu_{\kappa,\ell}$ is a deterministic array and
  - $\nu_{0,1} = 1$ and $\nu_{0,n} = 0$ for $n \geq 2$;
  - $\nu_{\kappa,1} = 1$ for $\kappa \geq 0$;
  - $\nu_{\kappa+1,n+1} = \nu_{\kappa+1,n} + \nu_{\kappa,n+1} \mod 2$.
- For $\kappa \geq 1$ and $n \geq 1$,

$$\nu_{\kappa,n} = \binom{n+\kappa-2}{n-1} \mod 2.$$

The binomial connection (Sierpinski triangle)

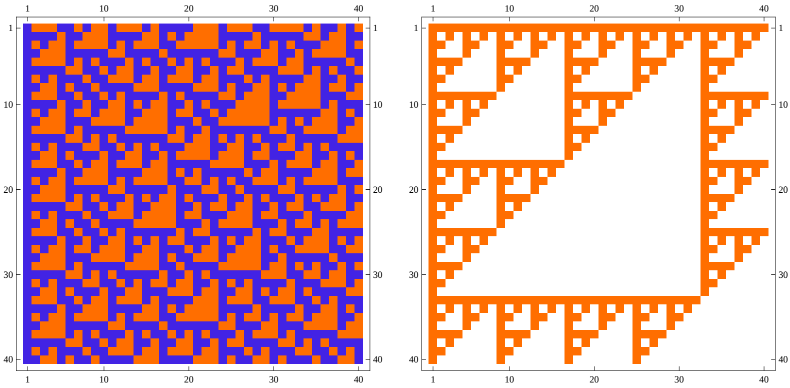

Figure : A graphical representation of $\eta_{\kappa,n}$ (left) and $\nu_{\kappa,n}$ (right)

A multi-dimensional extension

- $Y_{0,0} = X_0 = 0$ and $Y_{0,n} = X_n = \sum_{\ell=1}^{n} \xi_\ell$.
- $Y_{\kappa,0} = 0$ and $Y_{\kappa,n} = \sum_{\ell=1}^{n} \eta_{\kappa,\ell}$.
- $Y_{\kappa,n}$ is a simple symmetric random walk.
- $(X_n, Y_{\kappa,n})$ "behaves" like a two-dimensional random walk.
- In fact, $W_{\kappa,n} = (X_n, Y_{1,n}, \ldots, Y_{\kappa,n})$ "behaves" like a $(\kappa+1)$-dimensional random walk.

A multi-dimensional extension

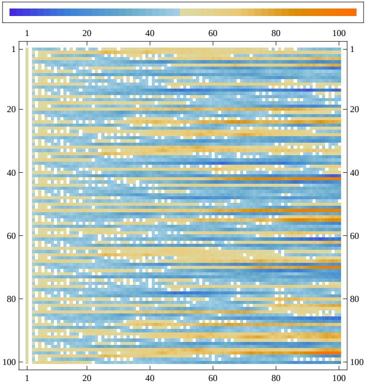

Figure : A graphical representation of $W_{\kappa,n}$ (left) and a multi-dimensional random walk (right)

A second functional central limit theorem

Theorem (Lucas, 1878)

A binomial coefficient $\binom{n}{m}$ is divisible by a prime p if and only if at least one of the base p digits of m is greater than the corresponding digit of n.

- $\binom{73}{25} = 2321476405329962052$ is even:

73	1001001_2
25	0011001_2

- For example, for $\kappa = 2^\omega$ and $2 \leq n \leq 2^\omega$, then $\nu_{\kappa,n} = 0$.

Theorem

$\mathfrak{W}_{\kappa,n}(t) = (\mathfrak{X}_n(t), \mathfrak{Y}_{1,n}(t)), \ldots, \mathfrak{Y}_{\kappa,n}(t))$ converges weakly to a $(\kappa + 1)$-dimensional Brownian motion.

