Derandomizing Arthur-Merlin games using hitting sets

Peter Bro Miltersen,

BRICS,

University of Aarhus.

Talk based on joint work with

N.V. Vinodchandran,
BRICS, University of Aarhus

appearing at FOCS'99

This talk

1. The result
2. The proof technique
   (a) Andreev-Clementi-Rolim hashing based hitting set generator.
   (b) Our improvement:
   Hashing $\rightarrow$ Error correcting code + projection

The Impagliazzo-Wigderson Theorem (1997)

$$\exists L \in \mathbf{E}, \text{CircuitSize}(L \cap \{0,1\}^n) \geq 2^{\epsilon n}$$

$$\Downarrow$$

$$\mathbf{P} = \mathbf{BPP}$$

The I-W Theorem, paraphrased

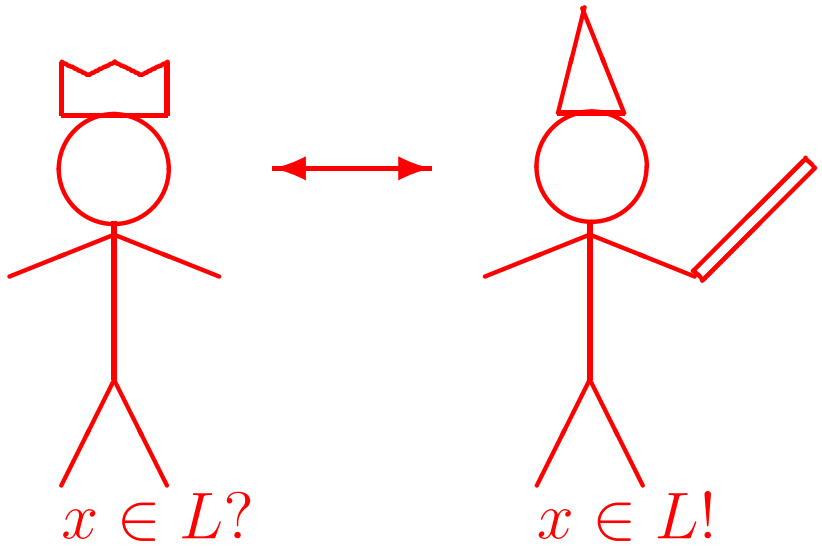
Randomization *never* speeds up computation

unless

Non-uniformity *always* speeds up computation

Circuit size $\approx$ deterministic time.

Interactive proof systems



$$\mathbf{DIP} = \mathbf{NP}$$

IP = PSPACE

$$\mathbf{NP} \subseteq \mathbf{MA} \subseteq \mathbf{AM} \subseteq \mathbf{IP{=}PSPACE}$$

Goldreich-Zuckerman (1997)

$$\exists L \in \mathbf{E}, \text{CircuitSize}(L \cap \{0,1\}^n) \geq 2^{\epsilon n}$$

$$\Downarrow$$

$$\mathbf{MA} = \mathbf{NP}$$

Klivans-Van Melkebeek (1999)

$$\exists L \in \mathbf{NE} \cap \mathbf{coNE}, \text{SAT-CircuitSize}(L \cap \{0,1\}^n) \geq 2^{\epsilon n}$$

$$\Downarrow$$

$$\mathbf{AM} = \mathbf{NP}$$

Klivans-Van Melkebeek, attempted paraphrase

Randomization *never* speeds up Arthur-Merlin games

unless

Non-uniformity *always* speeds up strong nondeterministic computation

But: SAT-circuit size $\not\approx$ strong nondeterministic time.

Our result

$$\exists L \in \mathbf{NE} \cap \mathbf{coNE}, \mathrm{SVcircuitSize}(L \cap \{0,1\}^n) \geq 2^{\epsilon n}$$

$$\Downarrow$$

$$\mathbf{AM} = \mathbf{NP}$$

SV-circuit for f =

nondeterministic circuit for f *and*

nondeterministic circuit for $\neg f$.

Our result, paraphrased

Randomization *never* speeds up Arthur-Merlin games

unless

Non-uniformity *always* speeds up strong nondeterministic computation

SV-circuit size $\approx$ strong nondeterministic time.

