Recovering Simple Signals

Anna C. Gilbert

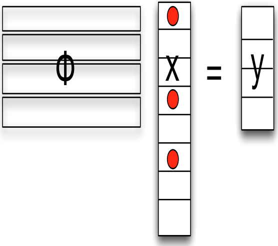
Department of Mathematics
University of Michigan

Joint work with B. Hemenway (U. Mich), A. Rudra (Buffalo), M. J. Strauss (U. Mich), and M. Wootters (U. Mich)

(Sparse) Signal recovery problem

signal or population length N k important features

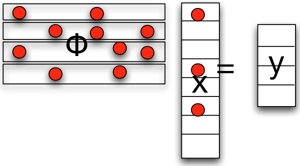


measurements or tests: length m

Under-determined linear system: $\Phi x = y$

Given Φ and y, recover information about x

Two main examples: group testing and compressed sensing



Group testing

Φ binary = pooling design

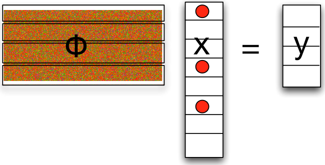
x binary, $1 \implies$ defective, $0 \implies$ acceptable

OUTPUT: defective set

success = number items found

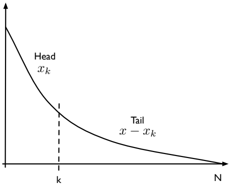
Arithmetic: OR

Two main examples: group testing and compressed sensing



Compressed sensing

Φ = measurement matrix
x signal, sparse/compressible
OUTPUT: $\widehat{x}$ good approximation to x
success = $\|x - \widehat{x}\|_2$ versus $\|x - x_k\|_2$
Arithmetic: $\mathbb{R}$ or $\mathbb{C}$



Design problems: matrices and algorithms

Design Φ with $m < N$ rows *and* recovery algorithm s.t.

$$\|x - \widehat{x}\|_2 \leq C\|x - x_k\|_2.$$

- **Adversarial or "for all"** recover all x that satisfy a geometric constraint:

  tail of x is really compressible [Candes, et al.'04, Donoho '04]

  $$\|x - x_k\|_1 \leq \sqrt{k}\|x - x_k\|_2$$

  block sparse/compressible [Eldar, et al.'09]
  sparsity patterns connected chain in binary tree
  i.e., model sparse/compressible [Baraniuk, et al.'09]
- **Probabilistic or "for each"**: recover all x that satisfy a statistical constraint

  fixed signal, recover whp over construction of Φ [GGIKMS'01]
  uniform distribution over k-sparse signals
  i.e., random signal model [Calderbank, et al.'08, Cevher, et al. '08, Sapiro, et al.'11]

Extremal models: pros and cons

- **Adversarial**

  places minimal assumptions on signal $\implies$ widely applicable
  positive results hard to come by
  unlikely that natural process is worst-case
  which geometric model?

- **Probabilistic**

  positive results easier to come by
  not as applicable
  debatable if natural process is oblivious to Φ or follow simple, fully specified random process
  which random process?

We need a middle ground: Compromise!

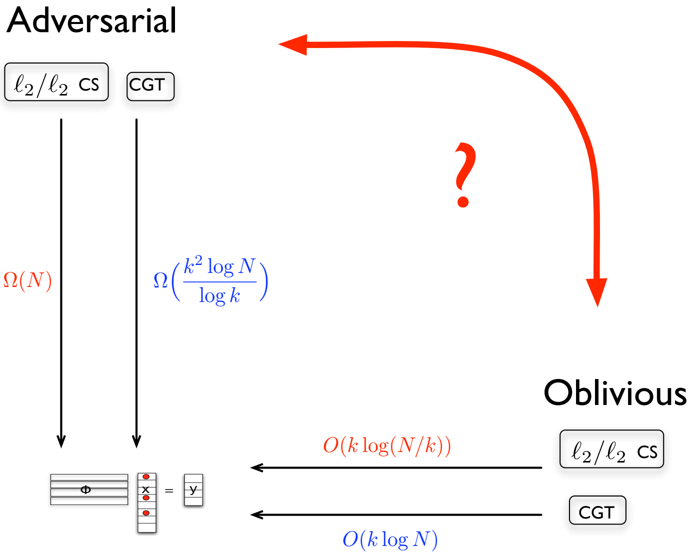
Feedback: never just measure, reconstruct once, and done

