Error-correcting codes vs. superconcentrator graphs

Anna Gál

Joint work with:
Kristoffer Hansen, Michal Koucky, Pavel Pudlak, Emanuele Viola

Asymptotically good codes

$C : \{0,1\}^k \to \{0,1\}^n$ is (ρ, δ)-good if

1. $k/n \geq \rho$
2. min distance of C is at least δn

What is the complexity of ENCODING asymptotically good codes?

Complexity of ENCODING good codes

[Gelfand, Dobrushin, Pinsker 1973] There exist good codes that can be encoded in linear time.

[Spielman 1995] Explicit good codes encodable (also decodable) in linear time and linear space.
Encoding by NC_1 circuits.

[Bazzi, Mitter 2005] Good codes cannot be encoded in linear time and sub-linear space.

Unbounded fan-in circuits - with aritrary gates

TRADEOFFS between # of wires and depth

· any code can be encoded in depth 1 with $O(n^2)$ wires

· for good codes $\Omega(n^2)$ wires are necessary in depth 1
regardless of types of gates
each input has to be connected to $\geq \delta n$ outputs

Tight bounds on # of wires for every fixed depth d

$d = 2$	$\Theta(n(\log n/\log\log n)^2)$	
$d = 3$	$\Theta(n\log\log n)$	
$d = 4, 5$	$\Theta(n\log^* n)$	
$d = 2i, 2i+1$	$\Theta(n\lambda_i(n))$	$(i \geq 2)$

Lower bounds: regardless of types of gates

Upper bounds: probabilistic, using only XOR gates.

"almost" linear number of wires in very small depth

$\lambda_{i+1} = \lceil \lambda_i^*(n) \rceil$ inverse Ackerman function

* operation:

number of times to iterate λ_i to get a value ≤ 1

$$f^*(n) = min\{t \mid \underbrace{f(f(\ldots f(n)\ldots)}_{t \text{ times}} \leq 1\}$$

$\lambda_1(n) = \log(n)$

$\lambda_2(n) = \log^*(n)$

$\lambda_3(n) = (\log^*)^*(n)$

Our lower bounds hold for any good code regardless of types of gates allowed in the circuits.

We establish some connectivity properties that ANY circuit computing error-correcting codes must have.

Superconcentrator graphs

Definition [Valiant 1975]

Directed acyclic graph with n inputs V_1, n outputs V_2,
s.t. for any $1 \leq t \leq n$ and any $X \subseteq V_1$, $Y \subseteq V_2$ with
$|X| = |Y| = t$ $\exists$ t vertex disjoint paths from X to Y.

complete bipartite graph: n^2 edges
Valiant: superconcentrators with linear number of edges
(even in depth $O(\log n)$, e.g. Pippenger)

Spielman: connection between codes and superconcentrators

Relaxed superconcentrators

Definition [Pudlák 1994]

Directed acyclic graph with n inputs V_1, n outputs V_2,

s.t. for any $1 \leq t \leq n$ and

for RANDOM $X \subseteq V_1$, $Y \subseteq V_2$ with $|X| = |Y| = t$

$E_{X,Y}$[# of vert.disj.paths from X to Y] $\geq \delta t$

(different relaxation: [Dolev, Dwork, Pippenger, Wigderson 1983])

Connectivity property of circuits for codes

Definition "semi-relaxed" superconcentrator

Directed acyclic graph with k inputs V_1, n outputs V_2,

s.t. for any $1 \leq t \leq k$,

for ANY $X \subseteq V_1$ and RANDOM $Y \subseteq V_2$

with $|X| = |Y| = t$

E_Y[# of vert.disj.paths from X to Y] $\geq \delta t$

Theorem Any circuit encoding $C : \{0,1\}^k \to \{0,1\}^n$ with min. dist. δn must be a "semi-relaxed" superconcentrator (with parameter δ).

Connectivity properties

Superconcentrators: ANY $X \subseteq V_1$, ANY $Y \subseteq V_2$

Circuits for codes: ANY $X \subseteq V_1$, RANDOM $Y \subseteq V_2$

Relaxed s.c.: RANDOM $X \subseteq V_1$, RANDOM $Y \subseteq V_2$

Depth $d \geq 3$ We use known lower bounds on relaxed superconcentrators [Pudlák 1994]

$d = 3$	$\Omega(n \log \log n)$	
$d = 2i, 2i + 1$	$\Omega(n\lambda_i(n))$	$(i \geq 2)$

We prove matching upper bounds.

Depth $d = 2$

Superconcentrators: $\Theta(n(\log n)^2/\log\log n)$

[Radhakrishnan, Ta-Shma 2000]

Circuits for codes: $\Theta(n(\log n/\log\log n)^2)$

Relaxed superconcentrators: $\Theta(n\log n)$ [Pudlák 1994, DDPW 1983]

Lower bound proof sketch

Theorem Any circuit encoding $C : \{0,1\}^k \to \{0,1\}^n$ with min. dist. δn must be a "semi-relaxed" super-concentrator (with parameter δ).

Sketch of proof Start with any set X of inputs of size t, pick random Y one element at a time.

We show: as long as $|Y| < t$, with probability at least δ the next randomly chosen output will increase the number of vertex disjoint paths from X to the current Y by one.

Lower bound proof sketch

$|X| = t$, $|Y| < t$. Let the # of vertex disjoint paths from X to Y be $j < t$.

Suppose B is a set s.t. the # of vertex disjoint path from X to $Y \cup B$ is still j. By Menger's theorem, there is a set S of j vertices s.t. every path must contain a vertex from S.

But then, the values at the outputs $Y \cup B$ can take at most 2^j different values.

There are two different codewords $C(0x_1)$ and $C(0x_2)$,
that agree on $Y \cup B$,
there must be at least δn vertices outside $Y \cup B$.

Lower bound proof sketch

We need: at least δn vertices s.t. adding any of them increases # of paths.

Vertex v is bad if the # of vertex disjoint path from X to $Y \cup v$ is not larger than from X to Y.

BAD: set of all bad vertices.

Lemma # of vertex disjoint path from X to $Y \cup BAD$ is not larger than from X to Y.

Equivalent to a classical theorem from matroid theory.

Matroid Lemma [Hazel Perfect, 1968]

Collections of endpoints of vertex disjoint paths are independent sets of a matroid.

Note: not true for collections of vertex disjoint paths.

Recall: generator matrix of good codes is dense

must have $\Omega(n^2)$ nonzero entries

Corollary of our upper bounds:

There are dense generator matrices that can be obtained as the product of two "sparse" matrices.

with $O(n(\log n/ \log\log n)^2)$ nonzero entries

Upper bound: probabilistic construction

Depth 2 proof sketch:
Middle layer: $\log n$ groups of vertices (XOR gates).
i-th group: $2^i \log n$ vertices with fan-in $n/2^i$.
Total # of wires so far: $n(\log n)^2$.

Correspond to range detectors: i-th group "detects" weight $(2^{i-1}, 2^i]$: on inputs with such weight outputs a balanced string: constant fraction of 1's and 0's.

Each output gate is connected to one random gate in each group.
$n \log n$ wires to outputs.

For every nonzero message, at least one balanced group.
This implies, constant fraction of output XOR gates is odd.

Open problems

We have shown by a probabilistic construction the existence of good codes that can be encoded with "almost" linear number of wires in very small depth

OPEN:

Explicit construction of such codes

Complexity of decoding