Three-value model

- Suppose $(\xi_n)_n$ i.i.d. uniform on $\mathcal{U} = \{u, a, b\}$.
- In general, $ab \notin \mathcal{U}$. Therefore $\xi_1 \xi_2 \overset{d}{\neq} \xi_2$.
- How do we recycle in this case?
- Repace $\times$ with $\otimes$:

$\otimes$	u	a	b
u	u	a	b
a	a	b	u
b	b	u	a

- Observe that $u^{\otimes 3} = a^{\otimes 3} = b^{\otimes 3} = u$.
- If $\eta_n = \bigotimes_{\ell=1}^{n} \xi_\ell$ then $(\eta_n)_n \overset{d}{=} (\xi_n)_n$.
- This can be repeated and generalised.

General case

- $(\xi_n)_n$ i.i.d. uniform on $\mathcal{U} = \{u_0, u_1, \dots, u_{p-1}\}$, where p is prime and $\sum_{i=0}^{p-1} u_i = 0$.
- Form an Abelian (cyclic) group $(\mathcal{U}, \otimes)$.
- Define $\eta_{\kappa,n}$: $\eta_{1,n} = \bigotimes_{\ell=1}^{n} \xi_\ell$ and $\eta_{\kappa,n} = \bigotimes_{\ell=1}^{n} \eta_{\kappa-1,\ell} = \bigotimes_{\ell=1}^{n} \xi_{n-\ell+1}^{\otimes \nu_{\kappa,\ell}}$ with $\nu_{\kappa,n} = \binom{n+\kappa-2}{n-1} \mod p$.
- $(\eta_{\kappa,n})_n$ has the same distribution as $(\xi_n)_n$.
- Let $Y_{\kappa,n} = \sum_{\ell=1}^{n} \eta_{\kappa,\ell}$, with $Y_{\kappa,0} = 0$.

General case

Proposition

The vector $(\eta_{0,n+\kappa}, \dots, \eta_{\kappa,n+\kappa})$ is uniform over $\mathcal{U}^{\kappa+1}$ and is independent of $\mathcal{F}_{n-1}$. In particular, $\eta_{0,n+\kappa}, \dots, \eta_{\kappa,n+\kappa}$ are independent.

Let $\sigma^2 = \mathbb{E}[\xi_n^2] = \frac{1}{p}(u_0^2 + \cdots + u_{p-1}^2)$ and $\mathfrak{Y}_{K,n}(t) = \frac{1}{\sigma\sqrt{n}} Y_{K,\lfloor nt \rfloor}$.

Theorem

$\mathfrak{W}_{\kappa,n}(t) = (\mathfrak{X}_n(t), \mathfrak{Y}_{1,n}(t)), \dots, \mathfrak{Y}_{\kappa,n}(t))$ converges weakly to a $(\kappa+1)$-dimensional Brownian motion.

- What happens when p is not prime?
- We add sufficiently many zeroes...

Representing $(\phi_n)_n$

- Let $\eta_n = \phi_n(\xi_1, \ldots, \xi_n)$. Look for $(\phi_n)_n$ s.t. $(\eta_n)_n \overset{d}{=} (\xi_n)_n$.

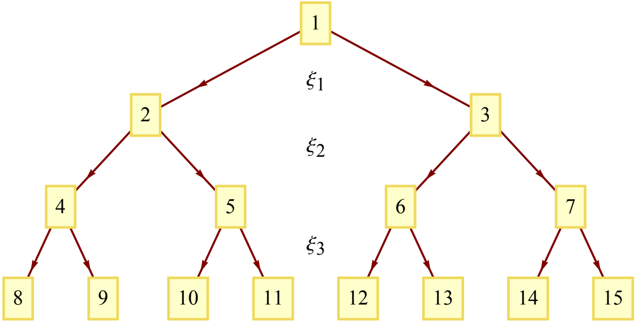

Figure : The tree representation of the functions ϕ_n.

