Shallow Circuits with High-Powered Inputs

Pascal Koiran

LIP, Ecole Normale Supérieure de Lyon

and

Department of Computer Science, University of Toronto

Workshop on Randomization, Relaxation and Complexity.

Banff International Research Station, March 2010.

Two central problems of complexity theory

1. Arithmetic complexity of the permanent
   (Valiant's algebraic version of P versus NP).

2. Derandomization of Polynomial Identity Testing.

- Problems turn out to be related.

- Progress on one may lead to progress on other problem
  (approach to problem 1 advocated by Agrawal, 2005).

Valiant's model: $\mathsf{VP}_K = \mathsf{VNP}_K$?

- Complexity of a polynomial f measured by number $L(f)$ of arithmetic operations (+,-,$\times$) needed to evaluate f:

L(f) = size of smallest arithmetic circuit computing f.

- $(f_n) \in \mathsf{VP}$ if number of variables, $\deg(f_n)$ and $L(f_n)$ are polynomially bounded. For instance, $(X^{2^n}) \notin \mathsf{VP}$.
- $(f_n) \in \mathsf{VNP}$ if $f_n(\overline{x}) = \sum_{\overline{y}} g_n(\overline{x}, \overline{y})$

  for some $(g_n) \in \mathsf{VP}$

  (sum ranges over all boolean values of $\overline{y}$).

  If $\mathrm{char}(K) \neq 2$ the permanent is a VNP-complete family:

$$\mathrm{PER}_n(X) = \sum_{\sigma \in S_n} \prod_{i=1}^{n} X_{i\sigma(i)}.$$

Constant-free version of Valiant's model

- Work with constant-free circuits (1 is the only constant).
- $(f_n) \in \mathsf{VP}^0$ if size and *formal degree* of circuits are polynomially bounded (Malod, 2003).

  Formal degree is an upper bound on $\deg(f_n)$:

  1. 1 for an input gate (variable or constant).
  2. Max of formal degrees of two inputs for $+, -$ gate.
  3. Sum of formal degrees for $\times$ gate.
- New goal: $\mathrm{PER}(X) \notin \mathsf{VP}^0$.

Polynomial Identity Testing

Given polynomial f, decide whether $f \equiv 0$.

If given by an arithmetic circuit: ACIT problem.

Schwartz-Zippel-DeMillo-Lipton lemma:
Let $f \in K[X_1, \ldots, X_n]$ of degree d.
If $f \not\equiv 0$ and $X_1, \ldots, X_n$ drawn independently at random from $S \subseteq K$:

$$\Pr[f(X_1, \ldots, X_n) = 0] \leq d/|S|.$$

"Natural" intuition about ACIT:
no efficient deterministic algorithm exists
(because we haven't found any).

Hardness versus randomness tradeoffs

Two roughly equivalent problems:

- derandomizing algorithms
- proving lower bounds.

For each problem we need **explicit constructions**.

From Kabanets-Impagliazzo (2004) :

- If ACIT can be derandomized:
  we have a lower bound for the permanent, or $\mathsf{NEXP} \not\subset \mathsf{P/poly}$.
- If we have a lower bound for the permanent:
  ACIT can be derandomized in subexponential time
  for circuits of logarithmic depth.

A possible approach to arithmetic circuit lower bounds ?
(Agrawal, 2005)

The black-box model

Only way to access f:

$$x \mapsto \boxed{\textbf{black box}} \to f(x).$$

Some problems studied in this model:
factorization, GCD, interpolation...

Two equivalent problems:

- derandomization of PIT in the black blox model.
- Construction of a *hitting set*.

A hitting set H for a family $\mathcal{F}$ of polynomials must contain a point x such that $f(x) \neq 0$ for every $f \not\equiv 0$ in $\mathcal{F}$.

Remark: Hitting sets $\not\Rightarrow$ derandomization in circuit model.

Hitting sets for sparse polynomials

- For the set of polynomials $f \in \mathbb{R}[X]$ with at most t monomials: any set $H \subseteq \mathbb{R}_+^*$ with $|H| = t$ is a hitting set
  Proof: apply Descarte's rule of signs.

