EXPANDER ℓ_0-DECODING

Rodrigo Mendoza-Smith Jared Tanner

Mathematical Institute, University of Oxford

1. Combinatorial Compressed Sensing

Combinatorial compressed sensing studies the problem of sampling and efficiently reconstructing a k-sparse signal $x \in \mathbb{R}^n$ from $m < n$ linear measurements of the form $y = Ax \in \mathbb{R}^m$ where A is an *expander matrix.*

Expander matrices: Let $A \in \mathbb{R}^{m\times n}$ be a sparse binary matrix with exactly $d << m$ ones per column, and let

$$\mathcal{N}_1(S) := \{i \in [m] : A_{ij} = 1 \text{ for exactly one } j \in S\}.$$

Then, A is a (k, ε, d)-expander matrix ($A \in \mathbb{E}_{k,\varepsilon,d}$) if for some $\varepsilon \in (0, 1)$ we have

$$|\mathcal{N}_1(S)| > (1 - 2\varepsilon)d|S| \;\; \forall \;\; S \in [n]^{(\leq k)}. \quad (1)$$

- Any sparse binary matrix with exactly d ones per column is a (k, ε, d)-expander for some k, ε.
- Expander matrices have a null-space property, so recovery is possible via convex relaxation.
- Expander matrices are low complexity and have an optimal measurement rate of $\mathcal{O}(k \log(n/k))$.

Prior art:

Algorithm 1: Iterative greedy CCS algorithms

Data: $A \in \mathbb{R}^{m\times n}$; $y \in \mathbb{R}^m$
Result: $\hat{x} \in \mathbb{R}^n$ s.t. $y = A\hat{x}$
$\hat{x} \leftarrow 0$, $r \leftarrow y$;
while *not converged* **do**
  Compute a score s_j and an update u_j $\forall$ $j \in [n]$;
  Select $S \subset [n]$ based on a rule on s_j;
  $\hat{x}_j \leftarrow \hat{x}_j + u_j$ for $j \in S$;
  k-threshold $\hat{x}$;
  $r \leftarrow y - A\hat{x}$;

Algorithm	Complexity
SMP [1]	$\mathcal{O}((nd + n\log n)\log \|x\|_1)$
SSMP [2]	$\mathcal{O}((\frac{d^3 n}{m} + n)k + (n\log n)\log \|x\|_1)$
LDDSR [3]	$\mathcal{O}((\frac{d^3 n}{m} + n)k)$
ER [4]	$\mathcal{O}((\frac{d^3 n}{m} + n)k)$

2a. Our work

- SMP updates multiple $\hat{x}_j$ per iteration. It is empirically fast, but unstable.
- SSMP, LDDSR, ER update a single $\hat{x}_j$ per iteration. They are slower than SMP, but stable.
- We propose an algorithmic model that stably updates multiple entries of $\hat{x}$ by choosing those $j \in [n]$ that would yield a decrease $\|r\|_0$.
- To be able to update multiple entries per iteration without compromising the region of recovery, we assume a *dissociated* signal model on x.

2b. Signal model

A signal $x \in \mathbb{R}^n$ is *dissociated* if

$$\sum_{j\in T_1} x_j \neq \sum_{j\in T_2} x_j \qquad \forall\, T_1, T_2 \subset \mathrm{supp}(x) \text{ s.t. } T_1 \neq T_2.$$

- x is dissociated almost surely if its nonzeros are drawn from a continuous distribution.

2c. Exp-ℓ_0-De

Algorithm 2: Serial-ℓ_0 [5]

