Active Search for Sparse Signals with Region Sensing

Yifei Ma* and Roman Garnett and Jeff Schneider***
*Carnegie Mellon University **Washington University in St. Louis
yifeim@cs.cmu.edu

Motivation

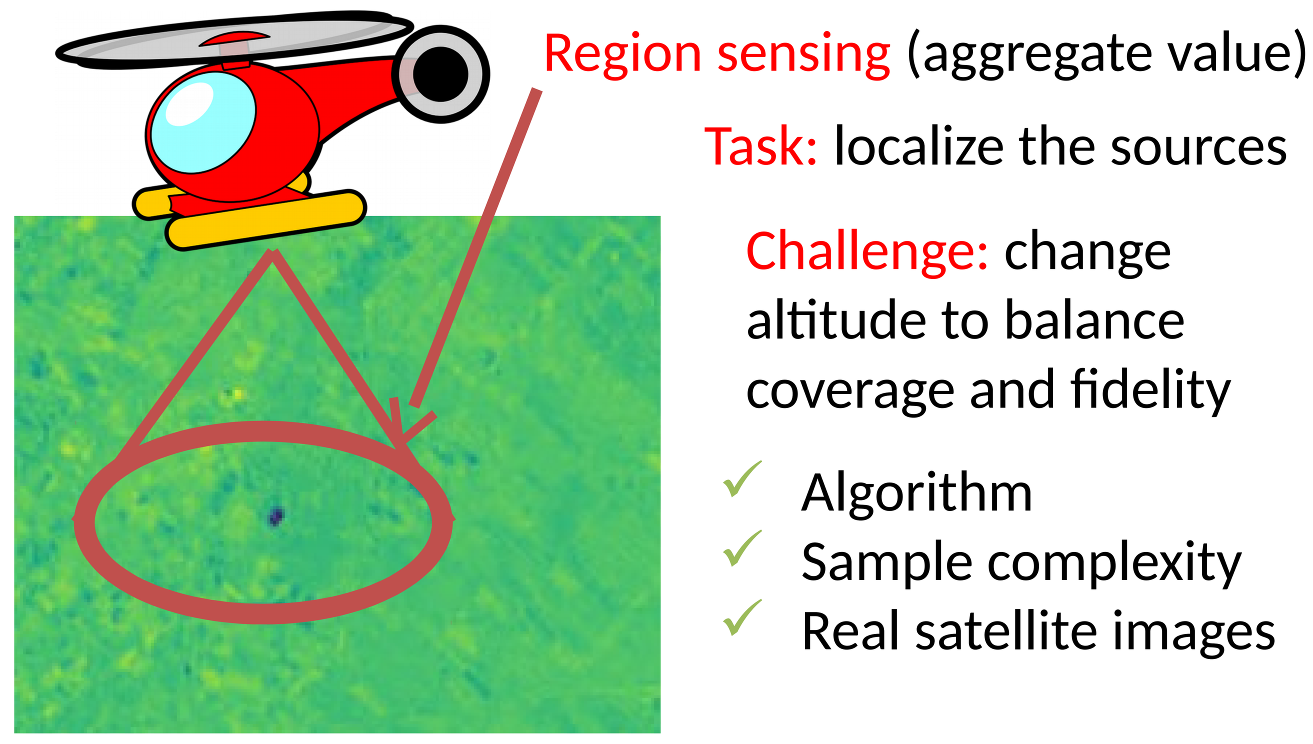
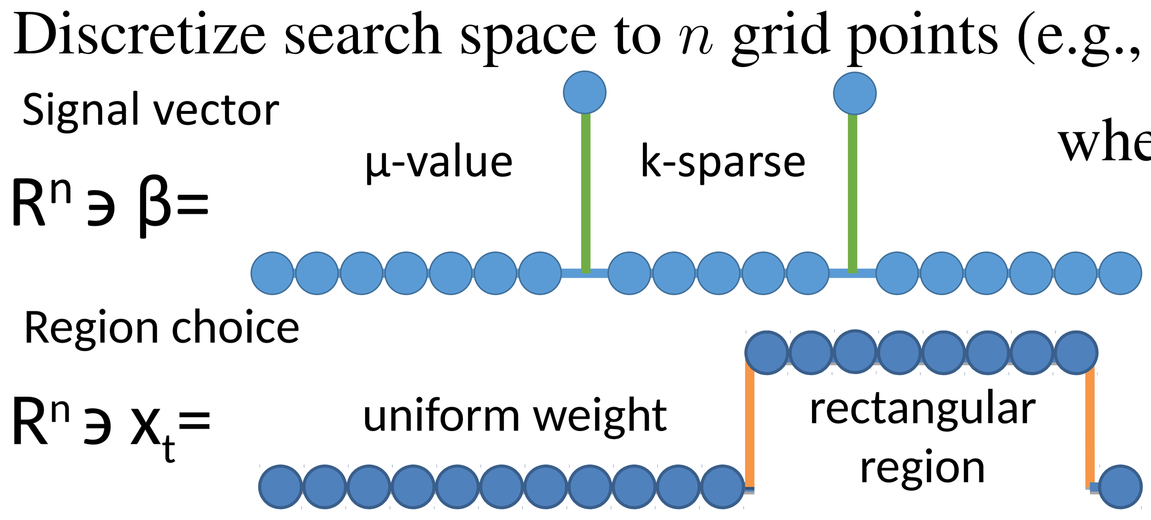


