A problem with two faces: Erdös and Keane

Kari Eloranta

X Lukuteorian päivät/Number theory days
28-29.5.2015
Aalto yliopisto/university

The problem

Definition

Consider the spaces of infinite 1-dimensional sequences of symbols from $S = \{1, 2, 3, \ldots, d\}$ *with an* **exclusion rule***:*

$$X_{(d,f)} = \left\{ x \in S^{\mathbf{Z}} \mid x_i \neq x_{i+f(n)},\ i \in \mathbf{Z},\ n \in \mathbf{N} \right\} \tag{1}$$

where $f : \mathbf{N} \to \mathbf{N}$ *is a strictly increasing function.*

One-sided case $X^{+}_{(d,f)}$: $\mathbf{Z}$ in (1) replaced by $\mathbf{N}$ or $\mathbf{N}_0$.

Mike Keane: What are the sequences in e.g. $X_{(d,n^2)}$ like?

Alternative: Turn integers into a graph G by connecting any two vertices with an edge iff their distance in the numberline equals to $f(n)$ for some n.

Paul Erdös: When is the chromatic number of G finite?

The problem

Definition

Consider the spaces of infinite 1-dimensional sequences of symbols from $S = \{1, 2, 3, \ldots, d\}$ *with an* **exclusion rule***:*

$$X_{(d,f)} = \left\{ x \in S^{\mathbf{Z}} \mid x_i \neq x_{i+f(n)},\ i \in \mathbf{Z},\ n \in \mathbf{N} \right\} \tag{1}$$

where $f : \mathbf{N} \to \mathbf{N}$ *is a strictly increasing function.*

One-sided case $X^{+}_{(d,f)}$: $\mathbf{Z}$ in (1) replaced by $\mathbf{N}$ or $\mathbf{N}_0$.

Mike Keane: What are the sequences in e.g. $X_{(d,n^2)}$ like?

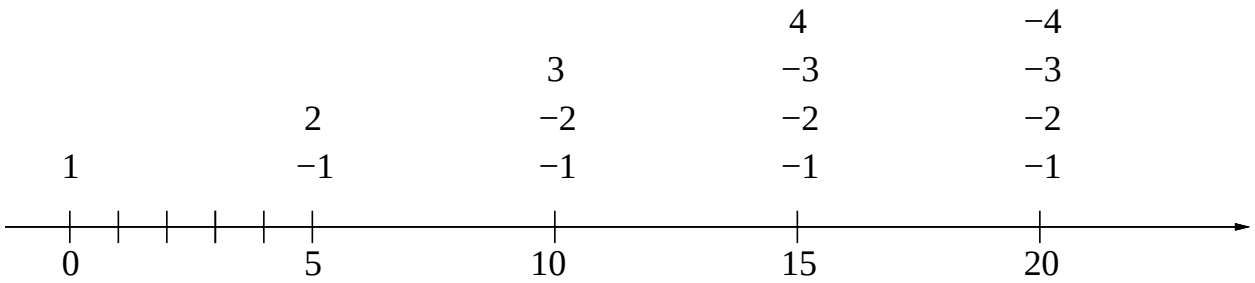
Alternative: Turn integers into a graph G by connecting any two vertices with an edge iff their distance in the numberline equals to $f(n)$ for some n.

Paul Erdös: When is the chromatic number of G finite?

Examples

(1) **(linear growth of f)** Let $S = \{1, 2\}$ and $f(n) = 2n$.
$x_0 = 1$ implies $x_{2k} = 2, \forall k \neq 0$. But $x_2 = 2$ implies $x_{2m} = 1,\ \forall m \neq 1$, a contradiction. So $X_{(2,2n)} = \emptyset$.