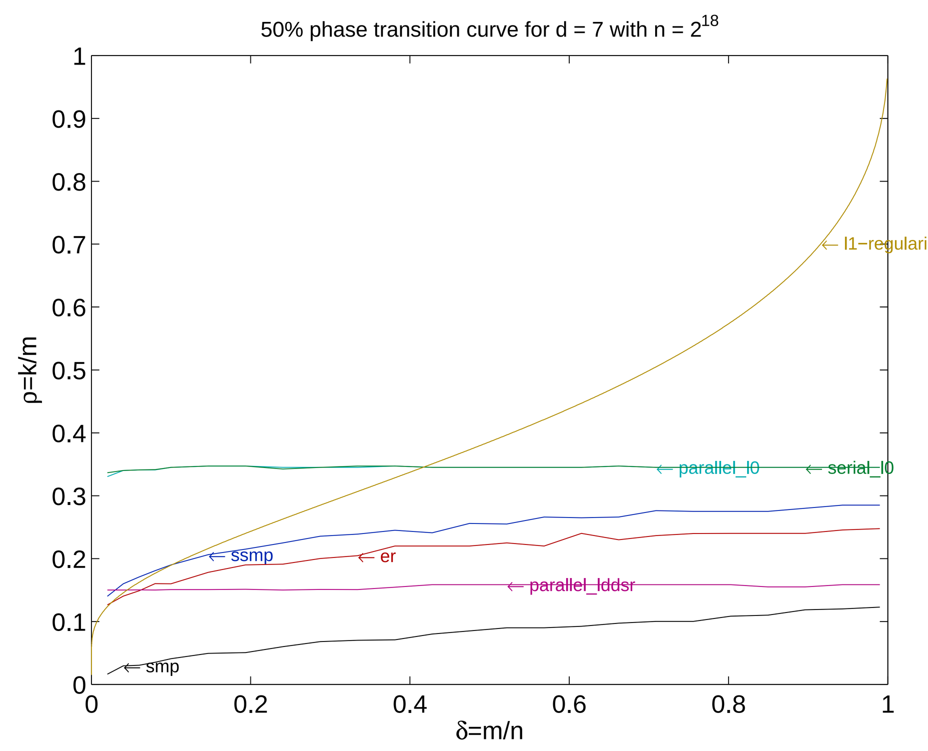
Data: $A \in \mathbb{R}^{m\times n}$; $y \in \mathbb{R}^m$; $\alpha \in [d]$
Result: $\hat{x} \in \mathbb{R}^n$ s.t. $y = A\hat{x}$
$\hat{x} \leftarrow 0$, $r \leftarrow y$;
while *not converged* **do**
  for $j \in [n]$ **do**
    $T \in \{\omega_j \in \mathbb{R} : \|r\|_0 - \|r - \omega_j a_j\|_0 > \alpha\}$;
    for $\omega_j \in T$ **do**
      $\hat{x}_j \leftarrow \hat{x}_j + \omega_j$;
  $r \leftarrow y - A\hat{x}$;

Algorithm 3: Parallel-ℓ_0 [5]

Data: $A \in \mathbb{R}^{m\times n}$; $y \in \mathbb{R}^m$; $\alpha \in [d]$
Result: $\hat{x} \in \mathbb{R}^n$ s.t. $y = A\hat{x}$
$\hat{x} \leftarrow 0$, $r \leftarrow y$;
while *not converged* **do**
  $T \leftarrow \{(j, \omega_j) \in [n] \times \mathbb{R} : \|r\|_0 - \|r - \omega_j a_j\|_0 > \alpha\}$;
  for $(j, \omega_j) \in T$ **do**
    $\hat{x}_j \leftarrow \hat{x}_j + \omega_j$;
  $r \leftarrow y - A\hat{x}$;

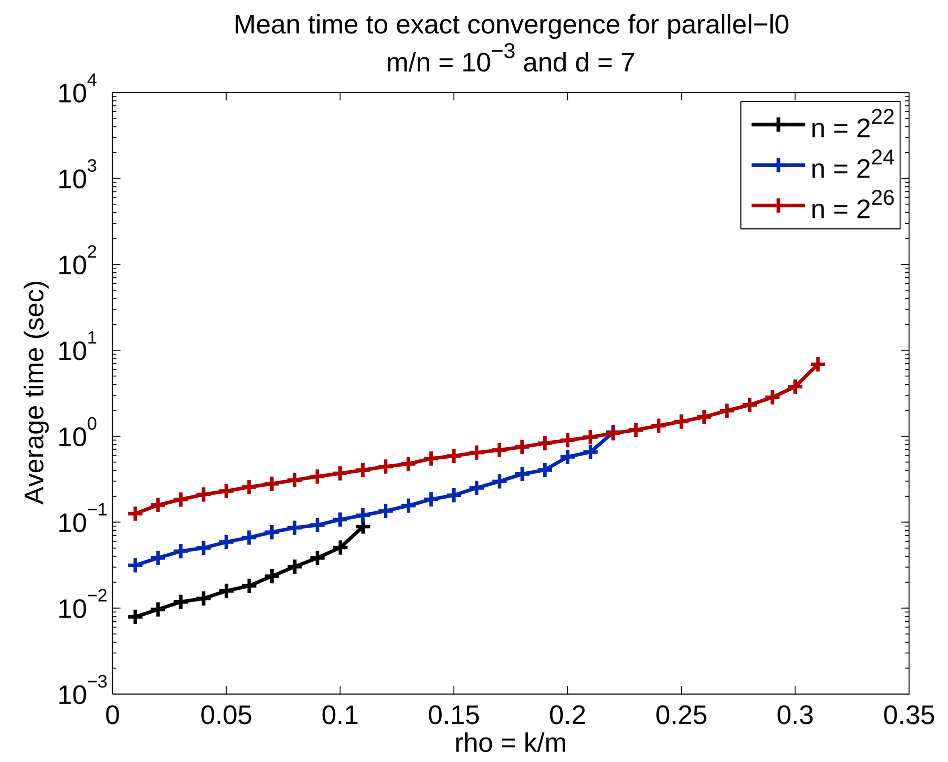
Theorem 0.1. *Let $A \in \mathbb{E}_{k,\varepsilon,d} \cap \mathbb{R}^{m\times n}$, and $x \in \mathbb{R}^n$ be a k-sparse dissociated signal. If $\varepsilon < 1/4$ and $\alpha = d/2$, then Algorithms 2-3 solve $y = Ax$ in $\mathcal{O}(dn \log k)$ operations.*