Figure 1: Use aggregate measurements on contiguous regions to find a sparse signal.

Measurement Model

Discretize search space to n grid points (e.g., 1d search)



where $S = \{j : \beta_j > 0\} = \{j : \beta_j = \mu\}$

such that $\|\mathbf{x}_t\|_2 = 1$

Sensing outcome at step t is $y_t = \mathbf{x}_t^\top \boldsymbol{\beta}^* + \varepsilon_t$, where $\varepsilon_t \sim \mathcal{N}(0, 1)$.

Objective

- ❑ Choose $\mathbf{X} = \{\mathbf{x}_t\}_{t=1}^T$ to discover S (let $\hat{S}_T$ be the estimate).
- ❑ Loss is $d(S, \hat{S}_T) = \frac{1}{k}|S \triangle \hat{S}_T|$, where $\triangle$ is the symmetric difference of two sets, $S \triangle \hat{S}_T = (S \setminus \hat{S}_T) \cup (\hat{S}_T \setminus S)$.

Proposed Algorithm: Region Sensing Index (RSI)

Require: n, k, μ
 use uniform prior $\pi_0(\boldsymbol{\beta})$ // (1)
 for $t = 1, \dots, T$ **do**
  pick $\mathbf{x}_t = \arg\max_{\mathbf{x}_t \in \mathcal{X}} I(\boldsymbol{\beta}; y_t \mid \mathbf{x}_t, \pi_{t-1})$ // (2)
  observe $y_t = \mathbf{x}_t^\top \boldsymbol{\beta} + \varepsilon_t$
  update $\pi_t(\boldsymbol{\beta}) \propto \pi_{t-1}(\boldsymbol{\beta}) p(y_t \mid \boldsymbol{\beta}, \mathbf{x}_{t-1})$ // (3)
Ensure: $\hat{S}_T = \arg\min_{\hat{S}} \mathbb{E}\left[d(S(\boldsymbol{\beta}), \hat{S}) \mid \boldsymbol{\beta} \sim \pi_T\right]$

(1) $\pi_0(\boldsymbol{\beta})$ is a uniform distribution on $\binom{n}{k}$ possible k-sparse signal vectors.
(2) For any choice of $\mathbf{x}_t$, compute mutual information between $\boldsymbol{\beta} \sim \pi_{t-1}$ and $(y_t \mid \boldsymbol{\beta}, \mathbf{x}_t) \sim \mathcal{N}(\mathbf{x}_t^\top \boldsymbol{\beta}, 1)$.
(3) Bayes rule based on collected data so until the current step.
(4) In practice, storing $\binom{n}{k}$ models may be infeasible; we find each signal location sequentially (RSI-A).