In fact $X_{(d,kn)} = \emptyset$ for all $d, k \geq 2$ just by exhausting S. $X_{(d,5n)} = \emptyset$:



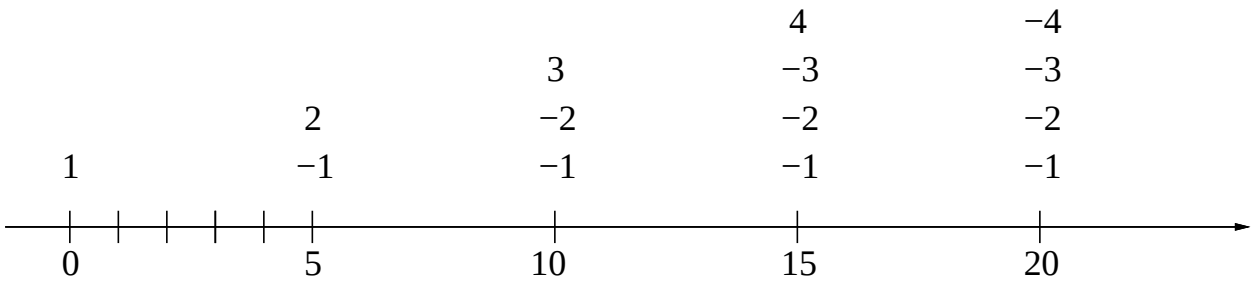
(2) **(linear growth)** For $S = \{1, 2\}$ and $f(n) = 2n - 1$ we have the periodic points $(12)^*$, hence $X_{(2,2n-1)} \neq \emptyset$. Favourable parity!

(3) **(powers)** $S = \{1, 2\}$ and $f(n) = n^r,\ r = 2, 3, \ldots$ If $x_0 = 1$ then $x_{2i} = 1,\ \forall i \in \mathbf{Z}$ so in particular $x_{2^r} = 1$, a contradiction. Therefore $X_{(2,n^r)} = \emptyset$.

Examples

(1) **(linear growth of f)** Let $S = \{1, 2\}$ and $f(n) = 2n$.
$x_0 = 1$ implies $x_{2k} = 2, \forall k \neq 0$. But $x_2 = 2$ implies $x_{2m} = 1,\ \forall m \neq 1$, a contradiction. So $X_{(2,2n)} = \emptyset$.

In fact $X_{(d,kn)} = \emptyset$ for all $d, k \geq 2$ just by exhausting S. $X_{(d,5n)} = \emptyset$:



(2) **(linear growth)** For $S = \{1, 2\}$ and $f(n) = 2n - 1$ we have the periodic points $(12)^*$, hence $X_{(2,2n-1)} \neq \emptyset$. Favourable parity!

(3) **(powers)** $S = \{1, 2\}$ and $f(n) = n^r,\ r = 2, 3, \ldots$ If $x_0 = 1$ then $x_{2i} = 1,\ \forall i \in \mathbf{Z}$ so in particular $x_{2^r} = 1$, a contradiction. Therefore $X_{(2,n^r)} = \emptyset$.

Examples

(4) **(no divisors)** Suppose there is $m \in \mathbf{N}$ which does not divide any of the values $f(n),\ n \in \mathbf{N}$. Then for $d \geq m$ one can have periodic points.
For example $X_{(3,2^n)}$ and $X^{+}_{(4,\{primes\})}$ are nonempty.

(5) **(lots of divisors)** $X^{+}_{(d,n!)}$, the cases $d = 2$ and 3 can easily be checked to be empty. But $X^{+}_{(4,n!)}$ could be non-trivial. There is a period (of length 25) which repeats until the exclusion would violate it for the first time at 5041.