- For the set polynomials $f \in \mathbb{C}[X]$ with at most t monomials, of degree at most d:
  let H be the set of all p-th roots of unity for all $p \in \mathcal{P}$,
  where $\mathcal{P}$ is a set of at least $t \log d$ prime numbers.
  Proof: If $f = 0$ on H then $f \equiv 0 \bmod (X^p - 1)$ for all $p \in \mathcal{P}$.
  Fix monomial $a_i X^{\alpha_i}$ in f.
  Then $p | (\alpha_j - \alpha_i)$ for some other monomial $a_j X^{\alpha_j}$.

Existence of hitting sets

Recall from Schwartz-Zippel lemma:

$$\Pr[f(X_1, \ldots, X_n) = 0] \leq 1/2$$

if $|S| \geq 2d$.

Let H = m random elements of S^n.
For $f \not\equiv 0$, $\Pr[f \equiv 0 \text{ on } H] \leq 1/2^m$.
Let $\mathcal{F}$ be a family of polynomials.
By union bound, H is *not* a hitting set with probability $\leq |\mathcal{F}|/2^m$:
take $m > \log |\mathcal{F}|$.

Remarks: same proof as $\mathsf{RP} \subseteq \mathsf{P}/\mathsf{poly}$ (Adleman, 1978);
good bounds also for some infinite families $\mathcal{F}$ (Heintz-Schnorr, 1980).

Lower bounds from (univariate) hitting sets

Let $H = \{a_1, \dots, a_k\}$ be a hitting set for $\mathcal{F}$, and

$$f(X) = \prod_{i=1}^{k} (X - a_i).$$

Then $f \notin \mathcal{F}$.

If H is explicit then f is explicit too!

Remarks:

1. This is a kind of indirect diagonalization.
2. Argument appears already in Heintz and Schnorr (1980).
3. Low-degree multivariate version in Agrawal (2005).
4. Our results are based on the univariate version.

Lower bounds for SPS polynomials

Main Theorem (informal statement):
Efficient deterministic constructions of hitting sets for sums of products of sparse polynomials imply that the permanent is not in VP^0.

SPS polynomials are of the form $f(X) = \sum_{i=1}^{k} \prod_{j=1}^{m} f_{ij}(X)$ where the f_{ij} are t-sparse.

We have seen efficient constructions for sparse polynomials, and products thereof (Descarte's rule).

Benefits of univariate method:

1. Would lead to lower bounds for the permanent, instead of polynomials with PSPACE coefficients (i.e., in VPSPACE).

2. Leads to refinements of Shub and Smale's τ-conjecture.

The τ-conjecture

For $f \in \mathbb{Z}[X_1, \ldots, X_n]$,
$\tau(f) =$ constant-free arithmetic circuit complexity of f.

Remark: If $(f_n) \in \mathsf{VP}^0$ then $\tau(f_n) \leq n^{O(1)}$;
converse not always true (take $f_n = X^{2^n}$ or $f_n = 2^{2^n}$).

For $f \in \mathbb{Z}[X]$, say that $f \in \mathcal{F}_\tau$ if $\tau(f) \leq \tau$.

Conjecture: Any nonzero $f \in \mathcal{F}_\tau$ has at most $p(\tau)$ integer roots, for some fixed polynomial p.

Theorem (Shub - Smale, 1995):
The τ-conjecture implies $\mathsf{P}_\mathbb{C} \neq \mathsf{NP}_\mathbb{C}$.

Two other consequences of the τ-conjecture

1. Hitting set $\{1, 2, 3, \ldots, p(\tau) + 1\}$ for $\mathcal{F}_\tau$.
2. $\tau(\mathrm{PER}_n)$ is not polynomially bounded in n (Bürgisser, 2007): otherwise, $\prod_{i=1}^{2^n}(X - i)$ would have polynomially bounded τ.