Representing $(\phi_n)_n$

- Let $\eta_n = \phi_n(\xi_1, \ldots, \xi_n)$. Look for $(\phi_n)_n$ s.t. $(\eta_n)_n \overset{d}{=} (\xi_n)_n$.

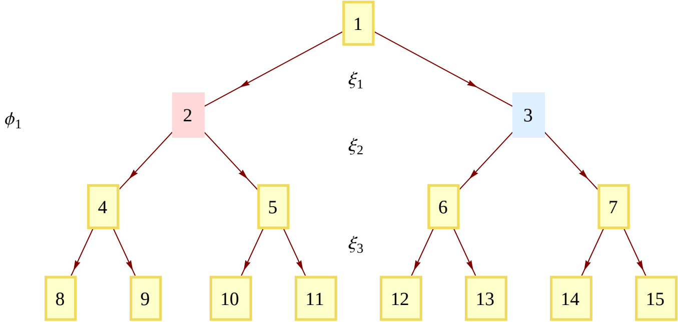

Figure : The tree representation of the functions ϕ_n.

Representing $(\phi_n)_n$

- Let $\eta_n = \phi_n(\xi_1, \ldots, \xi_n)$. Look for $(\phi_n)_n$ s.t. $(\eta_n)_n \stackrel{d}{=} (\xi_n)_n$.

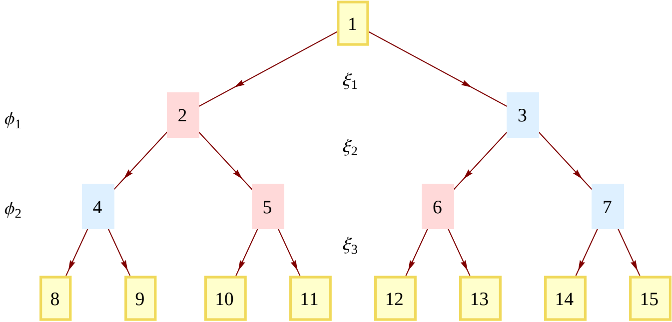

Figure : The tree representation of the functions ϕ_n.

Representing $(\phi_n)_n$

- Let $\eta_n = \phi_n(\xi_1, \ldots, \xi_n)$. Look for $(\phi_n)_n$ s.t. $(\eta_n)_n \overset{d}{=} (\xi_n)_n$.

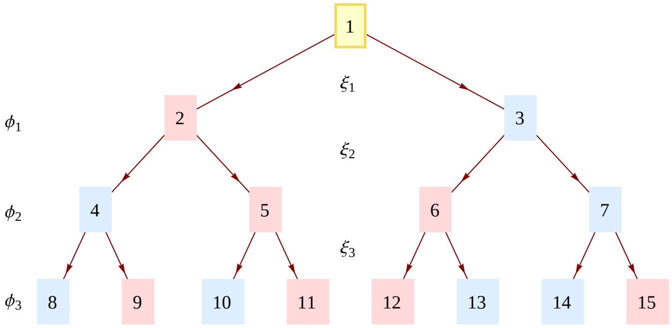

Figure : The tree representation of the functions ϕ_n.

Describing $(\phi_n)_n$

Theorem

Let $\eta_n = \phi_n(\xi_1, \ldots, \xi_n)$, then $(\eta_n)_n \overset{d}{=} (\xi_n)_n$ if and only if $\phi_n(x_1, \ldots, x_n)$ is of the following form:

$$\phi_1(x_1) = (-1)^{\alpha_1} x_1$$

and for $n \geq 2$, with $\mathbb{K}(n)$, the set of all non-empty subsets of $\{1, \ldots, n-1\}$,

$$\phi_n(x_1, \ldots, x_n) = (-1)^{\alpha_n} \left(\prod_{K \in \mathbb{K}(n)} x_{[K]}^{\beta_{n,K}} \right) x_n,$$

where $x_{[K]} = \max_{k \in K} x_k$ and $\alpha_n, \beta_{n,K} \in \{0, 1\}$.