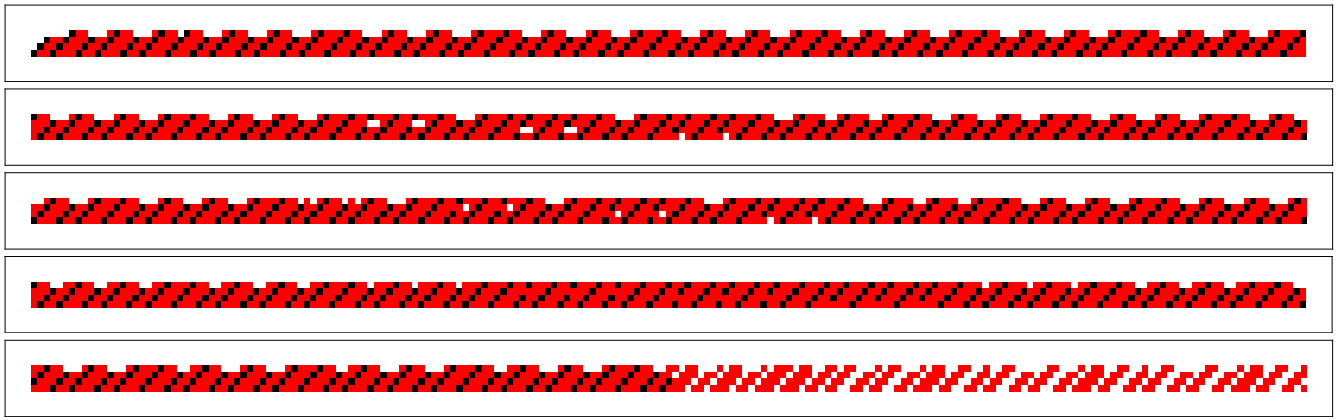


Figure: Lexicographically generated $X^{+}_{(4,n!)}$ (from $x_1 = 1$). Segments, from the top: 1-200, 4950-5150, 10000-10200, 362950-363150, 499900-500100. Periodicity is contradicted around values $n!, n = 7, 10, 11, 12 \ldots$ but clearly the sequence generation survives them for at least half a million steps.

Two results

Around 1990 Y. Katznelson's breakthrough in the lacunary case:

Theorem

If $\inf \frac{f(n+1)}{f(n)} > 1$ *then* $X_{(d,f)} \neq \emptyset$ *for a finite* d.

Proof via toral dynamics. Bounds for d exist (loose).

As for divisibility one can argue:

Theorem

If for any $m \in \mathbf{N}$ *there is* $n \in \mathbf{N}$ *such that* m *divides* $f(n)$, *then the words satisfying the exclusion do not form a context-free language. Hence the sequences do not form a regular language (sofic shift) either.*

Proof by showing that the validity of the appropriate Pumping Lemma is dependent on the (non)divisibility property.

Two results

Around 1990 Y. Katznelson's breakthrough in the lacunary case:

Theorem

If $\inf \frac{f(n+1)}{f(n)} > 1$ *then* $X_{(d,f)} \neq \emptyset$ *for a finite* d.

Proof via toral dynamics. Bounds for d exist (loose).

As for divisibility one can argue:

Theorem

If for any $m \in \mathbf{N}$ *there is* $n \in \mathbf{N}$ *such that* m *divides* $f(n)$, *then the words satisfying the exclusion do not form a context-free language. Hence the sequences do not form a regular language (sofic shift) either.*

Proof by showing that the validity of the appropriate Pumping Lemma is dependent on the (non)divisibility property.

Back to powers

Clearly for $d \geq 3$ and $r \geq 2$ no $X_{(d,n^r)}$ is context free. Moreover

- $X^+_{(3,n^2)} = \emptyset$ (hence also $X_{(3,n^2)} = \emptyset$) by an elementary argument.
- $X^+_{(4,n^2)} = \emptyset$ by a computer assisted proof. Max sequence length is 47.
- For $d = 5$ one can generate sequences of length at least 170.
- Random generation of sequences for $X^+_{(d,n^2)}$, $d = 5, 6, 7, 10, 15$ and 20 suggest strongly that all these spaces are empty...

Additive combinatorics/recurrence

For $A \subset \mathbf{N}$ let $A - A = \{a - a' \mid \text{any } a, a' \in A\}$ and $A^{(N)} = A \cap \{1, 2, \ldots, N\}$.

Question: If we insist that $f(n) \notin A - A$ for any natural n, what is A like?

For n^2 (Lovász's conjecture) Furstenberg and Sárközy showed in 1977-8:

Theorem

Given $\delta > 0$ there is $N_0(\delta)$ such that if $N \geq N_0(\delta)$ and $|A^{(N)}| \geq \delta N$ then there is natural n such that $n^2 \in A - A$.

The proofs were ergodic theoretic and Fourier analytic respectively.

In 1994 Balog, Pelikán, Pintz and Szemerédi proved furthermore

Theorem

For any natural $k \geq 2$ if $n^k \notin A - A$ for all n then $\frac{|A^{(N)}|}{N} \ll \frac{1}{(\log N)^{c \log\log\log\log N}}$.