Our main theorem in this special case (initial segments of $\mathbb{N}$):
similar statement for SPS polynomials,
instead of arbitrary arithmetic circuits:
If poly-size initial segments of $\mathbb{N}$ *form hitting sets for SPS polynomials, then permanent is not in* VP^0.

More precisely...

τ-conjecture for SPS polynomials

Consider $f(X) = \sum_{i=1}^{k} \prod_{j=1}^{m} f_{ij}(X)$ where the f_{ij} are t-sparse.
Let $\mathrm{size}(f) =$ number of monomials in this expression ($\leq kmt$).

Definition: $f \in \mathrm{SPS}_{s,e}$ if $\mathrm{size}(f) \leq s$, $\deg(f_{ij}) \leq e$,
and each integer coefficient of each f_{ij}:

(i) is of absolute value at most 2^e;

(ii) has $\leq s$ nonzero digits in its binary representation
(f_{ij} is a sparse polynomial with sparse coefficients).

Conjecture 1: If $f \in \mathrm{SPS}_{s,e}$ is nonzero,
f has at most $(s + \log e)^{O(1)}$ integer roots.

Remark: follows from the τ-conjecture since $\tau(f)$ is $(s + \log e)^{O(1)}$.

Theorem: Conjecture 1 implies that the permanent is not in VP^0.

τ-conjecture for SPS polynomials, strong from

Recall **Conjecture 1:** If $f \in \mathrm{SPS}_{s,e}$ is nonzero,
f has at most $(s + \log e)^{O(1)}$ integer roots.

We have a degree bound, sparse and bounded coefficients...
Are these things really relevant ??

Conjecture 2: Consider $f(X) = \sum_{i=1}^{k} \prod_{j=1}^{m} f_{ij}(X)$,
where the f_{ij} are t-sparse.
If f is nonzero, its number of integer roots is polynomial in kmt.

Remark: implies Conjecture 1 since $s \leq kmt$;
does not seem to follow from Shub and Smale's τ-conjecture.

There is an even wilder conjecture...

Real τ-conjecture

Conjecture 3: Consider $f(X) = \sum_{i=1}^{k} \prod_{j=1}^{m} f_{ij}(X)$,
where the f_{ij} are t-sparse.
If f is nonzero, its number of **real roots** is polynomial in kmt.

Remark: obvious for $k = 1$, open for $k = 2$;
could techniques from real analysis show that $\mathrm{PER} \notin \mathsf{VP}^0$?

If true, property would be specific to SPS polynomials:
Shub and Smale have observed that in general,
the number of real roots can be exponential in $\tau(f)$.

Chebyshev polynomials

- Let T_n be the Chebyshev polynomial of order n:

$$\cos(n\theta) = T_n(\cos\theta).$$

For instance $T_1(x) = x$, $T_2(x) = 2x^2 - 1$.

- T_n is a degree n polynomial with n real zeros on $[-1, 1]$.
- $T_{2^n}(x) = T_2(T_2(\cdots T_2(T_2(x))\cdots))$: n-th iterate of T_2.
  As a result $\tau(T_{2^n}) = O(n)$.

