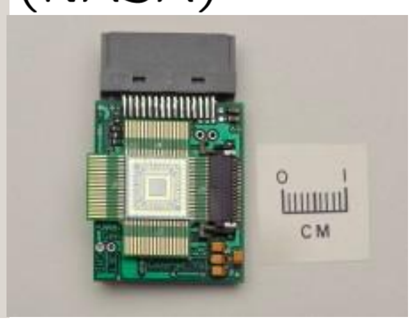
Active Sensing

Background and Motivation

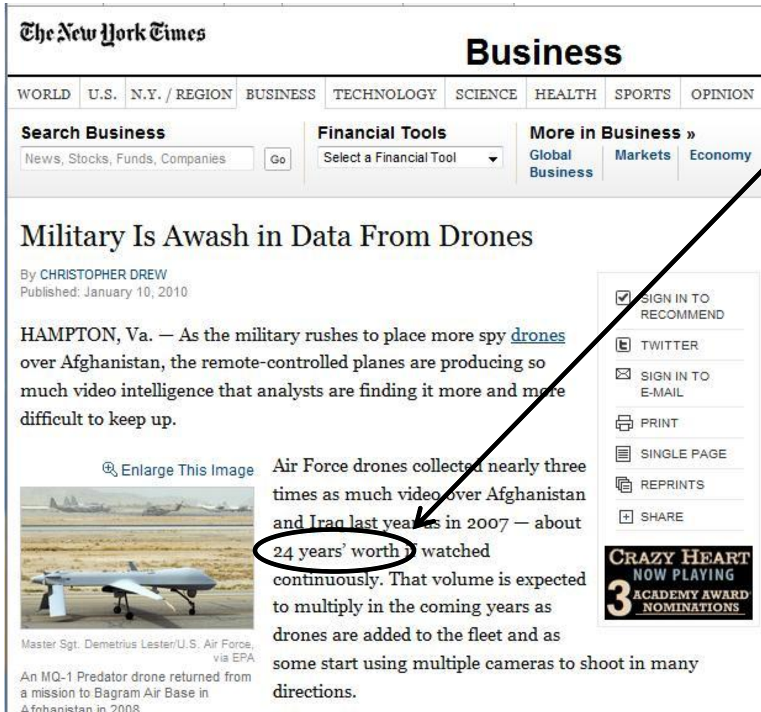
Sensors, Sensors Everywhere

Images, sound, GPS,
accelerometer, proximity,...
(Apple)

NH_3, Cl_2,...
(NASA)



Inevitable Data Deluge!



The New York Times

Business

WORLD | U.S. | N.Y. / REGION | BUSINESS | TECHNOLOGY | SCIENCE | HEALTH | SPORTS | OPINION

Search Business — News, Stocks, Funds, Companies — Go

Financial Tools — Select a Financial Tool

More in Business » Global Business | Markets | Economy

Military Is Awash in Data From Drones

By CHRISTOPHER DREW
Published: January 10, 2010

HAMPTON, Va. — As the military rushes to place more spy drones over Afghanistan, the remote-controlled planes are producing so much video intelligence that analysts are finding it more and more difficult to keep up.

Enlarge This Image

Master Sgt. Demetrius Lester/U.S. Air Force, via EPA

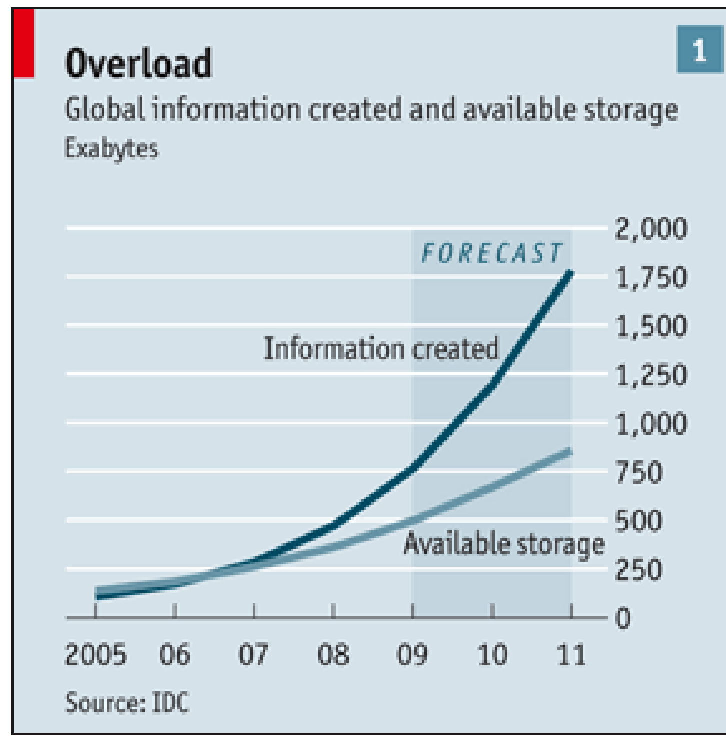
An MQ-1 Predator drone returned from a mission to Bagram Air Base in Afghanistan in 2008.

Air Force drones collected nearly three times as much video over Afghanistan and Iraq last year as in 2007 — about 24 years' worth if watched continuously. That volume is expected to multiply in the coming years as drones are added to the fleet and as some start using multiple cameras to shoot in many directions.

SIGN IN TO RECOMMEND
TWITTER
SIGN IN TO E-MAIL
PRINT
SINGLE PAGE
REPRINTS
SHARE

24 years/year !

YouTube Upload Rate (March 2010):
24 hours/minute ?!?!
(www.webpronews.com)



The Economist, February 2010

Challenges for Sensing/Processing Systems

Technology:

technologically impossible to sense/observe everything, everywhere, all the time
⇒incomplete, missing, or indirect data are the norm

Uncertainty:

experiments/measurements are noisy, corrupted or unreliable!
⇒info-processing and decision-making must be robust to uncertainty

Complexity:

systems can be ultra high-dimensional
⇒modeling/approximation is formidable, mathematically & computationally

Diversity:

data from disparate sources
⇒integration of info from sensors, experiments, databases, human intel, etc.

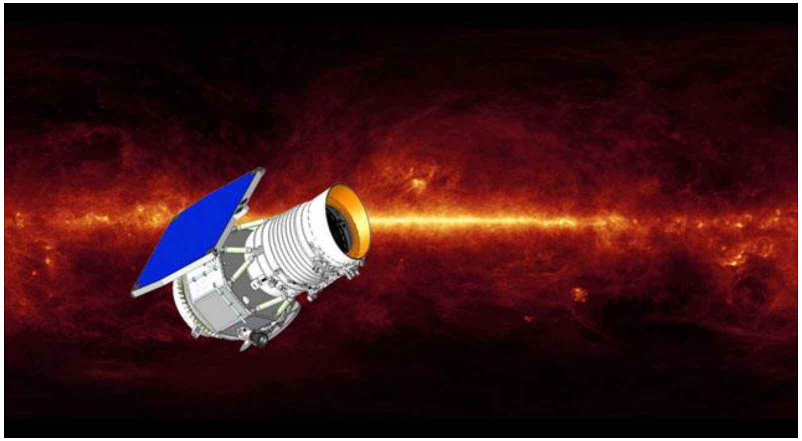
Approaches:

⇒Rethink Traditional Sensing Strategies
⇒Integration of Sensing and Processing

Active Sensing

-- A Few Examples --

Wide-field Infrared Survey Explorer (WISE)



→ Need to shield IR (heat) from its own instruments

→ Sensitive instruments housed in solid hydrogen

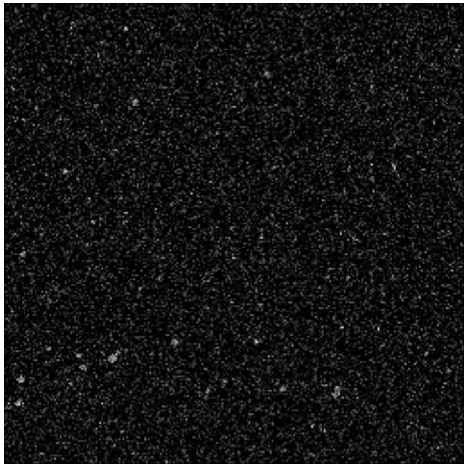
→ Expected lifetime: 10 months!