Additive combinatorics/recurrence

For $A \subset \mathbf{N}$ let $A - A = \{a - a' \mid \text{any } a, a' \in A\}$ and $A^{(N)} = A \cap \{1, 2, \ldots, N\}$.

Question: If we insist that $f(n) \notin A - A$ for any natural n, what is A like?

For n^2 (Lovász's conjecture) Furstenberg and Sárközy showed in 1977-8:

Theorem

Given $\delta > 0$ there is $N_0(\delta)$ such that if $N \geq N_0(\delta)$ and $|A^{(N)}| \geq \delta N$ then there is natural n such that $n^2 \in A - A$.

The proofs were ergodic theoretic and Fourier analytic respectively.

In 1994 Balog, Pelikán, Pintz and Szemerédi proved furthermore

Theorem

For any natural $k \geq 2$ if $n^k \notin A - A$ for all n then $\frac{|A^{(N)}|}{N} \ll \frac{1}{(\log N)^{c \log\log\log\log N}}$.

Corollary

$X^{+}_{(d,n^r)} = \emptyset$, *hence also* $X^{+}_{(d,n^r)} = \emptyset$ *for all* $r \in \mathbf{N}$.

Proof.

Given d symbols, $f(n) = n^k$, let $A_i = \{\{j\} | \, x_j = i\}$ and suppose that $A_i, \ i = 1, \ldots, d$ partition $\mathbf{N}$. Then the last Theorem implies that if the exclusion is to hold for the sequence $\{x_j\}$ for the given f, for sufficiently large n the densities of A_i's cannot add up to 1, a contradiction. □

These result have been extended for **intersective** polynomials: $f \in \mathbf{Z}[x]$ s.t. $f(n) \equiv 0 \pmod{q}$, $\forall q \in \mathbf{N}$. This is exactly the same condition as in the language characterization.

The image $\{f(n) | n \in \mathbf{N}\}$ for such f is a example of a **Poincaré sequence** which in turn is a **recurrence set**.

Finite sequences

Questions: How long sequences are possible e.g. when $f(n) = n^2$?
Termination patterns?

Generating samples from $X^+_{(d,n^2)}$:

Algorithm v2.0:

0. set $M \geq 1$, let $S_j = S$ at each $j \in \{1, \ldots, M\}$ and set $i = 1$.
1. if $S_i = \emptyset$ then **halt**,
   else pick uniformly a random symbol $s \in S_i$.
2. update $S_j \leftarrow S_j \setminus \{s\}$ for all $j = i + f(n) \in \{i+1, \ldots, M\}$, $n \in \mathbf{N}$.
3. if $i = M$ **halt** and call **full length**,

i.e. each coordinate is chosen independently and uniformly but in such a way as to respect the restrictions from all the relevant coordinates in its past.
M is an upper bound for the length, a guess.

Finite sequences

Questions: How long sequences are possible e.g. when $f(n) = n^2$?
Termination patterns?

Generating samples from $X^+_{(d,n^2)}$:

Algorithm v2.0:

0. set $M \geq 1$, let $S_j = S$ at each $j \in \{1, \ldots, M\}$ and set $i = 1$.
1. if $S_i = \emptyset$ then **halt,**
   else pick uniformly a random symbol $s \in S_i$.
2. update $S_j \leftarrow S_j \setminus \{s\}$ for all $j = i + f(n) \in \{i+1, \ldots, M\}$, $n \in \mathbf{N}$.
3. if $i = M$ **halt** and call **full length**,

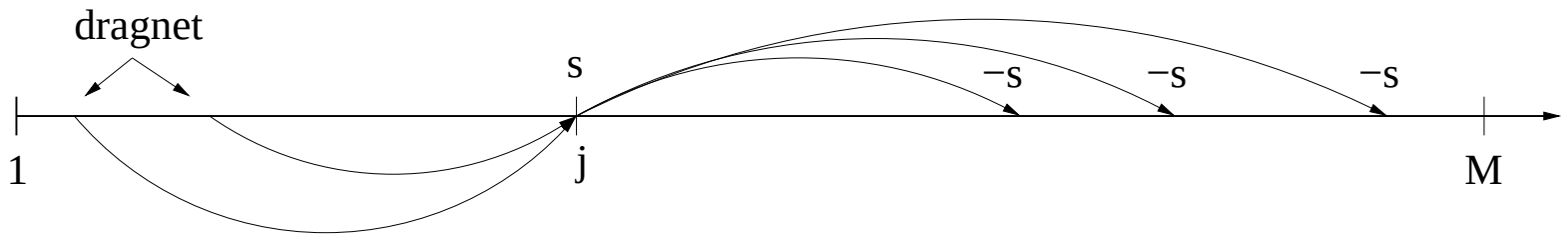
i.e. each coordinate is chosen independently and uniformly but in such a way as to respect the restrictions from all the relevant coordinates in its past.
M is an upper bound for the length, a guess.