Future signals depend on measurements of current signals

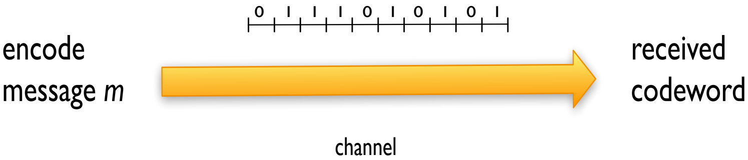
Benign dependency: store inventory

Adversarial dependency: radar detection of adversary and evasive action

We need a middle ground: lower bounds and separations?



Example: error-correcting codes

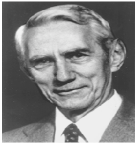


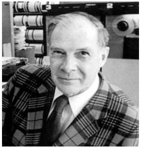
message $m \in \mathcal{M}$, over alphabet Σ

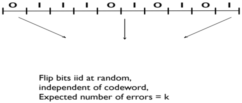
encode with codebook $\mathcal{C} \subset \Sigma^n$

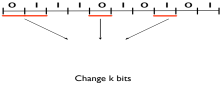
rate $= \frac{\log |\mathcal{M}|}{n \log |\Sigma|}$

Example: error-correcting codes extremal examples









- **Shannon:** channel is oblivious to message or codeword
  can prove existence of capacity-achieving codes
  rate > 0 when $\rho = 1/2$ random errors
- **Hamming:** adversarial process
  imposes strict conditions on codebook: distinct codewords must differ in at least a fraction of 2ρ positions for ρ fraction errors
  rate $= 0$ when $\rho > 1/4$

ECC: middle ground

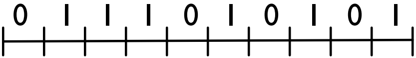

Adam Smith


Venkat Guruswami


Dick Lipton


Madhu Sudan


Silvio Micali



change $\leq k$ bits, **restrict** computation or information about codeword

- probabilistic polynomial time: practical but not an actual limitation
- LOGSPACE: "benign" processes only with small memory

Mallory: Adversarial model

- **Binary symmetric channel:** Entries 1 with prob. k/N and 0 with prob. $(N-k)/N$.
- **Oblivious:** Mallory generates x with no information about Φ
- **Information-theoretically bounded:** Mallory generates x with bounded mutual information with Φ.
  Algorithm M is information-theoretically bounded if $M(x) = M_2(M_1(x))$ where the output of M_1 consists of at most $O(\log(|x|))$.
- **Streaming log-space:** Mallory streams over *rows* of Φ, has only LOGSPACE to store information and to produce vector x.
- **Adversarial:** Mallory is fully malicious.

Example: Randomized algorithms against adversaries

String of length N, $N/2$ a's and $N/2$ b's

Randomized algorithm to produce position of a in vector:

0. Choose k positions at random
1. If a is in (at least) one of m positions, return position (**Success**)
   else, return $\emptyset$ (**Fail**)

Probability of success $= 1 - (1/2)^m$ on any *fixed* string

If Mallory *knows* which m positions (i.e, the random string used by the algorithm), she puts b's in those slots and Fail!

$A(x, r) =$ randomized algorithm, succeeds with prob. $1 - \epsilon$

$x \in \{0, 1\}^N$ input string and $r \in \{0, 1\}^m$ random string

Results

Combinatorial group testing		
Mallory	Num. Measurements	Reference
Adversarial	$\Omega(k^2 \log(N/k)/\log(k))$	[Furedi, and more]
Information-Theoretically bounded (logspace)	$O(k \log(N))$	new
Logspace streaming (one pass over the rows)	$\Omega(k^2/\log k)$	new
Deterministic $O(\log k \log N)$ space	$\Omega(k^2/\log k)$	new
Oblivious	$O(k \log(N))$	new
Binary symmetric channel	$\Omega(k \log(N/k))$, $O(k \log(N))$	new
Sparse signal recovery		
Mallory	Num. Measurements	Reference
Adversarial	$\Omega(N)$	[CDD'09]
Adversarial, but restricted so that $\|x - x_k\|_1 \leq \sqrt{k}\|x - x_k\|_2$	$O(k \log(N/k)$	[CRT'06, Donoho'06]
Information-Theoretically bounded (logspace)	$O(k \log(N/k))$	new
Logspace streaming (one pass over the rows)	$O(k \log(N/k))$	new
Oblivious	$O(k \log(N/k))$	[GLPS'10]