WISE Mission:
(http://www.nasa.gov/mission_pages/WISE/mission/index.html)

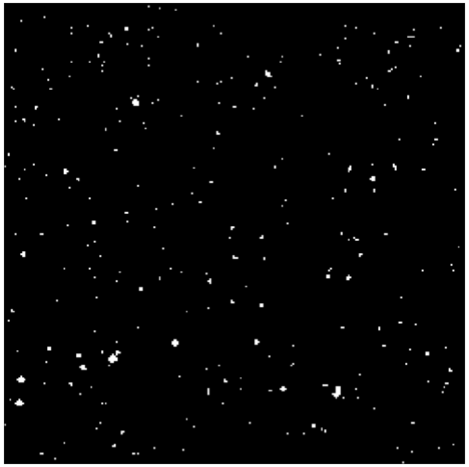
"…the infrared surveyor will spend six months mapping the whole sky. It will then begin a second scan to uncover even more objects and to look for any changes in the sky that might have occurred since the first survey. This second partial sky survey will end about three months later when the spacecraft's frozen-hydrogen cryogen runs out…"

Fornax Galaxy Cluster, Feb. 17 2010

Astronomical Imaging "On a Budget"

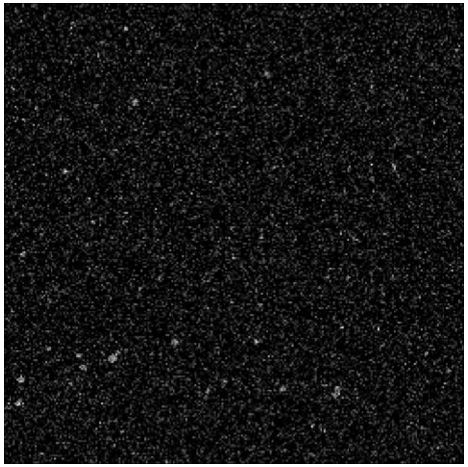


Noisy, non-adaptive sampling

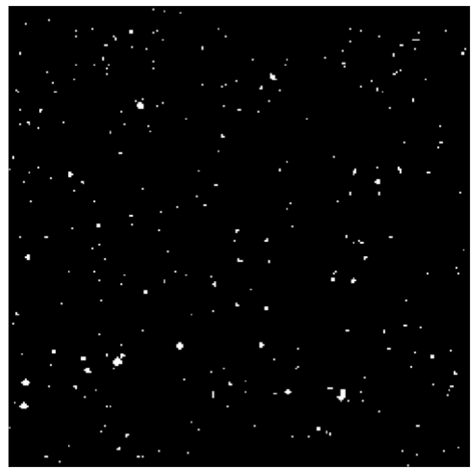


original signal
(~0.8% non-zero components)

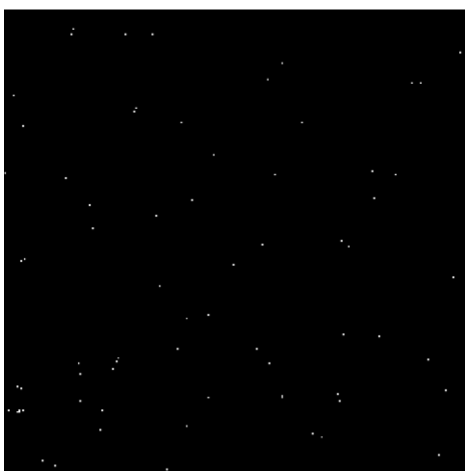
Astronomical Imaging “On a Budget”



Noisy, non-adaptive sampling

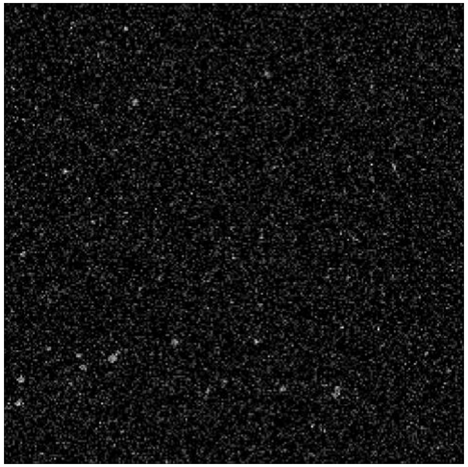


original signal
(~0.8% non-zero components)

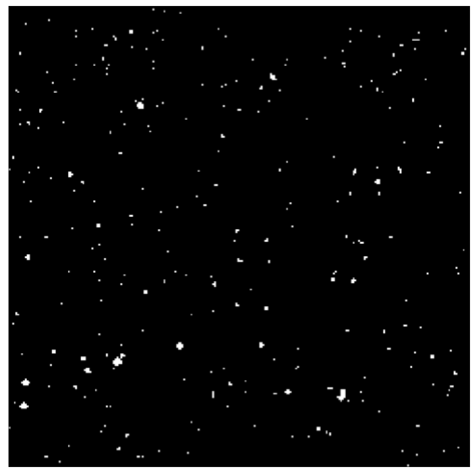


Recovery from non-adaptive samples
(1/20 “discoveries” are errors)

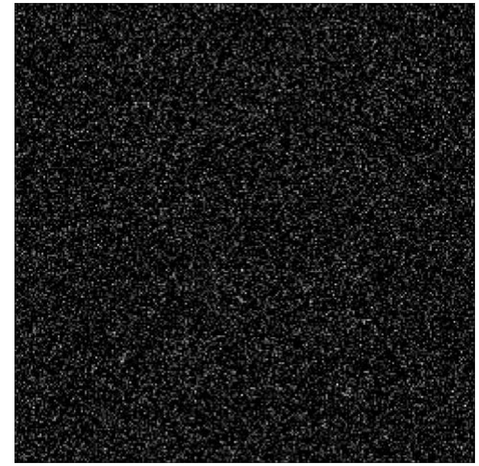
Astronomical Imaging “On a Budget”



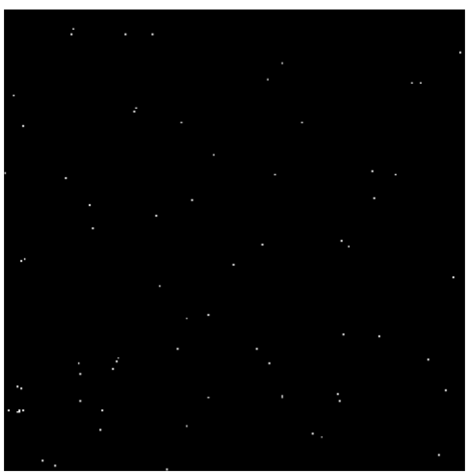
Noisy, non-adaptive sampling



original signal
(~0.8% non-zero components)