Probabilistic model

Dragnet D_j is the set of coordinates less than j restricting the assignment at j. Its cardinality is a step-function, equal to d from the start of the first **interval**.

For $f(n) = n^2$ this is at coordinate $j = d^2 + 1$ and the i^{th} interval is from $(d + i - 1)^2 + 1$ to $(d + i)^2$ (its length is $l_i = 2(d + i) - 1$).

If the sites on the dragnet D_j support the entire alphabet S then at site j there is a **full block**. First full block is possible at the start of the first interval.



Probabilistic model

Assume that all the symbols on $\{1, 2, \ldots, j-1\}$ have been laid out independently and uniformly from S. Then

Proposition

Let B_j be the event that one has the first full block at j in the i^{th} interval. Then

$$\mathbf{Pr}(B_j) = p_i = \frac{1}{d^{d+i-1}} \sum_{\substack{k_r \geq 1,\ r=1,\ldots,d \\ k_1+\cdots+k_d=d+i-1}} \binom{d+i-1}{k_1\ k_2\ \ldots\ k_d} \tag{2}$$

where the sum is d-fold over the given positive integers.

Recall the multinomial: $\binom{a}{b_1\ b_2\ \ldots\ b_d} = \frac{a!}{b_1!b_2!\cdots b_d!}$, $\sum_{i=1}^{d} b_i = a$.
Proof is combinatorics on the dragnet.

Probabilistic model

On the interval with dragnet cardinality $d + i - 1$ the sequence extension halts w.p. p_i and its length on the interval $\sim Geom(p_i)$.

Lemma

For an alphabet S of size d one has for all $i \geq 1$

$$1 - p_i < d \left(1 - \frac{1}{d}\right)^{d-1} \left(1 - \frac{1}{d}\right)^{i}.$$

For the proof of the Lemma one has to consider the entries on the $(d + i - 1)^{th}$ level (from the top) of Pascal's d-simplex. Multinomial Theorem gives the total sum but for $1 - p_i$ we need to bound its boundary sum. Note that $p_i \uparrow 1$ is obvious, but its geometric lower bound requires some work.

Probabilistic model

Theorem

Let the assumptions on the sequence be as above and $f(n) = n^2$. Then a full block materializes at j i.e. **Pr**(*sequence generation halts at* j) $=$

$$\begin{cases} 0 & 1 \leq j \leq d^2 \\ (1-p_1)^{[\,j-d^2-1]}\,p_1 & j \text{ in the first interval} \\ \left(\prod_{k=1}^{i-1}(1-p_k)^{l_k}\right)(1-p_i)^{[\,j-d^2-1-\sum_{k=1}^{i-1} l_k]}p_i & j \text{ in the } i^{th} \text{ interval, } i \geq 2 \end{cases}$$

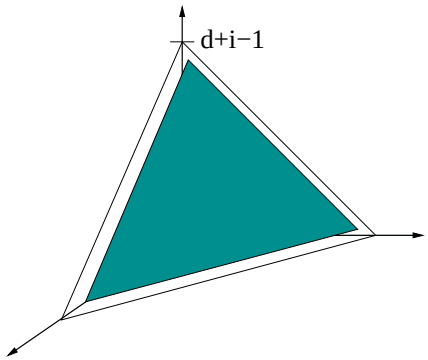
and the halting time distribution has a geometric tail.
The sequences generated are almost surely of finite length.

The Theorem follows by combining the geometric halting probabilities on the intervals, "uniformizing" them for a tail estimate (l_k are not equal) and finally using Borel-Cantelli.