Theoretical Guarantees in 1D Search

Design type	Region sensing	Algorithm	Bayes prior	Min T to get ϵ-risk	Sample complexity*
passive	yes	(any)	π_0	$T \geq \frac{n}{2}(1 - \frac{n-1}{n-k}\epsilon)$	$\Theta(n)$
		Point sensing	$(\mu \to \infty)$	$T \leq n(1 - \frac{n-1}{n-k}\epsilon)$	
active	no	(any)	$\tilde{\pi}_0$	$T \geq \frac{4n}{\mu^2}(1-\epsilon)^2$	$\Omega(\frac{n}{\mu^2})^\dagger$
	yes	CASS [1]	max risk (incl. π_0)	$T \leq 20\frac{n}{\mu^2}\log(\frac{8k}{\epsilon}) + 2k\log_2(\frac{n}{k})$	$\tilde{O}(\frac{n}{\mu^2} + k)^\ddagger$
		RSI (ours)	π_0	$\bar{T}_\epsilon \leq 50(\frac{n}{\mu^2} + \frac{k^2}{9}) \cdot \log_2(\frac{2}{\epsilon})\log(\frac{n}{\epsilon})$	$\tilde{O}(\frac{n}{\mu^2} + k^2)^\ddagger$

* Sample complexity assumes $\epsilon = 0.5$, $k \ll n$. † Shown for unconstrained sensingl; binary search requires $\Omega(\log_2(n) + k)$ additional measurements. ‡ $\log(n)$ terms are left out. $\bar{T}$ is defined differently, as the average sample size when the posterior risk can be bounded by ϵ in each case (instead of a fixed value regardless of outcomes).

Simulation Studies

Search space is $1d$; discretized to $n = 1024$ points. Signal is 1-sparse, with strength $\mu = 16$.

- ✓ **RSI**: most efficient choice of measurements.
- × **CASS** [1]: less efficient in general. not anytime; produces only turning points.*
- × **CS** [2]: unconstrained; still less efficient.
- × **Point**: best passive constrained sensing.

*To fully represent CASS, we show different choices of T given *a priori* (including the true value).