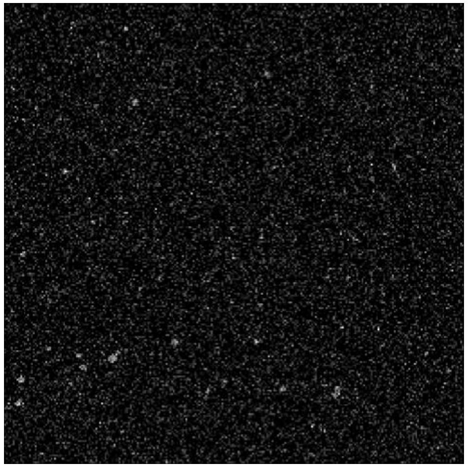


Noisy adaptive sampling

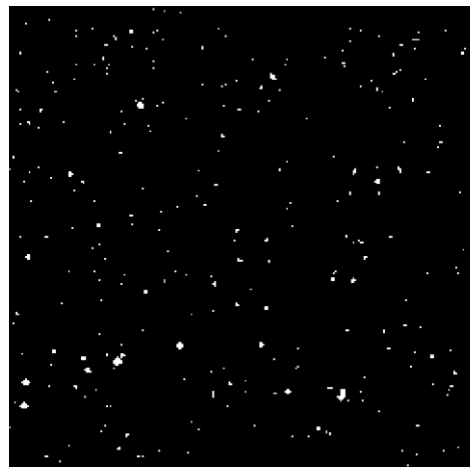


Recovery from non-adaptive samples
(1/20 “discoveries” are errors)

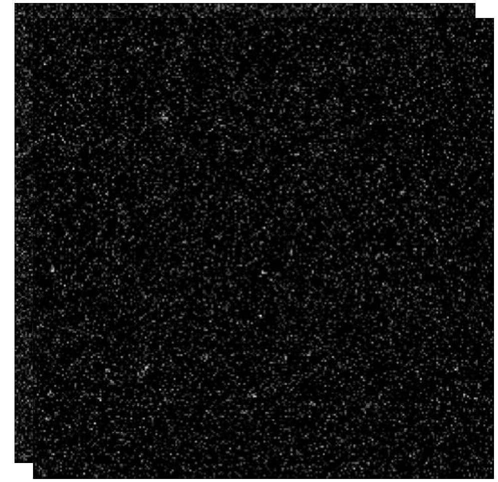
Astronomical Imaging "On a Budget"



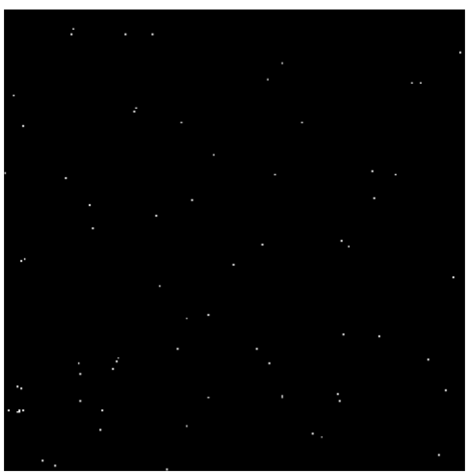
Noisy, non-adaptive sampling



original signal
(~0.8% non-zero components)

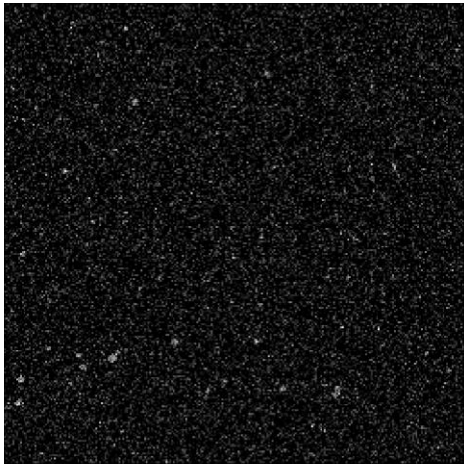


Noisy adaptive sampling

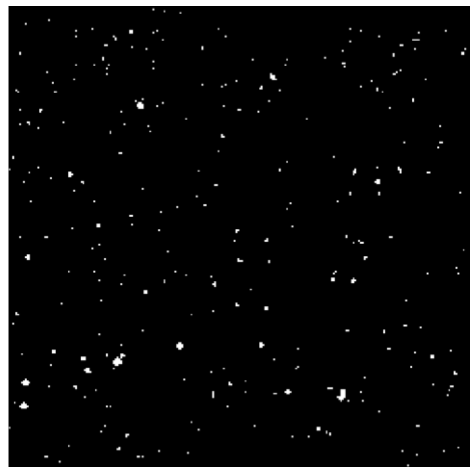


Recovery from non-adaptive samples
(1/20 "discoveries" are errors)

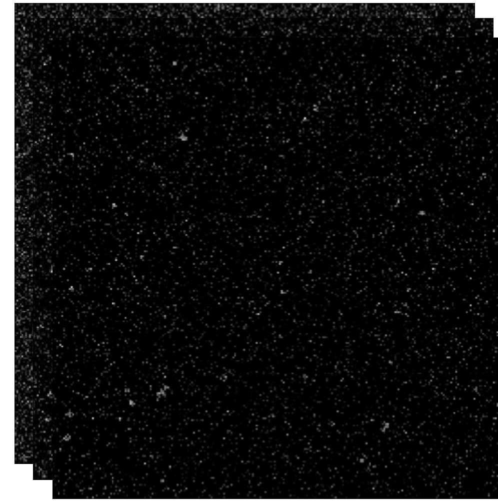
Astronomical Imaging “On a Budget”



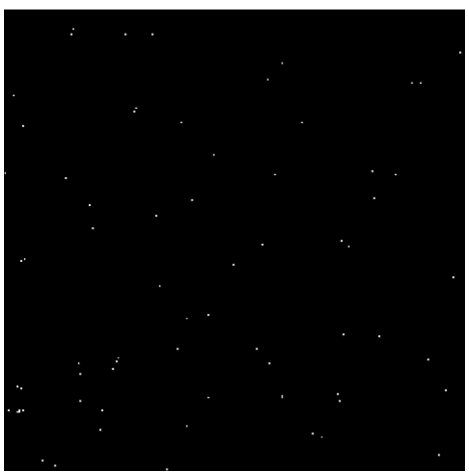
Noisy, non-adaptive sampling



original signal
(~0.8% non-zero components)

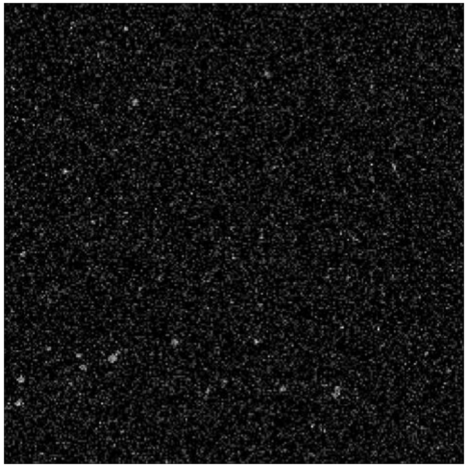


Noisy adaptive sampling

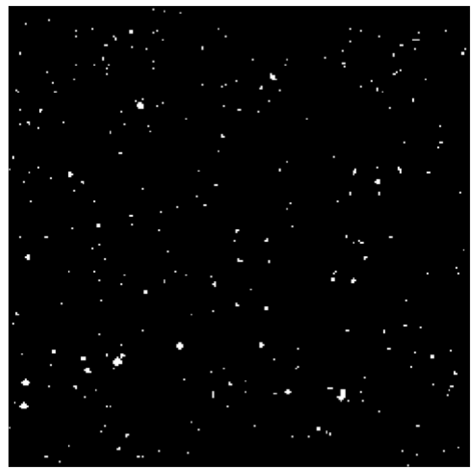


Recovery from non-adaptive samples
(1/20 “discoveries” are errors)