Remarks

- The sum

$$\sum_{\substack{k_r \geq 1,\ r=1,\ldots,d \\ k_1+\cdots+k_d=d+i-1}} \binom{d+i-1}{k_1\ k_2\ \ldots\ k_d}$$



has asymptotically an exponential number of summands both in d and i. To use the Theorem for large d and i one needs to find an efficient way to compute the p_i's.

- While $p_i \uparrow 1$ monotonically, the halting distribution is jagged: At the i^{th} jump

$$\frac{\mathbf{Pr}(\textit{halts at } (d+i)^2+1)}{\mathbf{Pr}(\textit{halts at } (d+i)^2)} = \frac{1-p_i}{p_i} p_{i+1} \to 0 \quad \textit{as} \quad i \to \infty$$

but far exceeds 1 earlier.

- The model applies verbatim to other jump sequences $f(n)$.

Reality check for n^2 case

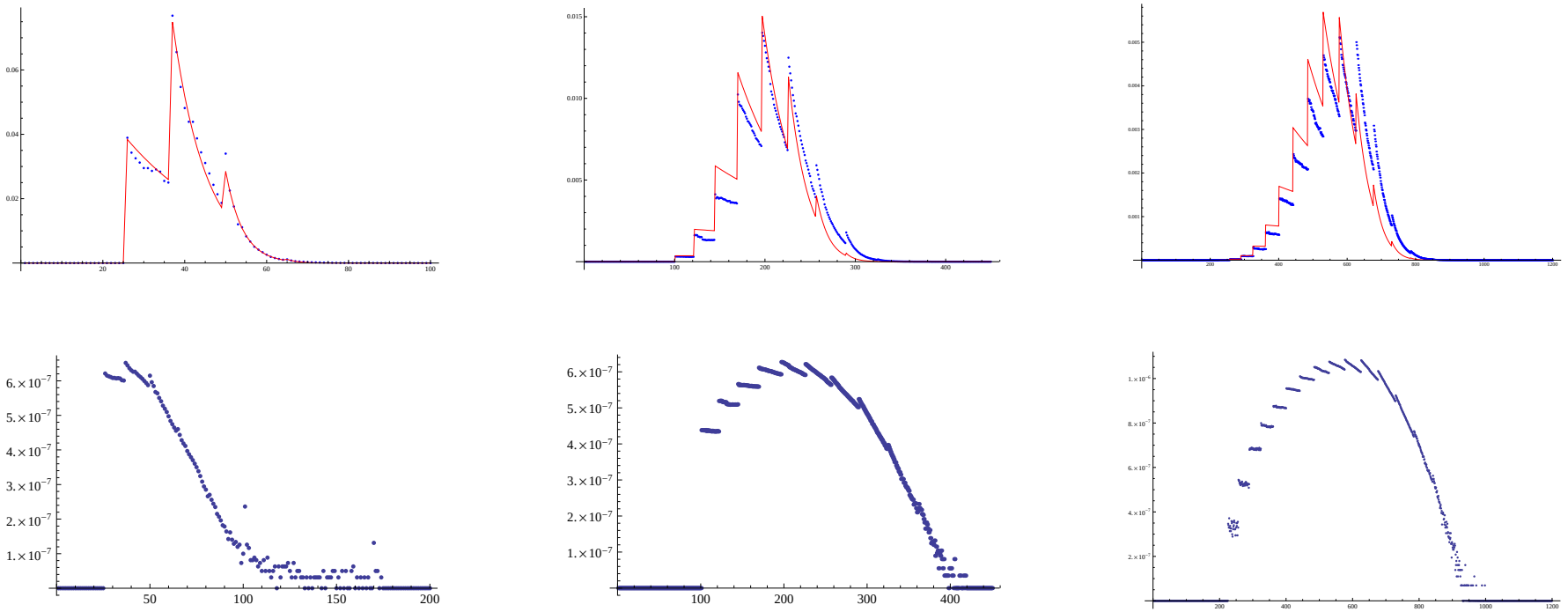


Figure: Top row: empirical (blue) and model (red) halting probability distributions. Bottom row: log of the blue data above. Columns left to right: $d = 5, 10$ and 15.