An example

$\eta_1 = \xi_1$, $\eta_2 = \xi_2 \xi_1$ and $\eta_n = \max(\xi_{n-2}, \xi_{n-1}) \xi_n$ for $n \geq 3$.

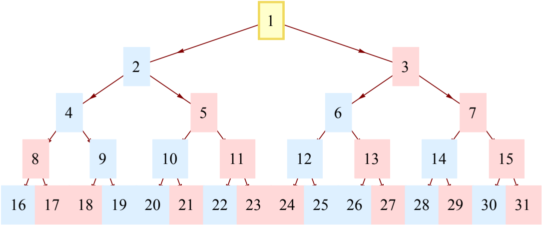

Figure : The tree representation of the functions $\max(x_{n-2}, x_{n-1}) x_n$.

Theorem

$(\mathfrak{X}_n(t), \mathfrak{Y}_n(t)))$ converges weakly to a correlated Brownian motion.

An example

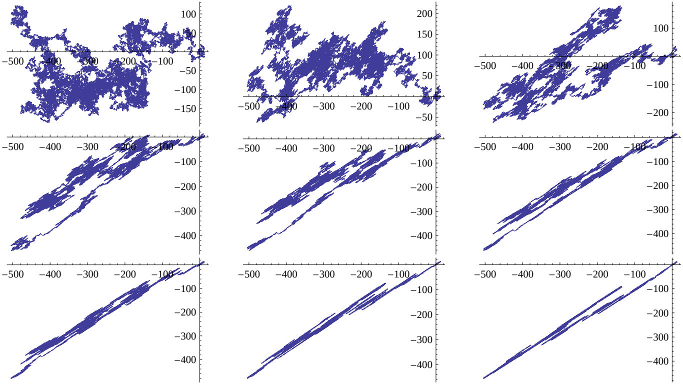

Figure : $(W_n)_n$ for $\phi_n(x_1, \ldots, x_n) = \max(x_{n-k}, \ldots, x_{n-1})x_n$.

A possible application – a greedy random bits generator

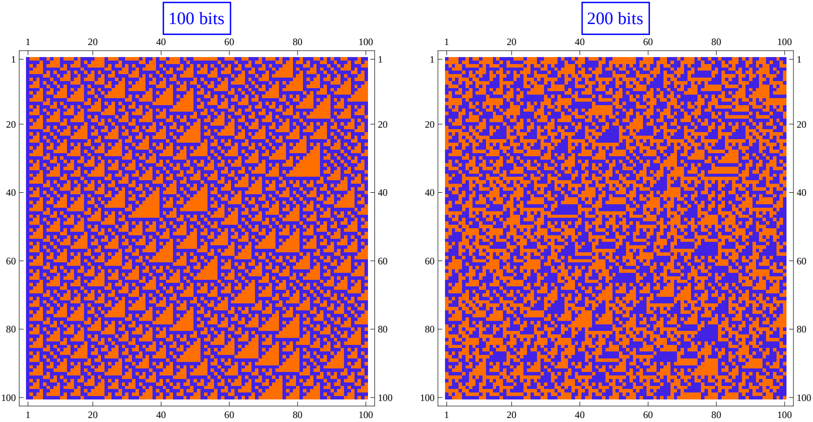

Figure : Simulate or compute?