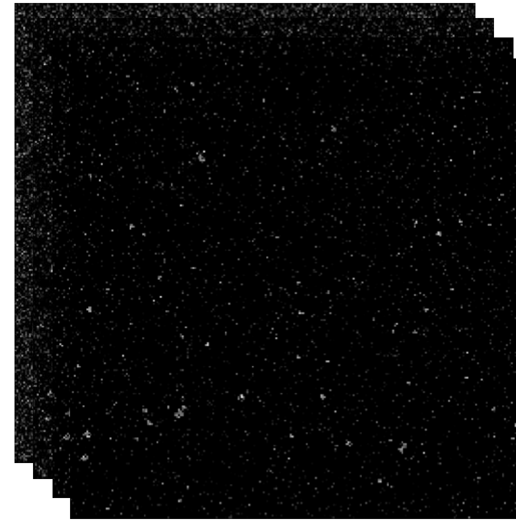
Astronomical Imaging “On a Budget”



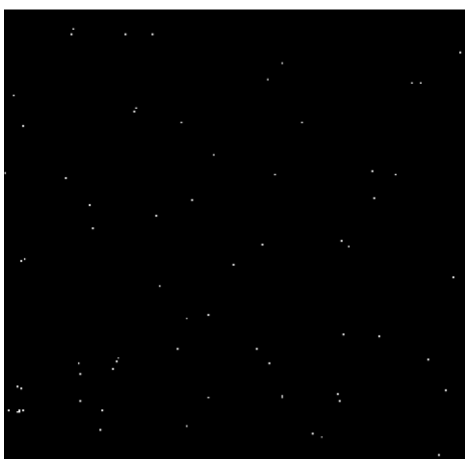
Noisy, non-adaptive sampling



original signal
(~0.8% non-zero components)

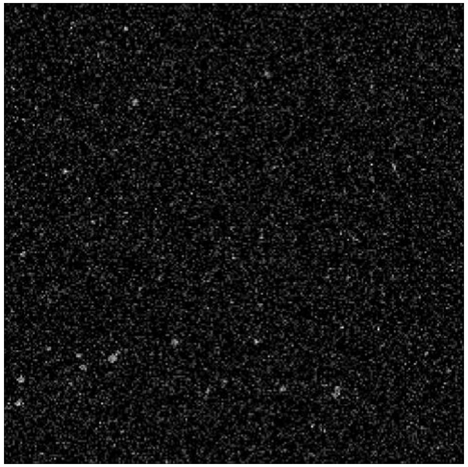


Noisy adaptive sampling

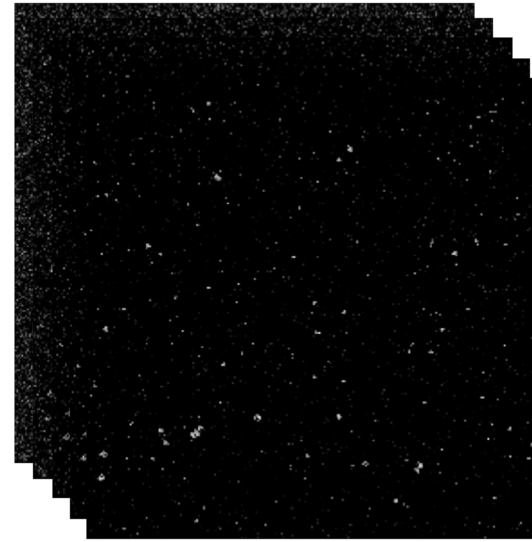


Recovery from non-adaptive samples
(1/20 “discoveries” are errors)

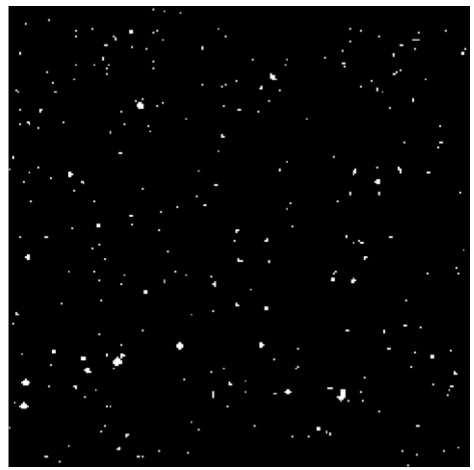
Astronomical Imaging "On a Budget"



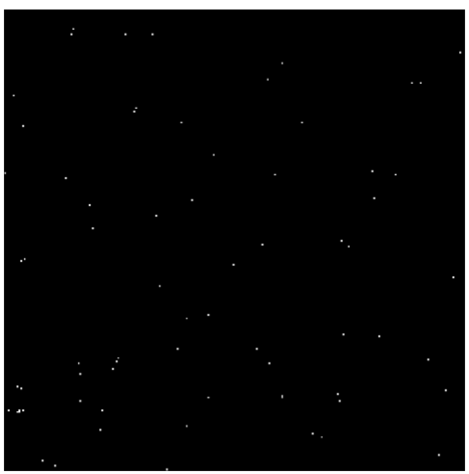
Noisy, non-adaptive sampling



Noisy adaptive sampling

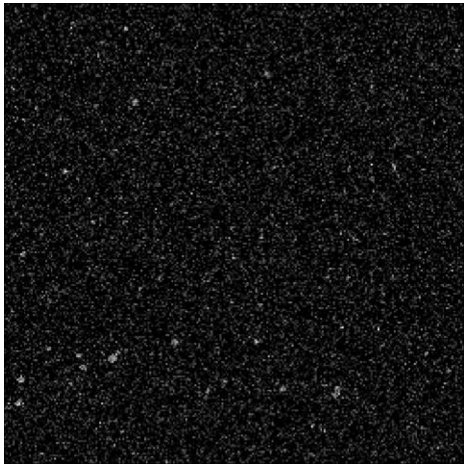


original signal
(~0.8% non-zero components)

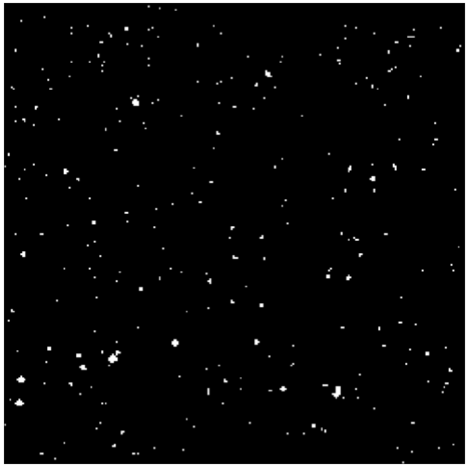


Recovery from non-adaptive samples
(1/20 "discoveries" are errors)

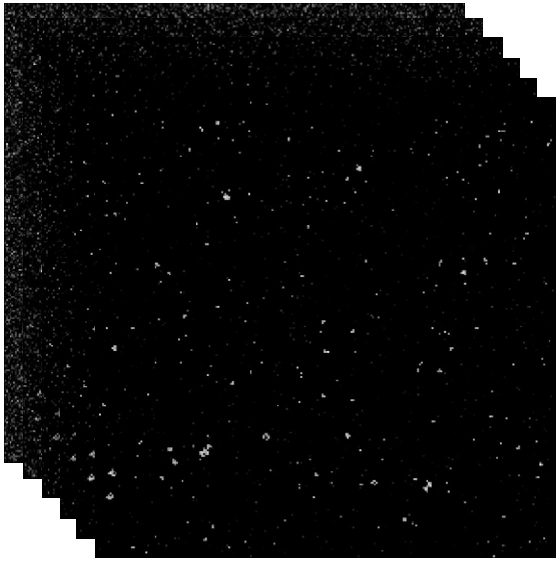
Astronomical Imaging “On a Budget”