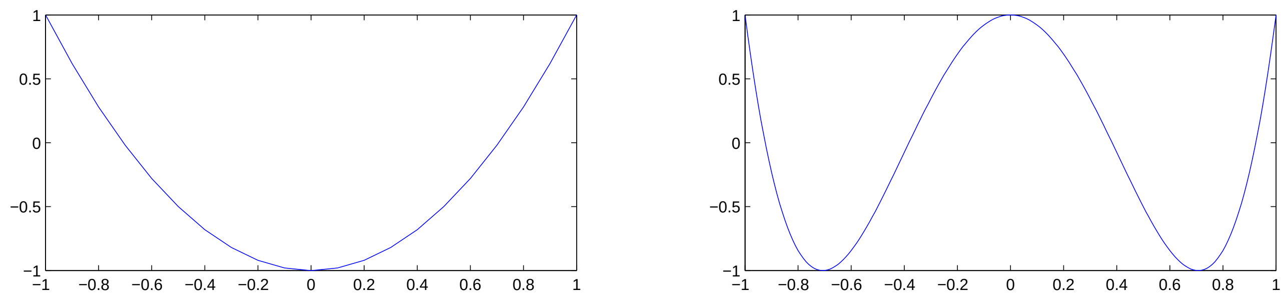

Figure 1: Plots of T_2 and T_4

A new ingredient: the chasm at depth 4

Depth reduction theorem (Agrawal and Vinay, 2008):
Any multilinear polynomial in n variables which has an arithmetic circuit of size $2^{o(n)}$ also has a depth-4 arithmetic circuit of size $2^{o(n)}$.

Remarks:

1. Depth-4 circuit $\equiv \Sigma\Pi\Sigma\Pi$ arithmetic formula;

2. Our polynomials are far from multilinear, but:

Depth-4 circuit with inputs of the form X^{2^i} or 2^{2^j} (*Shallow circuit with high-powered inputs*)

$\Updownarrow$

Sum of Products of Sparse Polynomials

Proof sketch (1/4)

Goal: If $\mathrm{PER} \in \mathsf{VP}^0$ then SPS polynomials of size $2^{o(n)}$ can compute multiples of $\prod_{i=1}^{2^n-1}(X+i)$.

Definition: A polynomial family (f_n) is in VNP^0 if for some family $(g_n) \in \mathsf{VP}^0$:

$$f_n(\overline{x}) = \sum_{\overline{y} \text{ boolean}} g_n(\overline{x}, \overline{y}).$$

Valiant's criterion: Let

$$f_n(x_1, \ldots, x_{p(n)}) = \sum_{i=0}^{2^{p(n)}-1} a_n(i) x_1^{i_1} \cdots x_{p(n)}^{i_{p(n)}}.$$

If $a : (1^n, i) \mapsto a_n(i) \in \{0, 1\}$ is in $\mathsf{P/poly}$ then $(f_n) \in \mathsf{VNP}^0$.

Proof sketch (2/4)

The counting hierarchy: $\mathsf{C}_0\mathsf{P} = \mathsf{P}$; $\mathsf{C}_1\mathsf{P} = \mathsf{PP}$ where $A \in \mathsf{PP}$ iff there exists a polynomial p and $B \in \mathsf{P}$ such that for x of length n:

$$x \in A \Leftrightarrow |\{y \in \{0,1\}^{p(n)};\ \langle x, y\rangle \in B\}| > 2^{p(n)-1}.$$

$\mathsf{C}_2\mathsf{P} = \mathsf{PP}^{\mathsf{PP}}$, $\mathsf{C}_3\mathsf{P} = \mathsf{PP}^{\mathsf{C}_2\mathsf{P}}$,...

Two consequences of $\mathrm{PER} \in \mathsf{VP}^0$:

(i) $\mathsf{CH} \subseteq \mathsf{P}/\mathsf{poly}$.

(ii) (almost) completeness of the permanent:
for any $(f_n) \in \mathsf{VNP}^0$ we have $(2^{p(n)} f_n) \in \mathsf{VP}^0$
for some polynomially bounded sequence $p(n) \in \mathbb{N}$.

Proof sketch (3/4)

Expand product: $g_n(X) = \prod_{i=1}^{2^n-1}(X+i) = \sum_{\alpha=0}^{2^n-1} a_n(\alpha)X^\alpha.$

Binary expansion: $a_n(\alpha) = \sum_{i=0}^{2^{c.n}-1} a_n(i,\alpha)2^i.$

Hence:

$$\begin{aligned} g_n &= \sum_{\alpha=0}^{2^n-1}\sum_{i=0}^{2^{c.n}-1} a_n(i,\alpha)2^i X^\alpha \\ &= h_n(X^{2^0}, X^{2^1}, \ldots, X^{2^{n-1}}, 2^{2^0}, 2^{2^1}, \ldots, 2^{2^{c\cdot n-1}}) \end{aligned}$$

where $h_n(X_1, \ldots, X_n, Z_1, \ldots, Z_{c\cdot n})$ is the multilinear polynomial

$$\sum_\alpha \sum_i a_n(i,\alpha) X_1^{\alpha_1} \cdots X_{\cdot n}^{\alpha_{c\cdot n}} Z_1^{i_1} \cdots Z_{c\cdot n}^{i_{c\cdot n}}.$$