A possible application – a greedy random bits generator

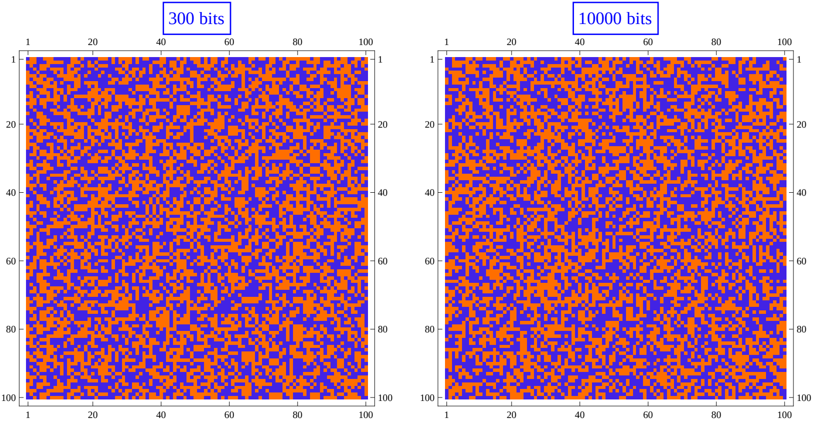

Figure : Simulate or compute?

A connection – cellular automata

- Model could be regarded as a long-range cellular automaton.

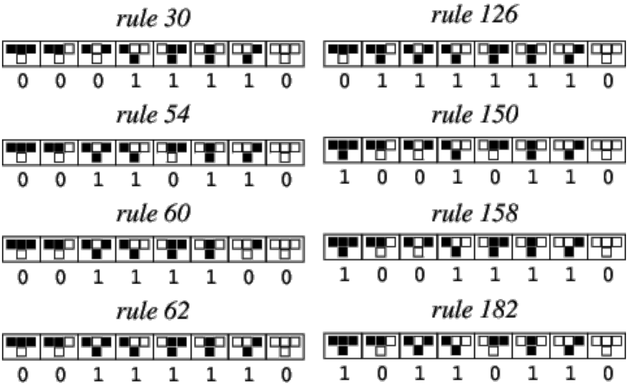

Figure : Elementary cellular automata

A connection – cellular automata

- Model could be regarded as a long-range cellular automaton.

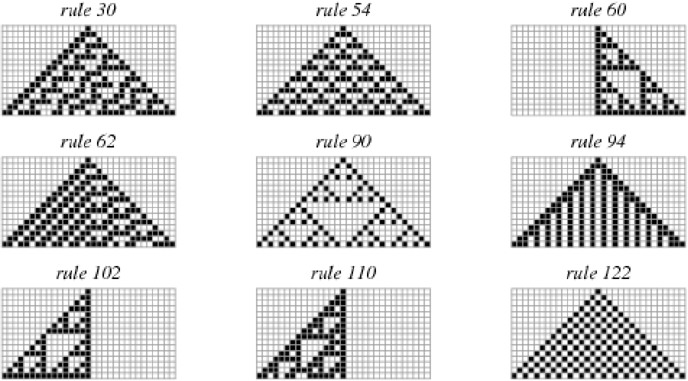

Figure : Elementary cellular automata

A connection – cellular automata

- Up to a "sliding" of the columns, it is also equivalent to CA60.

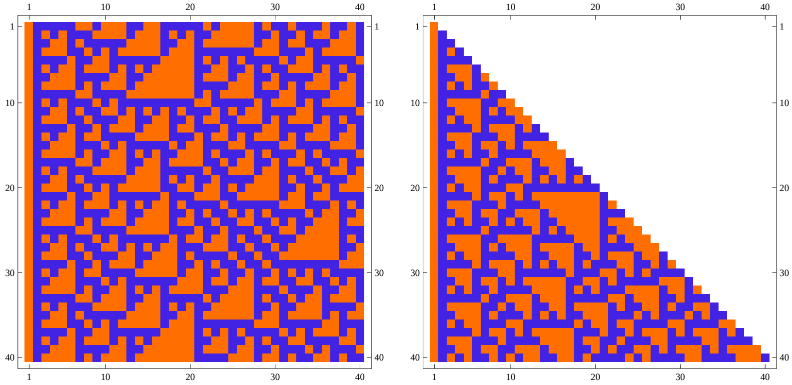

Figure : Cellular automaton 60

A related question – percolation (work in progress)