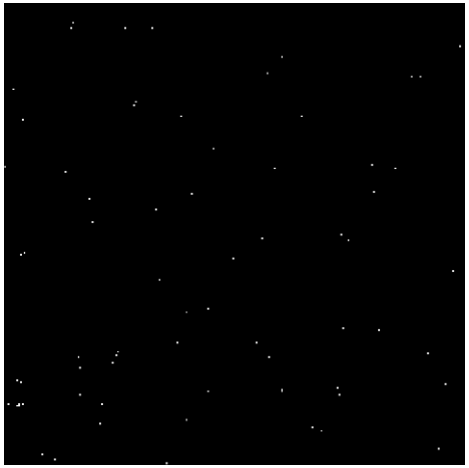
Noisy, non-adaptive sampling



original signal
(~0.8% non-zero components)

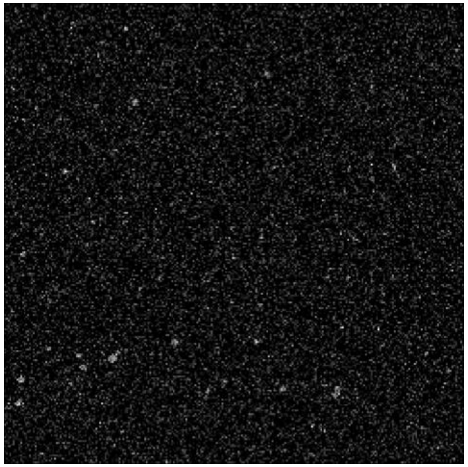


Noisy adaptive sampling

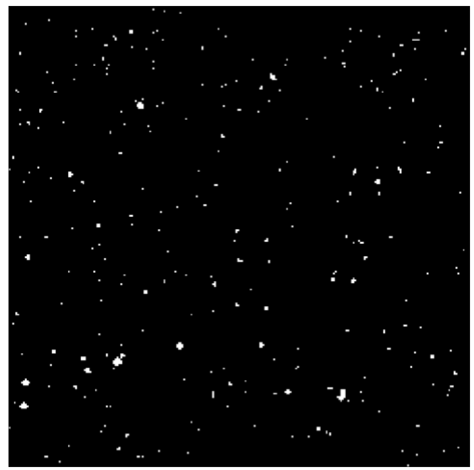


Recovery from non-adaptive samples
(1/20 “discoveries” are errors)

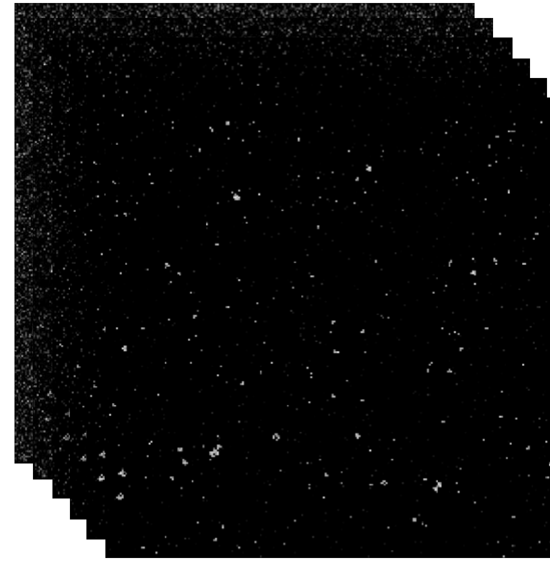
Astronomical Imaging “On a Budget”



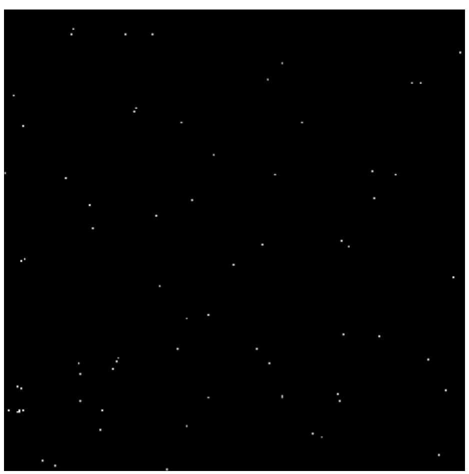
Noisy, non-adaptive sampling



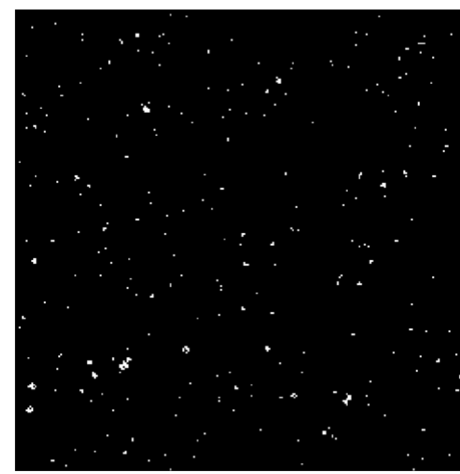
original signal
(~0.8% non-zero components)



Noisy adaptive sampling

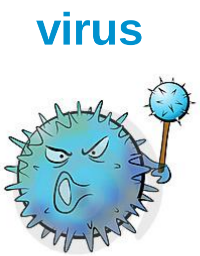


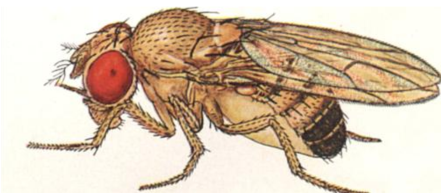
Recovery from non-adaptive samples
(1/20 “discoveries” are errors)