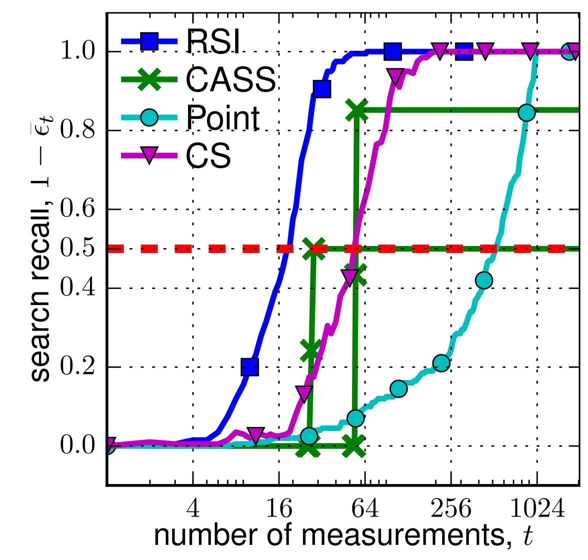


Figure 2: Search progress

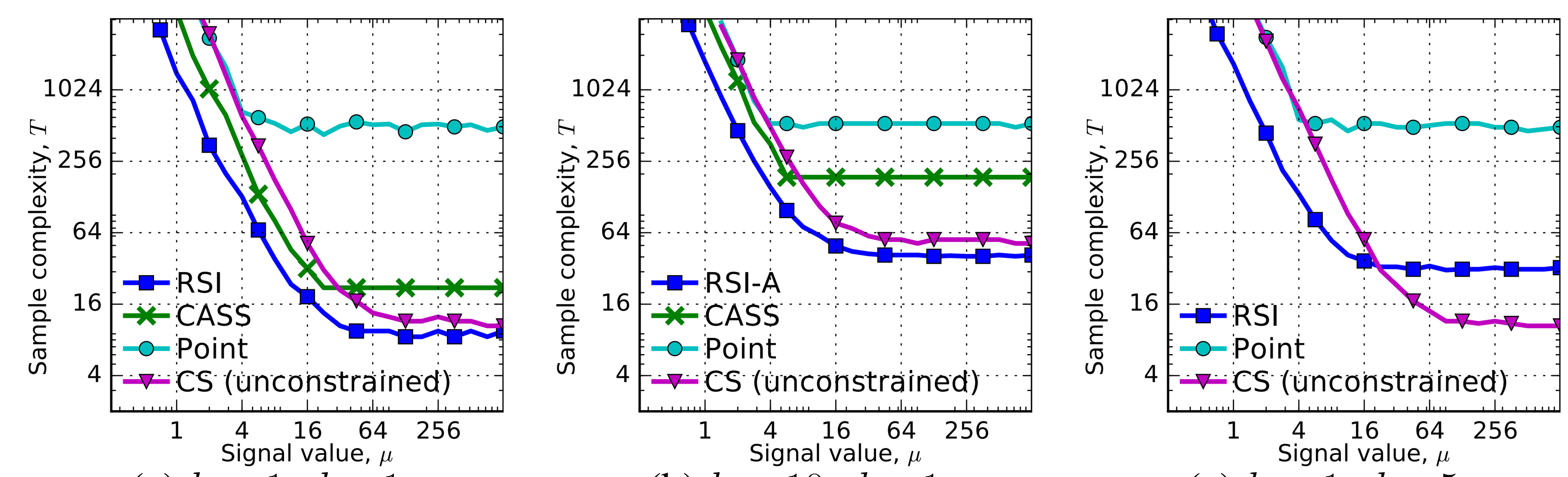


Figure 3: Number of measurements to achieve Bayes error of 0.5 or less. Fix $n = 1024$.

- ✓ **RSI** usually* uses the least number of measurements of order $T = \tilde{O}(\frac{n}{\mu^2})$.
- × **CASS** requires choosing T appropriately to achieve the comparable (albeit worse) performance; requires $d = 1$.
- × **Point** passive region sensing has worse rates, $T = O(n)$.
- × **CS** passive *unconstrained* sensing can also get near-optimal $T = \tilde{O}(\frac{n}{\mu^2})$ [3].

*The only exception in $d = 5$ seems due to a restriction that regions have to choose from a hierarchical sequence of increasingly finer grid boxes with dyadic side lengths.