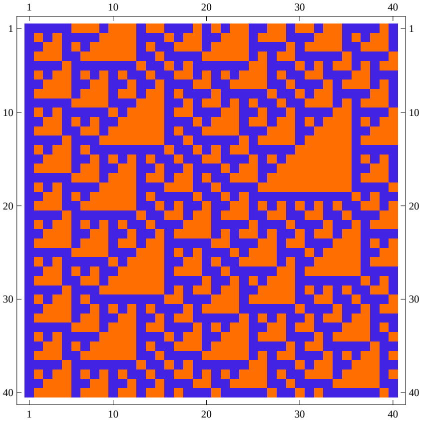

Figure : $p = 0.5$

A related question – percolation (work in progress)

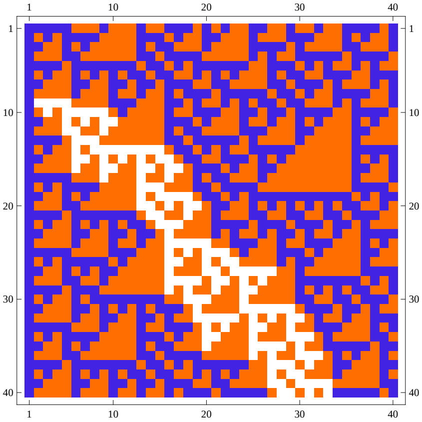

Figure : $p = 0.5$

A related question – percolation (work in progress)

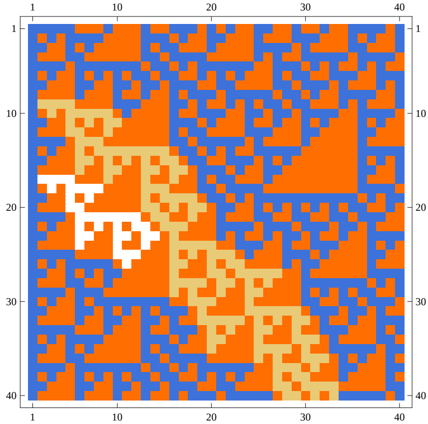

Figure : $p = 0.5$

A related question – percolation (work in progress)

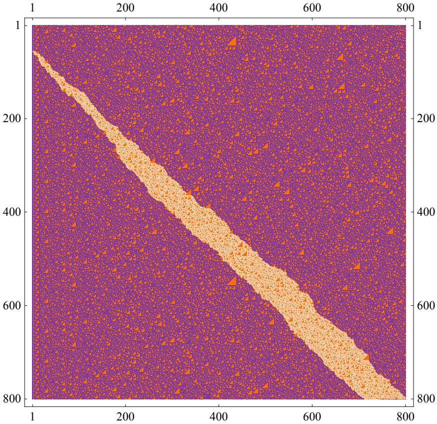

Figure : $p = 0.5$

A related question – percolation (work in progress)

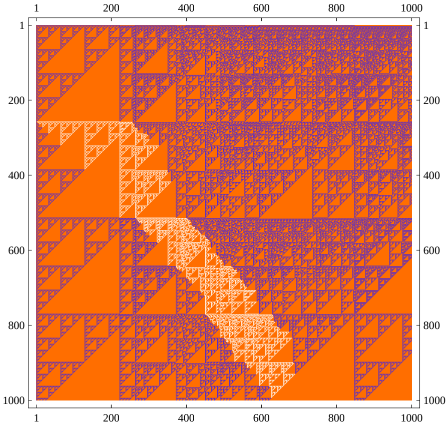

Figure : $p = 0.99$

References

- Collevecchio A., Hamza K. & Shi M. (2015) *Bootstrap Random Walks*, arXiv.org, 1508.02840 21pp.