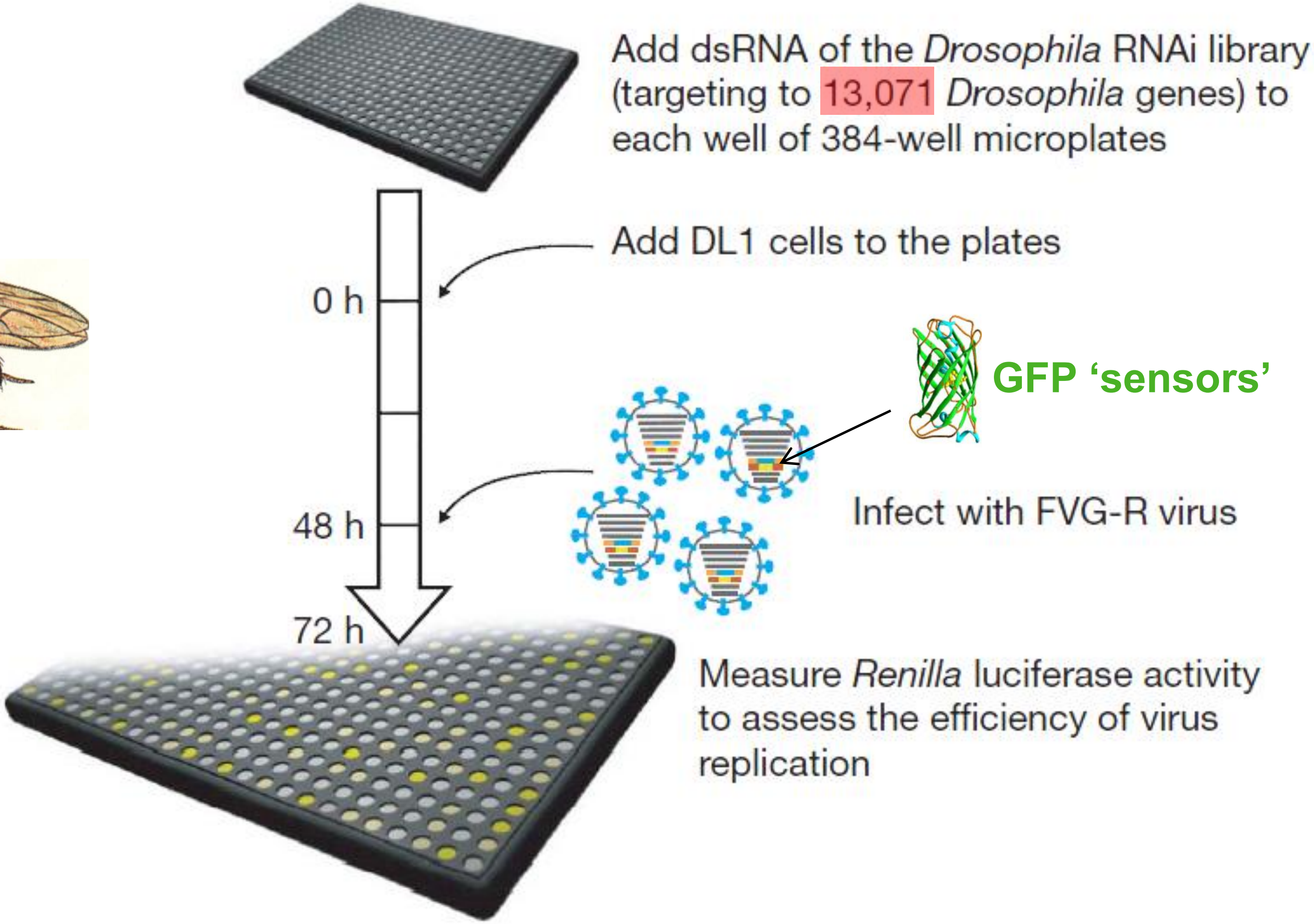
Recovery from adaptive samples
(1/20 “discoveries” are errors)

Functional Genomics: Virus-Host Interaction





fruit fly

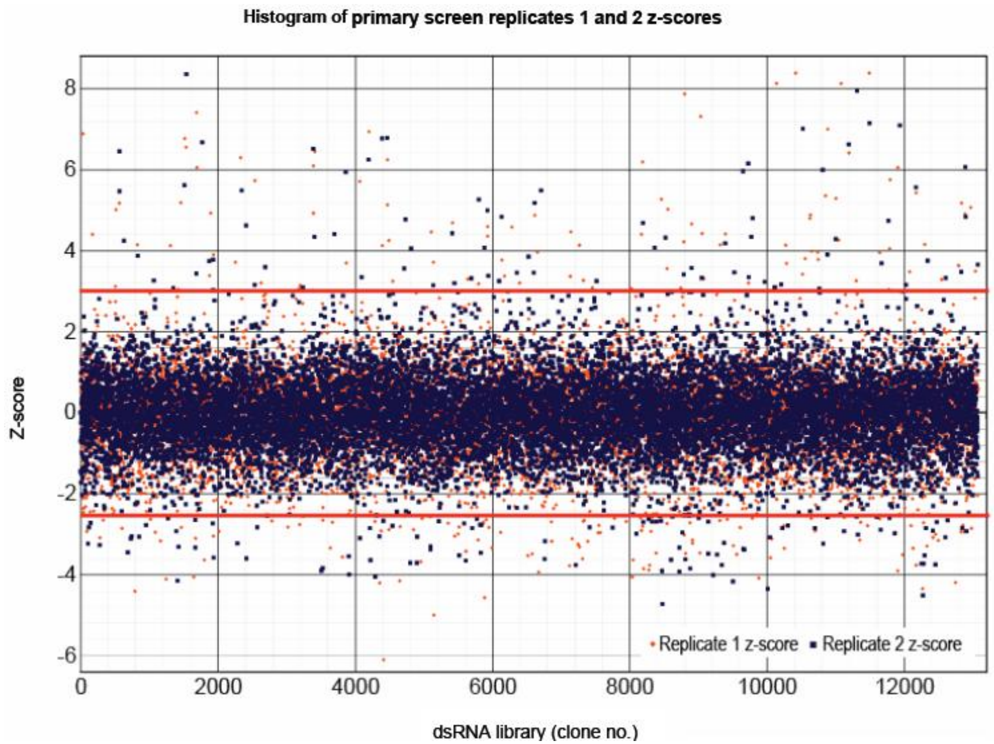


L. Hao, et al., "Drosophila RNAi Screen Identifies Host Genes Important for Influenza Virus Replication," Nature, 2008.

Adaptive Experimentation

How to confidently determine ~ 100 out of $\sim 13k$ genes that are hijacked for virus replication from extremely noisy data?

Stage 1: Assay all $13k$ genes twice, keep all with significant fluorescence in one or both assays for 2nd stage ($13k \rightarrow 299$)



z-score ≥ 3
(increased luciferase expression)

z-score ≤ -2.5
(inhibited luciferase expression)

Stage 2: Assay remaining genes multiple times, retain only those with statistically significant fluorescence in at least one of the trials ($299 \rightarrow 112$)

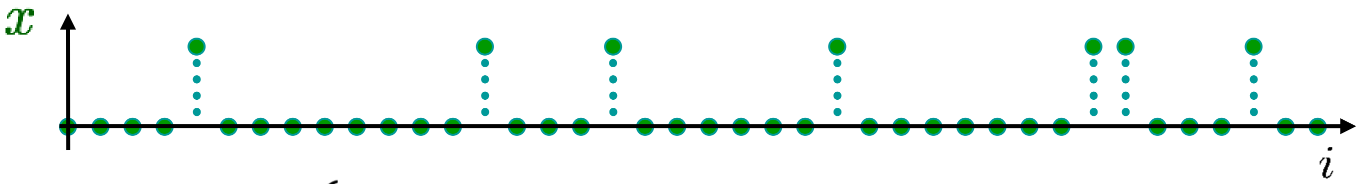
See also related work on two-stage and multi-stage methods in gene expression studies:
(Satagopan and Elston 2003; Zehetmayer, Bauer, & Posch 2005; Muller, Pahl, & Schafer 2007; Zehetmayer, Bauer, & Posch 2008)

Sparse Recovery

-- Preliminaries and Formalization --

A Sparse Signal Model

Signals of interest are vectors $x \in \mathbb{R}^n$

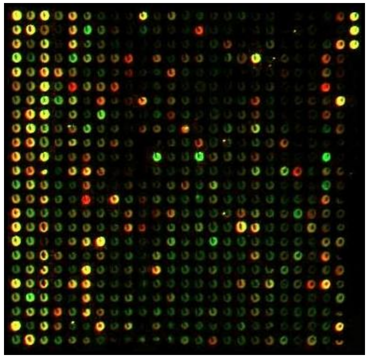


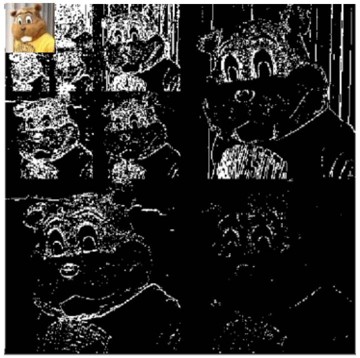
$$x_i = \begin{cases} \mu > 0 & i \in \mathcal{S} \\ 0 & i \notin \mathcal{S} \end{cases}$$