BPP, MA, AM

To derandomize	we want languauge in	with no small
BPP	**E**	Circuits
MA	$\mathbf{NE} \cap \mathbf{coNE}$	Circuits
AM	$\mathbf{NE} \cap \mathbf{coNE}$	SV-Circuits

Inside the Impagliazzo-Wigderson proof

$\exists$ efficient procedure with
Input: n
Output: Truthtable of $f : \{0,1\}^n \to \{0,1\}$ with
circuit size at least (maximum possible)$^\epsilon$

$\Downarrow$ lots of work

Pseudorandom generators exist

$\Downarrow$ immediate

P = BPP

Pseudorandom generators

$S \subset \{0,1\}^n$ is *pseudorandom set*

if

$$\Pr_{x \in S}[C(x) = 1] \approx \Pr_{x \in \{0,1\}^n}[C(x) = 1]$$

for all small circuits C

Andreev-Clementi-Rolim 1997

$\exists$ efficient procedure with
Input: n, m
Output: Truthtable of $f : \{0,1\}^n \to \{0,1\}^m$
with *maximum* circuit size

$\Downarrow$

Hitting set generators exist

$\Downarrow$ immediate $\Downarrow$ ACR'95

P = RP **P = BPP**

easy!!

Hitting set generators

$H \subset \{0,1\}^n$ is *hitting set*

if

$$\Pr_{x \in \{0,1\}^n}[C(x) = 1] > 1 - n^{-3} \Rightarrow \Pr_{x \in H}[C(x) = 1] > 0$$

for all small circuits C

$1 - n^{-3}$ is the *threshold* of the set.

$f : \{0,1\}^{2\log n} \to \{0,1\}^n$ has max circuit complexity

$\Rightarrow S = \text{Image}(f)$ is a hitting set.

Assume to the contrary S does not hit C.

$Z = \{z | C(z) = 0\}$. $|Z| \leq 2^n/n^3,\ S \subseteq Z$

Universal family $\mathcal{H} \ni h : \{0,1\}^n \to \{0,1\}^{n-1}$

$x \neq y \Rightarrow \Pr[h(x) = h(y)] = 2^{-n-1}$

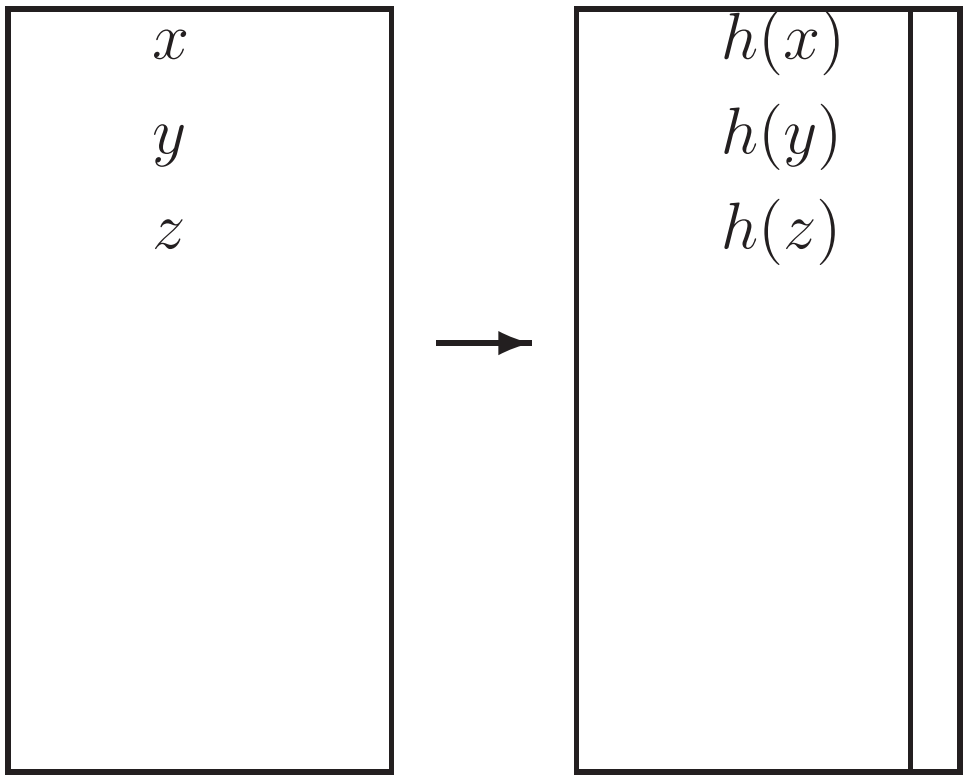
$\exists h \in \mathcal{H} : x \in S, y \in Z, x \neq y \Rightarrow h(x) \neq h(y)$

