Erasures vs. Errors in Local Decoding and Property Testing

Sofya Raskhodnikova*, Noga Ron-Zewi#, Nithin Varma*

*Boston University, USA #University of Haifa, Israel

Overview of Results

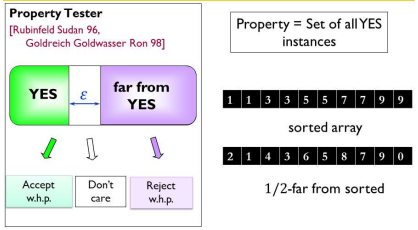
- **Local list decoding in the presence of erasures**
  - Local list erasure-decoding Hadamard Code
  - Constant vs. $\Omega(\log n)$ separation between erasure-resilient testing and tolerant testing
- **Approximate local list decoding in the presence of erasures**
  - Constant vs. $n^{\Omega(1)}$ separation between erasure-resilient testing and tolerant testing
- **Relationship between local decoding in the presence of erasures and in the presence of errors**

Locally (Unique or List) Decodable Codes

- **Locally decodable codes** [BabaiFortnowLevinSzegedy91, GemmellLiptonRubinfeldSudanWigderson91, GemmellSudan92, BlumLubyRubinfeld93, PolishchukSpielman94, BeimelDihaKirshtevitzRaymond02, Yekhanin08, Ben-AsherEfremenkoTaShma10, DvirGopalanYekhanin11, Efremenko12, …]
  - Each message bit can be decoded with high probability by querying a few bits of the codeword
- **Locally list decodable codes** [GoldreichLevin89, SudanTrevisanVadhan01, GuruswamiRudra08, GopalanKlivansZuckerman08, Ben-AsherEfremenkoTaShma10, KoppartyShpilkaYekhanin11, KoppartyRon-ZewiSaraf13, HemenwayRon-ZewiWootters17, GoldreichKoppartyOliveiraRon-ZewiSaraf17, KoppartyRon-ZewiSarafWootters18, …]
  - Given oracle access to codeword, with high probability, obtain a list of descriptions of local decoders of each candidate message
  - Decodes from a larger fraction of corruptions

Can locally decodable codes perform better with erasures than with errors?

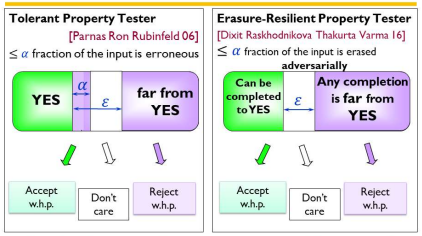
Property Testing



Property = Set of all YES instances

Two objects are at distance ε = they differ in an ε fraction of places

Tolerant and Erasure-Resilient Property Testing

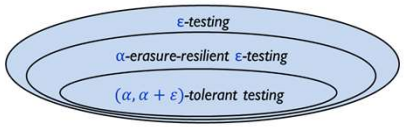


(α, ε)-tolerant testing α-erasure-resilient ε-testing

Relationships Between Models

Containments are strict:

- [Fischer Fortnow 05]: standard vs. tolerant
- [Dixit Raskhodnikova Thakurta Varma 16]: standard vs. erasure-resilient
- **Our Result**: erasure-resilient vs. tolerant



18

Erasure-resilient testing vs. Tolerant testing

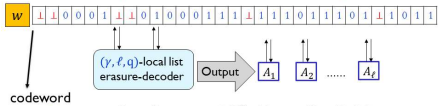
Theorem: There exists a property R of n-bit strings such that

- **erasure-resilient testing** R has constant query complexity;
- **tolerant testing** R needs $n^{\Omega(1)}$ queries.

This Poster
constant vs. $\Omega(\log n)$

A Locally List Erasure-Decodable Code

- An error-correcting code $\mathcal{C}_n: \Sigma^n \to \Sigma^N$, usually $N \gg n$
- Parameters: γ fraction of erasures, list size ℓ and q queries.



– w.h.p., for every $x \in \Sigma^n$ with encoding $\mathcal{C}_n(x)$ that agrees with w on all non-erased bits, one of the algorithms A_j, given oracle access to w, simulates oracle access to x;

– each algorithm A_j makes at most q queries to w.

Hadamard Code

Hadamard: $\{0,1\}^k \to \{0,1\}^{2^k}$; $\text{Hadamard}(x) = (\langle x, y \rangle)_{y \in \{0,1\}^k}$

Type of Corruptions	Corruption Tolerance γ	Number of Queries	List Size	Reference
Errors	$0 \le \gamma < 1/2$	$\Theta\left(\frac{1}{(1/2-\gamma)^2}\right)$	$\Theta\left(\frac{1}{(1/2-\gamma)^2}\right)$	[Goldreich Levin89, Blinovsky86, Guruswami Vadhan10, GrinbergShaltiel Viola18]
Erasures*	$0 \le \gamma < 1$	$\Theta\left(\frac{1}{1-\gamma}\right)$	$O\left(\frac{1}{1-\gamma}\right)$	[new, GrinbergShaltiel Viola18]

If fraction of errors is $\ge$ 1/2, impossible to decode Hadamard codes.