Dependencies, squared

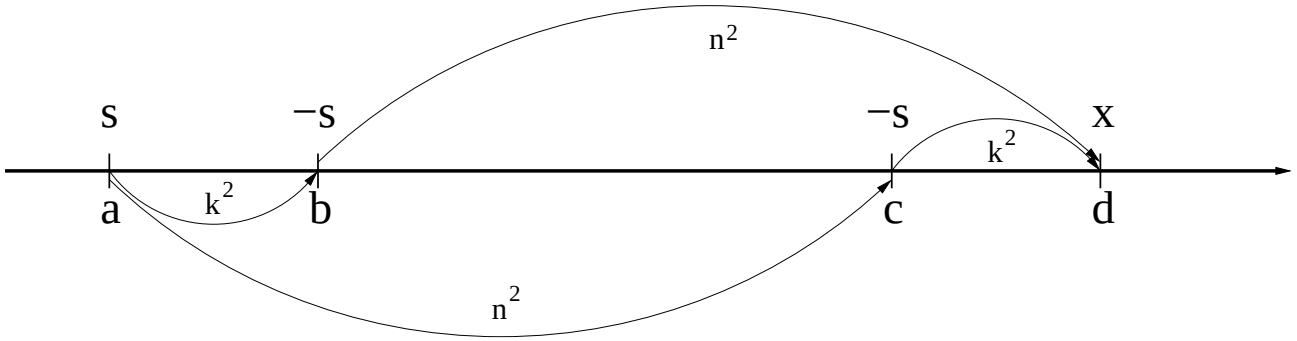


Figure: A dependency mechanism affecting the termination probability. Dragnet is at d.

$$\mathbf{Pr}\left(seq.\ term.\ at\ d \mid k^2 + n^2 \neq m^2\right) < \mathbf{Pr}\left(model\ term.\ at\ d \mid k^2 + n^2 \neq m^2\right)$$

$$\mathbf{Pr}\left(seq.\ term.\ at\ d \mid k^2 + n^2 = m^2\right) > \mathbf{Pr}\left(model\ term.\ at\ d \mid k^2 + n^2 = m^2\right)$$

For exact analysis one would need to account the Pythagorean triples. However as k and n vary, the non-triples case is far more likely to occur than the triples case. So termination probabilities of the independent model should major the observed ones.

Statistics of the sequence lengths for n^2

Symbols d	Empirical mean	Empirical std. dev.	Model mean	Model std. dev.	Sequences
4	27.2542	5.13374	23.992	5.23924	$50 \cdot 10^6$
5	39.5672	8.28983	39.2172	8.22516	$80 \cdot 10^6$
6	60.8247	13.5813	59.3666	11.9713	$80 \cdot 10^6$
7	89.4687	18.5912	84.982	16.5113	$30 \cdot 10^6$
10	209.315	38.2887	199.562	35.1369	$20 \cdot 10^6$
15	566.87	92.2796	543.291	84.4349	$10 \cdot 10^6$
20	1156.57	170.829	$*$	$*$	$5 \cdot 10^6$

Table: Data from randomly generated one-sided sequences and the probabilistic model. Asterisks are due to missing coefficients (for i large).

$d \to \infty$ limiting behavior?

Based on the above one might venture to...

Conjecture

Consider the case of d symbols and $f(n) = n^2$. Suppose $T^{(d)}$ is the halting instant of the Algorithm v2.0. For sufficiently rapidly growing $M(d)$ there are positive constants a and b such that as $d \to \infty$

$$\mathbf{Pr}\left(\frac{T^{(d)} - ad^{5/2}}{bd^{15/7}} \leq x\right) \longrightarrow \Phi(x) \qquad \forall x \in \mathbf{R}$$

where Φ is the cumulative distribution function of the standard normal $N(0,1)$.

Such CLT should hold for the probabilistic model as well (with parameters but not exponents adjusted)

$M(d)$ just needs to outgrow the off-set rate $d^{5/2}$.