Small circuit for f

Advice: C, h, Table of $x \rightarrow h(f(x))$.

On input x, inspect $h(f(x))$. There are two possiblilities for $f(x)$. Take the one making C zero.

This contradicts f having maximum possible circuit complexity.


x
y
z
$h(x)$
$h(y)$
$h(z)$

Why only one bit compression?

1. We are making a "sensitive" hitting set with fairly small threshold.

2. We have to search through all possible preimages of $h(f(x))$

Solution: Deterministic amplification using dispersers (done by ACR).

Solution: Switch to hardness assumption in terms of *nondeterministic* circuits
(almost, but not quite, done by ACR).

2nd Andreev-Clementi-Rolim theorem

Different assumption

$\Downarrow$ same as before

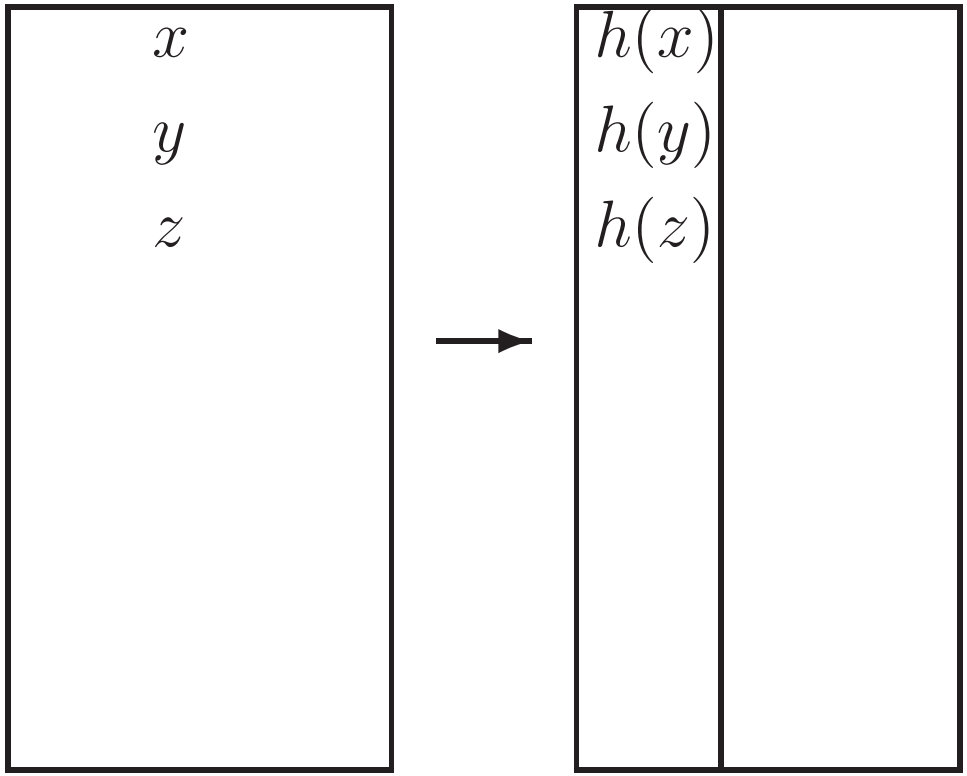
Hitting set generator with very big threshold exists