Real-World Datasets

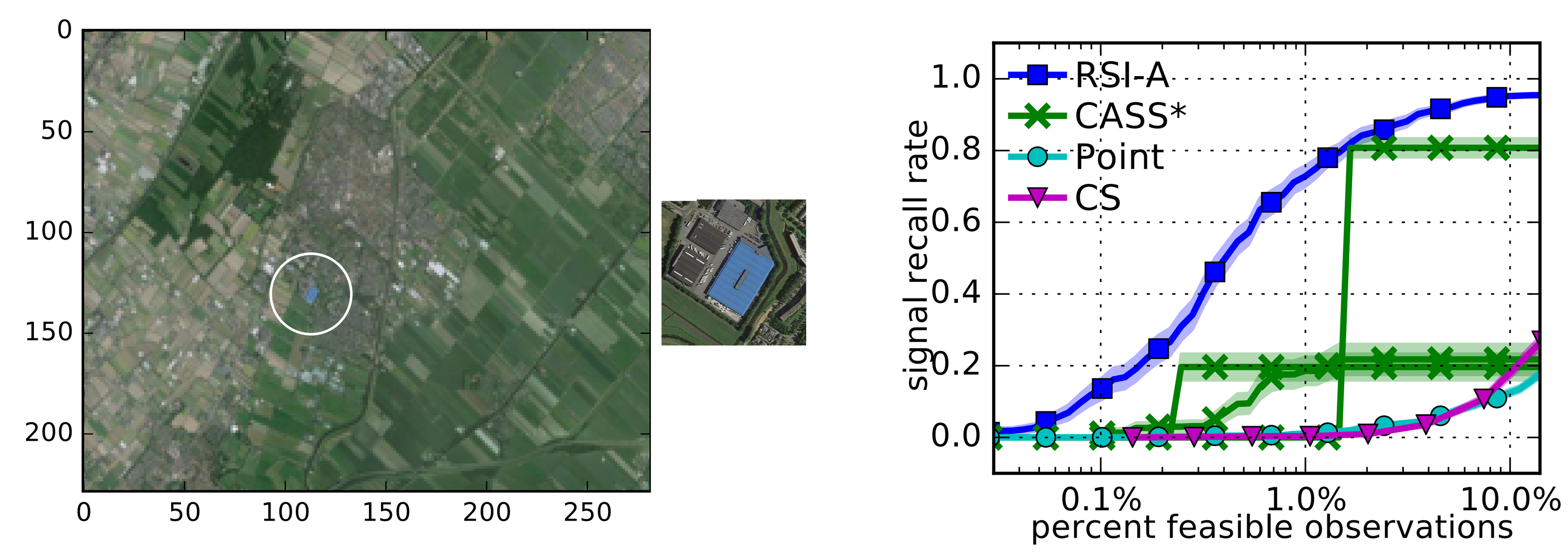


Figure 4: Active search for blue colors on real satellite images. Left, an example image and the signals to search for (circled). Right, search progress as more measurements are taken.

- ❑ Simulate search and rescue in open areas based on life jacket colors.
- ❑ Used a blue filter on the RGB values, yielding the heatmap in Figure 1.
- ❑ Performance RSI (ours) > CASS ≫ CS ≈ Point

Discussion and Related Work

Method / Paper	Region sensing	Sparse signals	Active method	Anytime search	Robust to non-iid noise
Bayesian optimization [4]	×	×	✓	✓	✓
Compressive sensing [2]	×*	✓**	×	×	×
CASS [1]	✓	✓	✓	×†	×‡
RSI (ours)	✓	✓	✓	✓	✓

* Compressive sensing requires unconstrained sensing that does not incorporate region constraints.
** "No active method (e.g. [1]) can *fundamentally* improve sample efficiency beyond logarithmic factors" — [3]. This is only true in the case of unconstrained sensing.
† CASS is a branch-and-bound algorithm that produces results only near the end.
‡ CASS requires repetitive measurements on the same region to control branching error, which is not practical in static environments.

References

[1] Matthew L Malloy and Robert D Nowak. Near-optimal adaptive compressed sensing. *Information Theory, IEEE Transactions on*, 60(7):4001–4012, 2014.

[2] Bradley Efron, Trevor Hastie, Iain Johnstone, Robert Tibshirani, et al. Least angle regression. *The Annals of statistics*, 32(2):407–499, 2004.

[3] Ery Arias-Castro, Emmanuel J Candes, Mark Davenport, et al. On the fundamental limits of adaptive sensing. *Information Theory, IEEE Transactions on*, 59(1):472–481, 2013.

[4] José Miguel Hernández-Lobato, Matthew W Hoffman, and Zoubin Ghahramani. Predictive entropy search for efficient global optimization of black-box functions. In *Advances in Neural Information Processing Systems*, 2014.