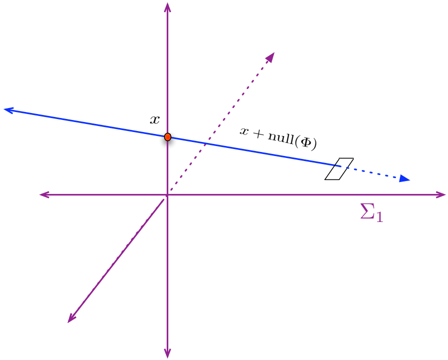
Sketch of results for CS

Intuition: geometry of null space of Φ

Info.-theory bounded adversary: judicious use of simple lemma and existing algorithms and matrix constructions (Gaussian, Bernoulli, and hashing)

Streaming adversary: communication complexity arguments

Intuition: $2k$ measurements for exact k-sparse signals



Example: Φ is 2×3 matrix

$m = 2$, $N = 3$, $k = 1$

$\dim(\mathrm{null}(\Phi)) = 1$

$\dim(\Sigma_1) = 1$

For unique solution to $\Phi x = y$,

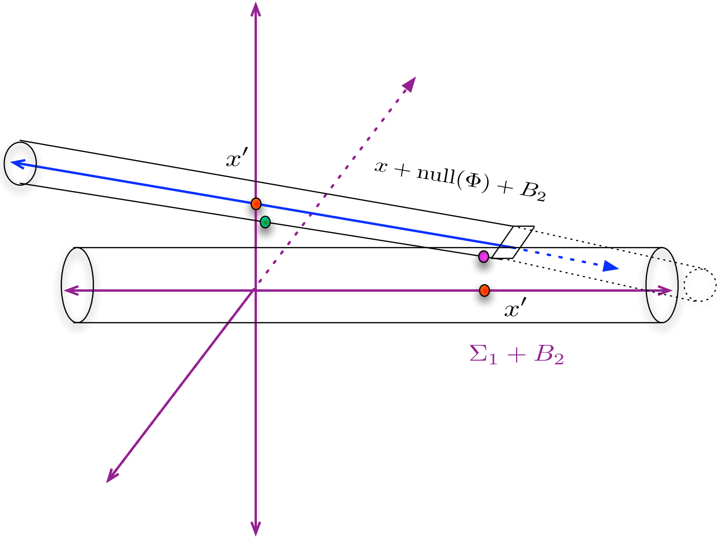
x exactly 1-sparse

$x + \mathrm{null}(\Phi) \cap \Sigma_1 = \{x\}$

$\iff \mathrm{null}(\Phi) \cap (-x + \Sigma_1) = 0$

$\iff \mathrm{null}(\Phi) \cap \Sigma_{2k} = 0$

Intuition: null-space condition [CDD'09]



Lemma: randomized algorithms and Mallory

Lemma

$A(x, r)$ randomized algorithm with success probability $1 - \epsilon$.

Let $\ell < N$ (space assigned to Mallory–runs $M(r)$).

Fix $0 < \alpha < 1$.

For any such Mallory, $A(M(r), r)$ succeeds with probability at least

$$\min\left\{1 - \alpha, 1 - \frac{\ell}{\log(\alpha/\epsilon)}\right\}$$

over the choice of r.

Conclusions

For CS: same (small) number of measurements against adversaries as for oblivious with ℓ_2/ℓ_2 error guarantees

For CGT: more interesting adversaries and different numbers of measurements

Good compromise between extremes

Alternative to *statistical* or *geometric* models

Recognition of measure + reconstruct feedback

(as opposed to measure same signal *repeatedly* or measure several *simultaneously*)