$\Downarrow$ explicit dispersers

Hitting set generators exist

$\Downarrow$ immediate $\quad$ $\Downarrow$ (ACR'95)

P = RP $\quad$ **P = BPP**


x
y
z
$h(x)$
$h(y)$
$h(z)$

Without new ideas, we have

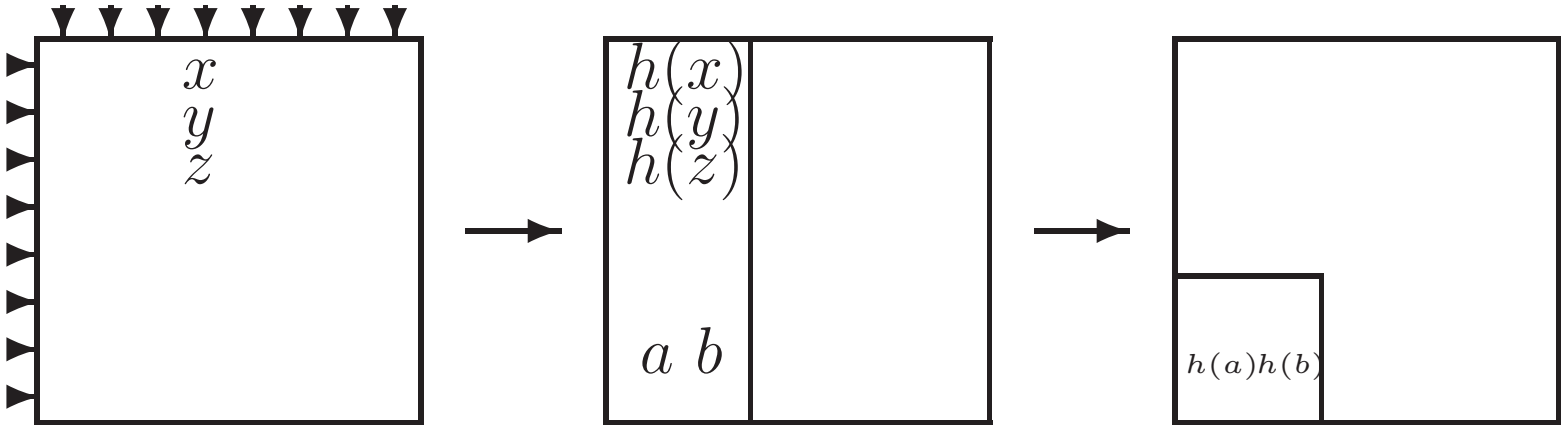
$$\exists L \in \mathbf{NE} \cap \mathbf{coNE}, \mathrm{SVcircuitSize}(L \cap \{0,1\}^n) \geq 2^{0.99n}$$

$$\Downarrow$$

$$\mathbf{AM} = \mathbf{NP}$$

Improvement idea

Show that the set of rows and *columns* is a hitting set



Obstacle

Hashing *horisontally* destroys the structure of the data when viewed *vertically*.

How can we hash the rows of a matrix without destroying the columns?

Answer

Hashing $\rightarrow$ error correcting code + projection.

$$x \in \{0,1\}^n \mapsto h(x) \in \{0,1\}^m$$

$$\downarrow$$

$$x \in \{0,1\}^n \mapsto e(x) \in \{0,1\}^{n'} \mapsto \pi(e(x)) \in \{0,1\}^m$$