*An improvement in dependence on γ was suggested by Venkat Guruswami.

3CNF Properties: Hard to Test, Easy to Decide

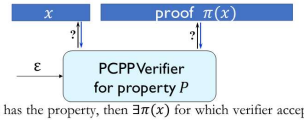
- Formula ϕ_n : 3CNF formula on n variables, $\Theta(n)$ clauses
- Property $R_{\phi_n} \subseteq \{0,1\}^n$: set of satisfying assignments to ϕ_n

Theorem [Ben-Sasson Harsha Raskhodnikova 05]
There exists sufficiently small ε^*, ε^* -testing R_{ϕ_n} requires $\Omega(n)$ queries.

- R_{ϕ_n} decidable by a $O(n)$-size circuit.

Testing with Advice: PCPs of Proximity (PCPPs)

[Ben-Sasson Goldreich Harsha Sudan Vadhan 06, Dinur Reingold 06]

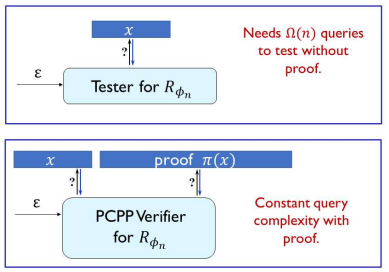


- If x has the property, then $\exists \pi(x)$ for which verifier accepts.
- If x is ε-far, then $\forall \pi(x)$ verifier rejects with probability $\ge 2/3$.

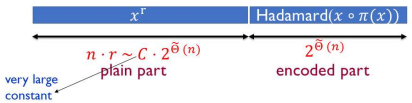
Theorem [Dinur 07]
Every property decidable with a circuit of size m has PCPP with proof length $\tilde{O}(m)$ and constant query complexity.

3CNF properties have efficient PCPPs

3CNF properties have efficient PCPPs



Separating Property P



- x satisfies the hard 3CNF property R_{ϕ_n}
- $\pi(x)$ is the proof on which the PCPP verifier accepts x
- r is the number of repetitions to make the length of plain part a large multiple of the length of encoded part

Separating Property: Erasure-Resilient Testing

Idea: If a constant fraction (say, 1/4) of the encoding is preserved, we can locally list erasure-decode the encoded part.

Inputs: $\varepsilon \in (0,1), \alpha \in [0,1)$, oracle access to $y \in \{0,1\}^N$

1. Locally list erasure-decode Hadamard to get a list of algorithms.
2. For each algorithm:
   - **repeat** $\Theta(\frac{1}{\varepsilon(1-\alpha)})$ times: // repetition check
     - Check whether a random nonerased bit in plain part equals the corresponding bit in decoding.
   - **repeat** $\Theta(1)$ times: // PCPP check
     - Check whether PCPP verifier accepts (decoded) $x \circ \pi(x)$.
3. **Accept** if, for some algorithm on the list, both checks pass.

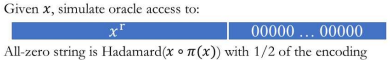
Constant query complexity!

Hardness of Tolerant Testing P

Theorem
Tolerant testing P needs $\tilde{\Omega}(\log N)$ queries.

Idea: Reduce standard testing of 3CNF property (in the absence of proof) to tolerant testing of the separating property.

- Given x, simulate oracle access to:



- All-zero string is Hadamard$(x \circ \pi(x))$ with 1/2 of the encoding bits are erroneous!

$\Omega(n) \approx \tilde{\Omega}(\log N)$ queries are needed to tolerant test P.

Errors and Erasures in Local Decoding

- **Local decoding implies local erasure-decoding**
  - locally decodable from at most an α fraction of errors ⇒ locally decodable from at most an 2α fraction of erasures
  - Also holds for local list decoding and approximate local list decoding
- **Local erasure-decoding implies local decoding** (up to some parameters)
  - locally decodable from at most an α fraction of erasures using q queries ⇒ locally decodable from at most an $\alpha/O(q^2 \cdot 9^q)$ fraction of errors using $O(q \cdot 3^q)$ queries

Open Questions

- Even stronger separation -- constant vs. linear ?
- Separation between errors and erasures for a "natural" property?
- Constant-query, constant list size, local list erasure-decodable codes with inverse polynomial rate ?