Termination details for n^2

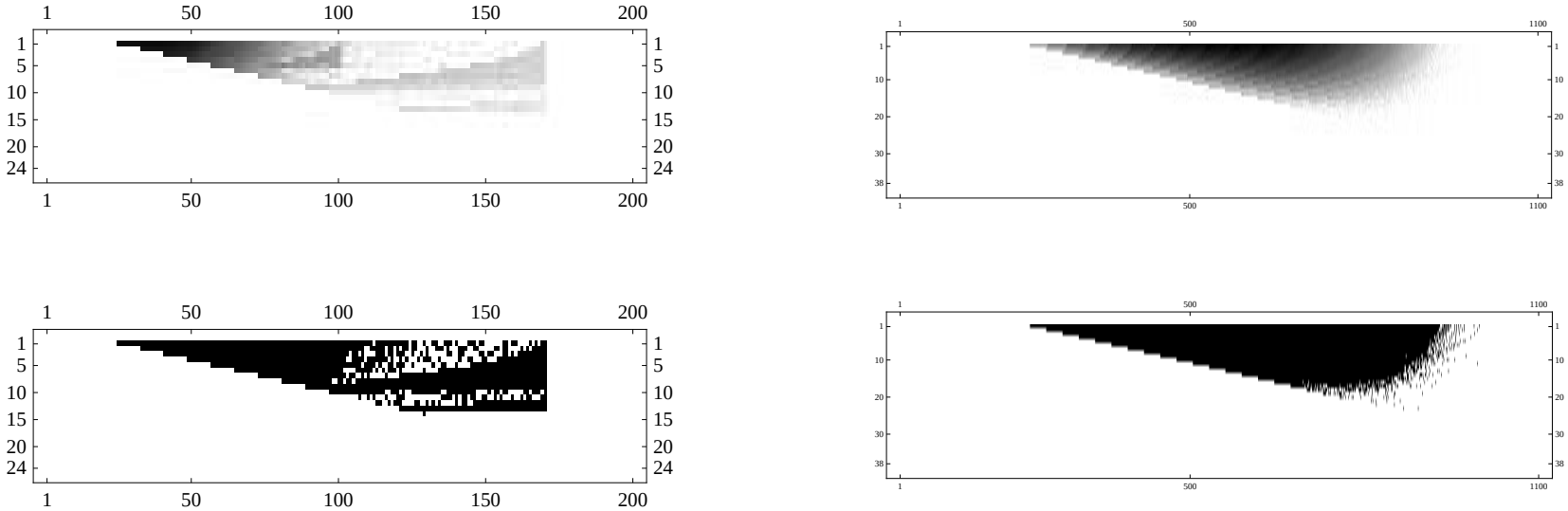


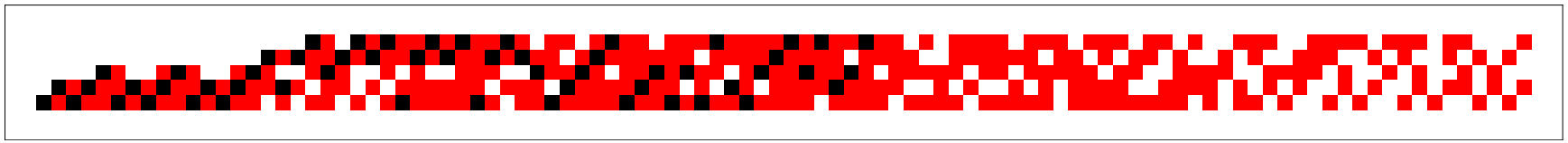
Figure: Distribution of the instant when the upcoming termination can be seen for the first time (x) and how far ahead it will be (y). Log and sign, top and bottom resp. for $d = 5$ and 15 (20 and 10 million samples). No boundary effect from the right (actual horizon M used much higher).

Algorithm termination seems to turn into a random process as d increases.

Rest

arXiv:math-ph/1204.3439
or
www.math.aalto.fi/∼kve/research.html

Thank you!



A **context-free language** is recognized by a non-deterministic pushdown automaton (i.e. has *one* stack). Such language satisfies a **Pumping Lemma**:

Lemma

Any sufficiently long string s, say $|s| \geq k$, can be written as $s = uvxyz$ such that
(i) $|vxy| \leq k$,
(ii) $|vy| \geq 1$,
(iii) $uv^n xy^n z$ is an allowed string for all natural n.

If either v or y vanishes but the other is non-trivial (hence (ii) is still valid) this reduces to the Pumping Lemma of **regular languages** (languages recognized by a finite state automaton).