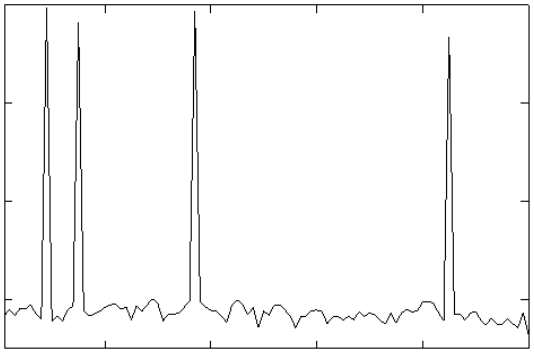
signal support set

$$\text{Sparse} \Leftrightarrow |\mathcal{S}| \ll n$$

number of nonzero
signal components







Observation Model

$$y_{i,j} = \begin{cases} x_i + \gamma_{i,j}^{-1/2}\, z_{i,j}, & \gamma_{i,j} > 0,\ i = 1,\ldots,n,\quad j = 1,\ldots,k \\ 0 & \gamma_{i,j} = 0,\ i = 1,\ldots,n,\quad j = 1,\ldots,k \end{cases}$$

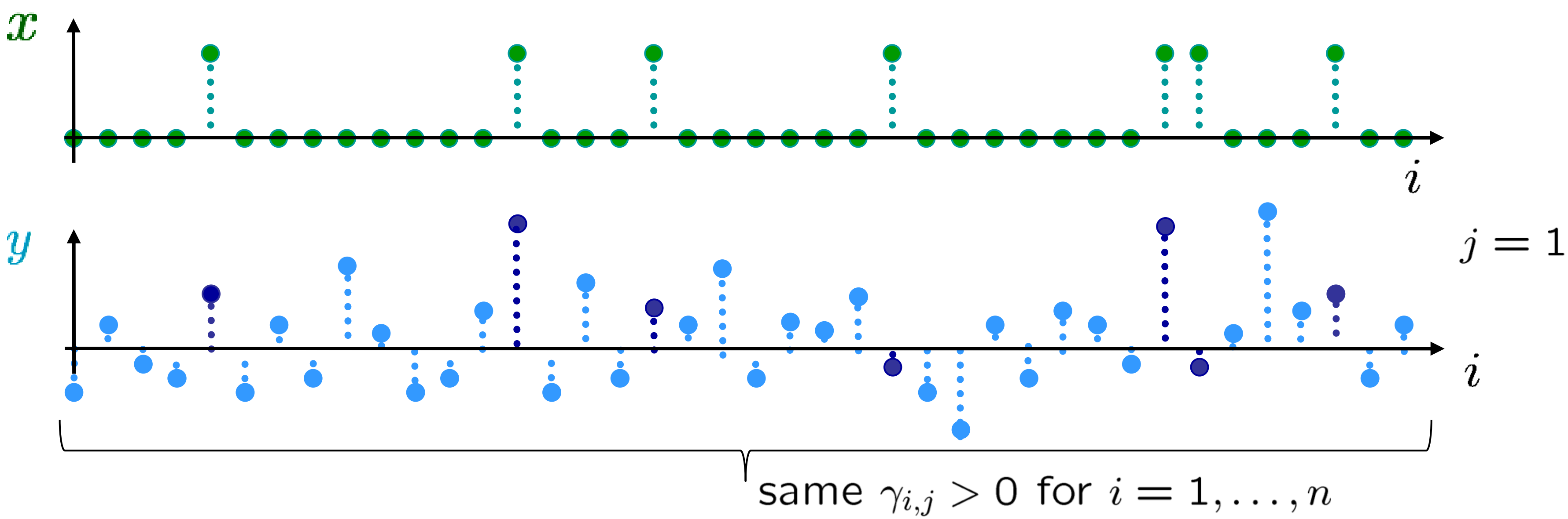
$z_{i,j} \overset{\text{iid}}{\sim} \mathcal{N}(0,1)$

j indexes the observation step

k is the total number of steps

$\gamma_{i,j} \geq 0$ is the *precision* of observation $y_{i,j}$

$$\Rightarrow y_{i,j} \sim \mathcal{N}\left(x_i, 1/\gamma_{i,j}\right) \text{ when } \gamma_{i,j} \neq 0$$



Observation Model

$$y_{i,j} = \begin{cases} x_i + \gamma_{i,j}^{-1/2}\, z_{i,j}, & \gamma_{i,j} > 0,\ i = 1,\ldots,n,\quad j = 1,\ldots,k \\ 0 & \gamma_{i,j} = 0,\ i = 1,\ldots,n,\quad j = 1,\ldots,k \end{cases}$$

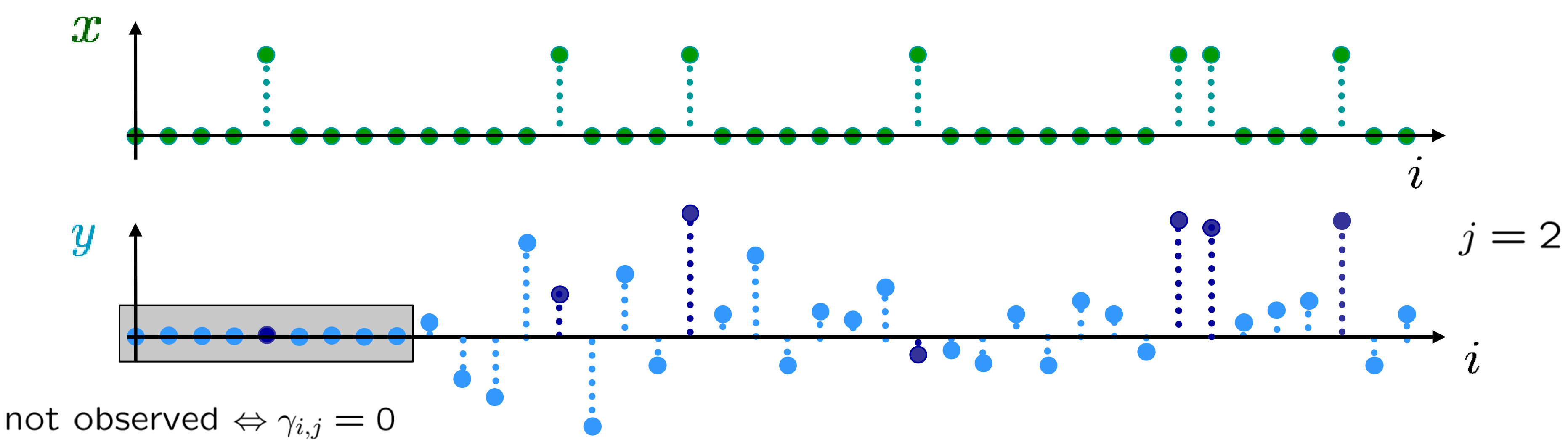
$z_{i,j} \overset{\text{iid}}{\sim} \mathcal{N}(0,1)$

j indexes the observation step

k is the total number of steps

$\gamma_{i,j} \geq 0$ is the *precision* of observation $y_{i,j}$

$\Rightarrow y_{i,j} \sim \mathcal{N}\left(x_i, 1/\gamma_{i,j}\right)$ when $\gamma_{i,j} \neq 0$