We would like to apply Valiant's criterion…

Proof sketch (4/4)

Recall: $h_n = \sum_\alpha \sum_i a_n(i,\alpha) X_1^{\alpha_1} \cdots X_n^{\alpha_n} Z_1^{i_1} \cdots Z_{c\cdot n}^{i_{c\cdot n}}$.

The $a_n(i,\alpha)$ can be computed in CH (Bürgisser),
and $\mathsf{CH} \subseteq \mathsf{P}/\mathsf{poly}$ since $\mathrm{PER} \in \mathsf{VP}^0$.

Hence $(h_n) \in \mathsf{VNP}^0$ (Valiant's criterion),
$2^{p(n)} h_n \in \mathsf{VP}^0$ since $\mathrm{PER} \in \mathsf{VP}^0$ (second application of hypothesis),
and $2^{p(n)} h_n$ has depth-4 circuits of size $2^{o(n)}$ (Agrawal - Vinay).

Substitution of powers 2^{2^i} and X^{2^j} in $h_n \Rightarrow$

$2^{p(n)} \prod_{i=1}^{2^n-1} (X+i)$ can be written as a SPS polynomial of size $2^{o(n)}$. $\square$

Algebraic number generators

This is a sequence $(f_i)_{i\geq 1}$ of nonzero polynomials of $\mathbb{Z}[X]$:
$f_i(X) = \sum_\alpha a(\alpha, i)X^\alpha$ where

1. $\deg(f_i) \leq i^c$ and $|a(\alpha, i)| \leq 2^{i^c}$ for some constant c;
2. The $a(\alpha, i)$ can be computed *efficiently*, i.e.,

$$L(f) = \{(\alpha, i, j);\ \text{ the } j\text{-th bit of } a(\alpha, i) \text{ is equal to } 1\}$$

is in P...or in $\mathsf{P/poly}$ ...or even in $\mathsf{CH/poly}$.

Example: $L(f) \in \mathsf{P}$ for $f_i(X) = X - i$, $X^i - 1$ or $X^i - 2^i X + i^2 + 1$.

Remarks: A generator generate the roots of the f_i;
We will consider hitting sets made of the roots of an initial segment of the f_i.

Statement of main theorem

Consider $f(X) = \sum_{i=1}^{k} \prod_{j=1}^{m} f_{ij}(X)$ where the f_{ij} are t-sparse; $\mathrm{size}(f) =$ number of monomials in this expression ($\leq kmt$).

Recall the **Definition:** $f \in \mathrm{SPS}_{s,e}$ if $\mathrm{size}(f) \leq s$, $\deg(f_{ij}) \leq e$, and each coefficient of each f_{ij}:

(i) is of absolute value at most 2^e;

(ii) has $\leq s$ nonzero digits in its binary representation (f_{ij} is a sparse polynomial with sparse coefficients).

Theorem: Let (f_i) be an algebraic number generator, and H_m the set of all roots of the polynomials f_i for all $i \leq m$. If there exists a polynomial p such that $H_{p(s+\log e)}$ is a hitting set for $\mathrm{SPS}_{s,e}$ then the permanent is not in VP^0.

To Be Done...

- Real τ-conjecture: prove or disprove.
- Real τ-conjecture, case $k = 2$: prove or disprove.
- Case $k = 2$, continued:
  give a deterministic algorithm to test identities of the form

$$F_1 \times \cdots \times F_m = G_1 \times \cdots \times G_m$$

  where the F_i and G_i are sparse;
  construct hitting sets (real or otherwise).
- Adapt to univariate setting recent results on deterministic PIT for circuits of bounded depth (3 or 4) and bounded k (as above, k = fan-in of output gate).