3a. High phase transition



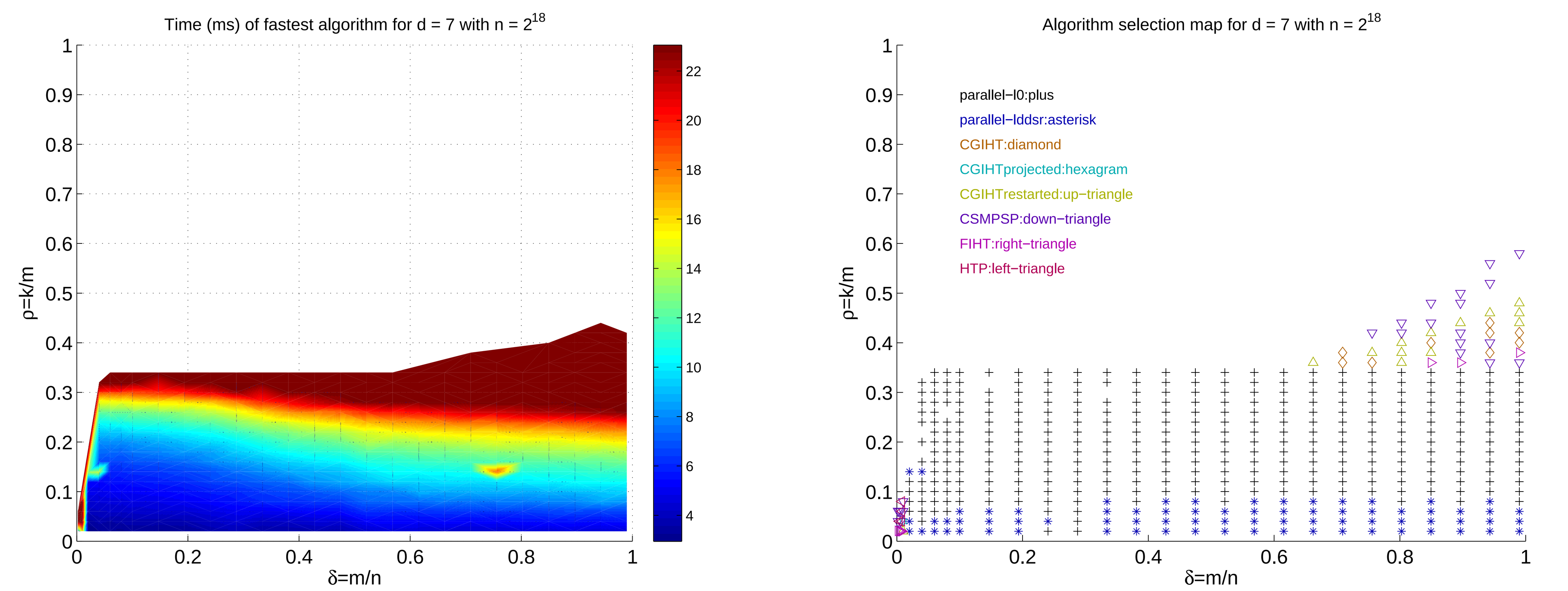
Algorithms 2-3 have substantially higher phase transitions than the rest of the CCS algorithms, and higher than ℓ_1-regularisation for $\delta := m/n \lesssim 0.4$.

3b. High phase transition $\delta \approx 0$



For fixed $\delta << 1$, we observe an increasing phase transition as $n \to \infty$. Also, the time increases linearly with n.

3c. Fastest CS algorithms for dissociated signals



When implemented in parallel, Algorithm 3 is the fastest compressed sensing algorithm for dissociated x, except for regions of low $\rho := k/m$ where a parallel version of LDDSR with multiple updates per iteration is the fastest. However, the convergence of this version of LDDSR is also given by Theorem 0.1.

References

[1] Radu Berinde, Piotr Indyk, and Milan Ruzic. Practical near-optimal sparse recovery in the l1 norm. In *Communication, Control, and Computing, 2008 46th Annual Allerton Conference on*, pages 198–205. IEEE, 2008.

[2] Radu Berinde and Piotr Indyk. Sequential sparse matching pursuit. In *Communication, Control, and Computing, 2009. Allerton 2009. 47th Annual Allerton Conference on*, pages 36–43. IEEE, 2009.

[3] Weiyu Xu and Babak Hassibi. Efficient compressive sensing with deterministic guarantees using expander graphs. In *Information Theory Workshop, 2007. ITW'07. IEEE*, pages 414–419. IEEE, 2007.

[4] Sina Jafarpour, Weiyu Xu, Babak Hassibi, and Robert Calderbank. Efficient and robust compressed sensing using high-quality expander graphs. *arXiv preprint arXiv:0806.3802*, 2008.

[5] Rodrigo Mendoza-Smith and Jared Tanner. Expander ℓ_0-decoding. 2015.