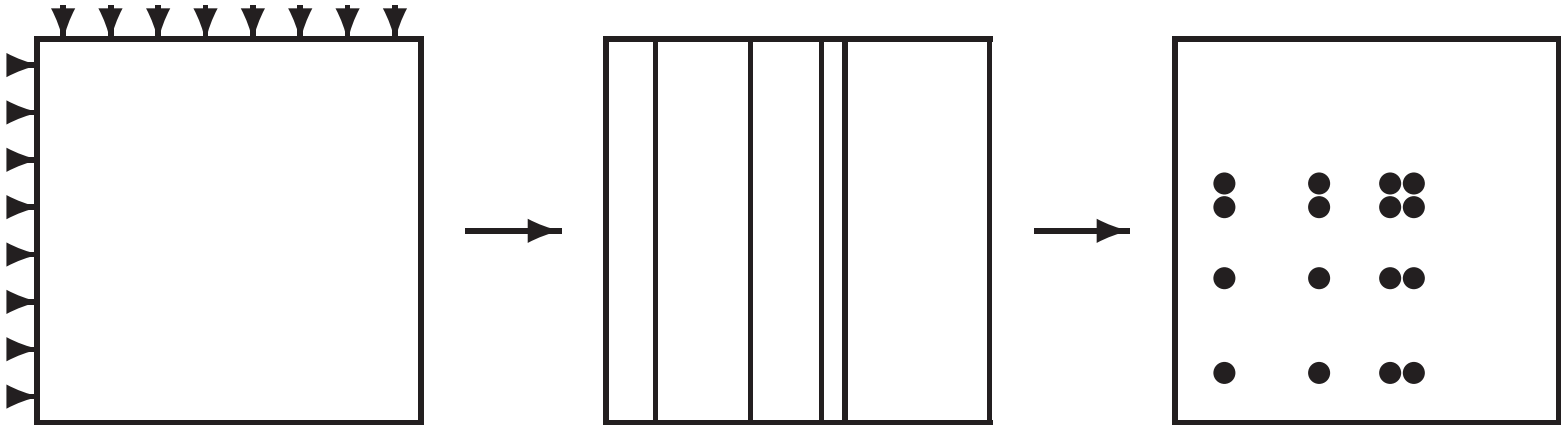
General useful trick. Other applications.

Deterministic Dictionaries
(Miltersen 1998, Hagerup 1999, Pagh 2000)

Perfect hashing circuits (Miltersen 1998)

Shannon function for partially defined Boolean functions
(Miltersen 1999)

Rows and columns are encoded in error correcting code



Let the projection π be to the set of indices S

Small circuit for $f(i,j)$:

Advice: C, π_S, Table of $f_{|S\times S}$.

For each $s \in S$, guess and verify $f(*,s)$. Keep $f(i,s)$.

We now have $f(i,S)$.

Guess and verify $f(i,*)$. Keep $f(i,j)$.

Multidimensional version

Wanted: Encoding of truth table as k-dimensional bit-array so

1. The orginal truth table is in there
2. Blowup is small.
3. Every 1-dimensional slice of the bit-array is a codeword in an error correcting code.

A k-cubic code

Low degree extension is k-cubic code over $\mathbf{F}$

Message: $f \in \mathbf{F}^N$

$f : S^k \to \mathbf{F},\ S \subset \mathbf{F},\ |\mathbf{F}| \approx |S|^2$

Code $\tilde{f} : \mathbf{F}^k \to \mathbf{F}$.

$\tilde{f}$: The unique polynomial of degree at most $|S| - 1$ in *each* variable agreeing with f on S^k.

Recipe: Hitting set for **AM**

Ingredient: T, truth table of worst case hard function

1. Take k-dimensional low degree extension $\tilde{T}$ of T over $\mathbf{F}$
2. Enumerate all one-dimensional slices of $\tilde{T}$.
3. Replace each field element with Boolean encoding.
4. Put the resulting set through disperser.

Curious fact

First step of Impagliazzo-Wigderson proof is *also* to take the low degree extension of the hard function.

Property used there: It is a *locally decodable* code.

Property used here: It is a *k-cubic* code.

Open problem

Can we show:

$$\exists L \in \mathbf{NE} \cap \mathbf{coNE}, \mathrm{SVcircuitSize}(L \cap \{0,1\}^n) \geq n^{\omega(1)}$$

$$\Downarrow$$

$$\mathbf{AM} \subseteq \mathbf{NSUBEXP}$$

Best using compression technique

$$\exists L \in \mathbf{NE} \cap \mathbf{coNE}, \mathrm{SVcircuitSize}(L \cap \{0,1\}^n) \geq 2^{n^{\frac{1}{2}+\epsilon}}$$

$$\Downarrow$$

$$\mathbf{AM} \subseteq \mathbf{NTIME}(2^{\mathrm{polylog}n})$$