Observation Model

$$y_{i,j} = \begin{cases} x_i + \gamma_{i,j}^{-1/2}\, z_{i,j}, & \gamma_{i,j} > 0,\ i = 1, \ldots, n, \quad j = 1, \ldots, k \\ 0 & \gamma_{i,j} = 0,\ i = 1, \ldots, n, \quad j = 1, \ldots, k \end{cases}$$

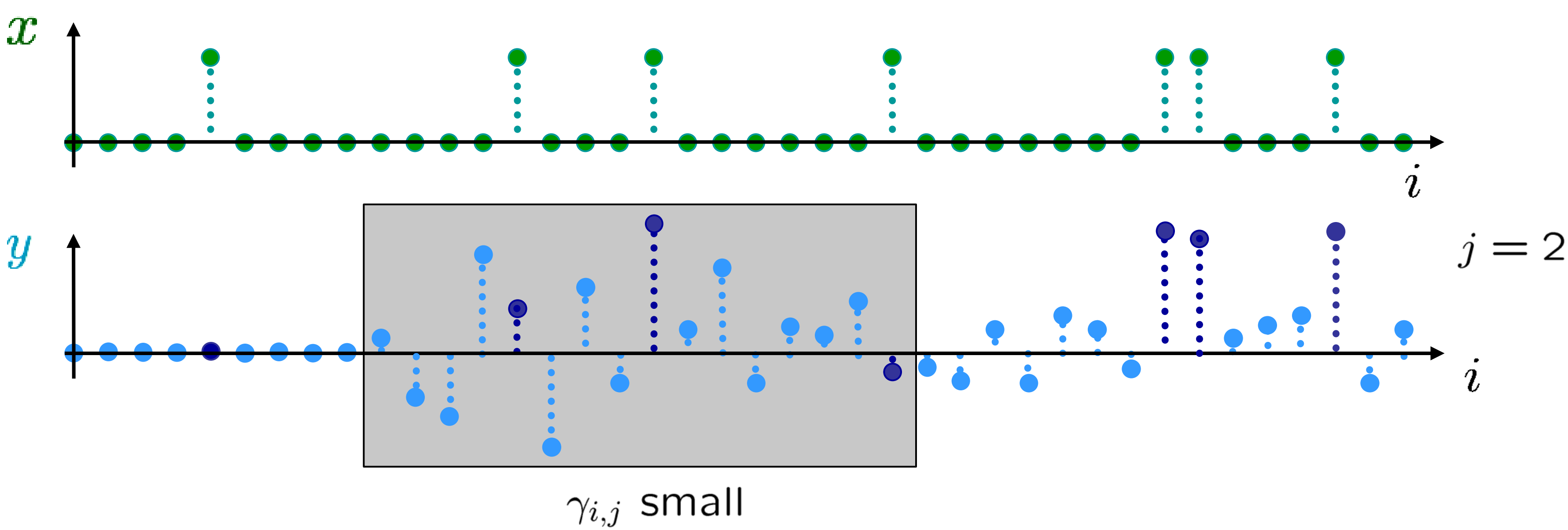
$z_{i,j} \overset{\text{iid}}{\sim} \mathcal{N}(0,1)$

j indexes the observation step

k is the total number of steps

$\gamma_{i,j} \geq 0$ is the *precision* of observation $y_{i,j}$

$\Rightarrow y_{i,j} \sim \mathcal{N}\left(x_i, 1/\gamma_{i,j}\right)$ when $\gamma_{i,j} \neq 0$



Observation Model

$$y_{i,j} = \begin{cases} x_i + \gamma_{i,j}^{-1/2}\, z_{i,j}, & \gamma_{i,j} > 0,\ i = 1, \dots, n, \quad j = 1, \dots, k \\ 0 & \gamma_{i,j} = 0,\ i = 1, \dots, n, \quad j = 1, \dots, k \end{cases}$$

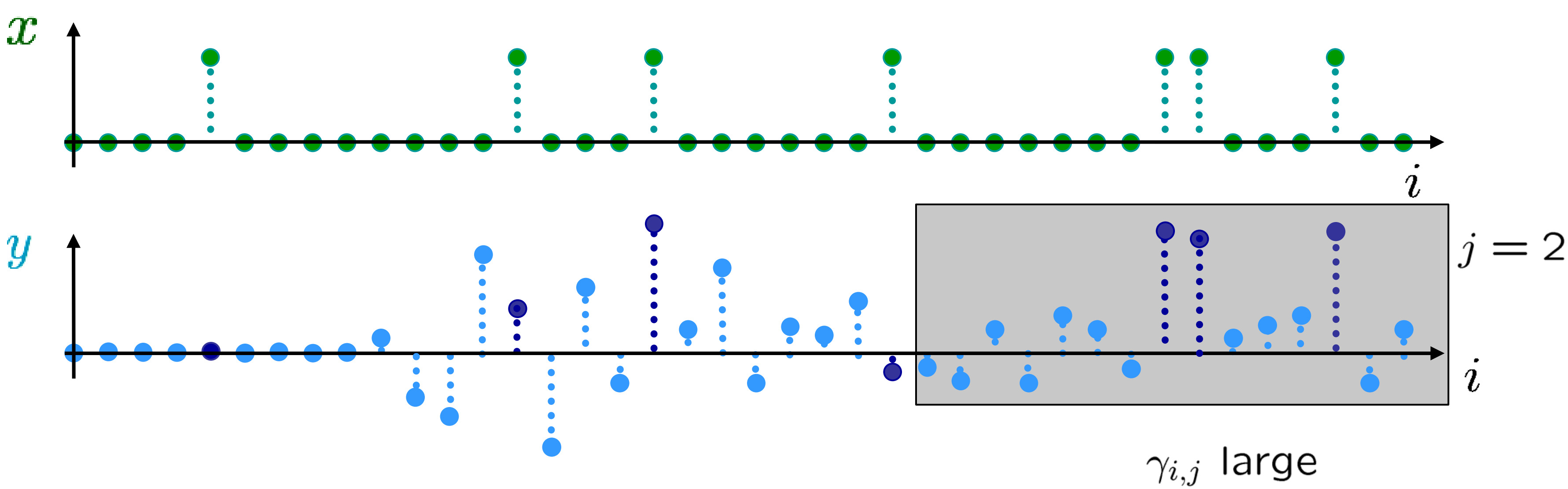
$z_{i,j} \overset{\text{iid}}{\sim} \mathcal{N}(0,1)$

j indexes the observation step

k is the total number of steps

$\gamma_{i,j} \geq 0$ is the *precision* of observation $y_{i,j}$

$\Rightarrow y_{i,j} \sim \mathcal{N}\left(x_i, 1/\gamma_{i,j}\right)$ when $\gamma_{i,j} \neq 0$



Observation Model

$$y_{i,j} = \begin{cases} x_i + \gamma_{i,j}^{-1/2} \, z_{i,j}, & \gamma_{i,j} > 0, \ i = 1, \ldots, n, \quad j = 1, \ldots, k \\ 0 & \gamma_{i,j} = 0, \ i = 1, \ldots, n, \quad j = 1, \ldots, k \end{cases}$$

$z_{i,j} \stackrel{\text{iid}}{\sim} \mathcal{N}(0,1)$

j indexes the observation step

k is the total number of steps

$\gamma_{i,j} \geq 0$ is the *precision* of observation $y_{i,j}$

$$\Rightarrow y_{i,j} \sim \mathcal{N}\left(x_i, 1/\gamma_{i,j}\right) \text{ when } \gamma_{i,j} \neq 0$$

Precision is increased (decreased) by:
– Averaging more (fewer) repeated samples
– Longer (shorter) observation times

Total precision subject to a global constraint:

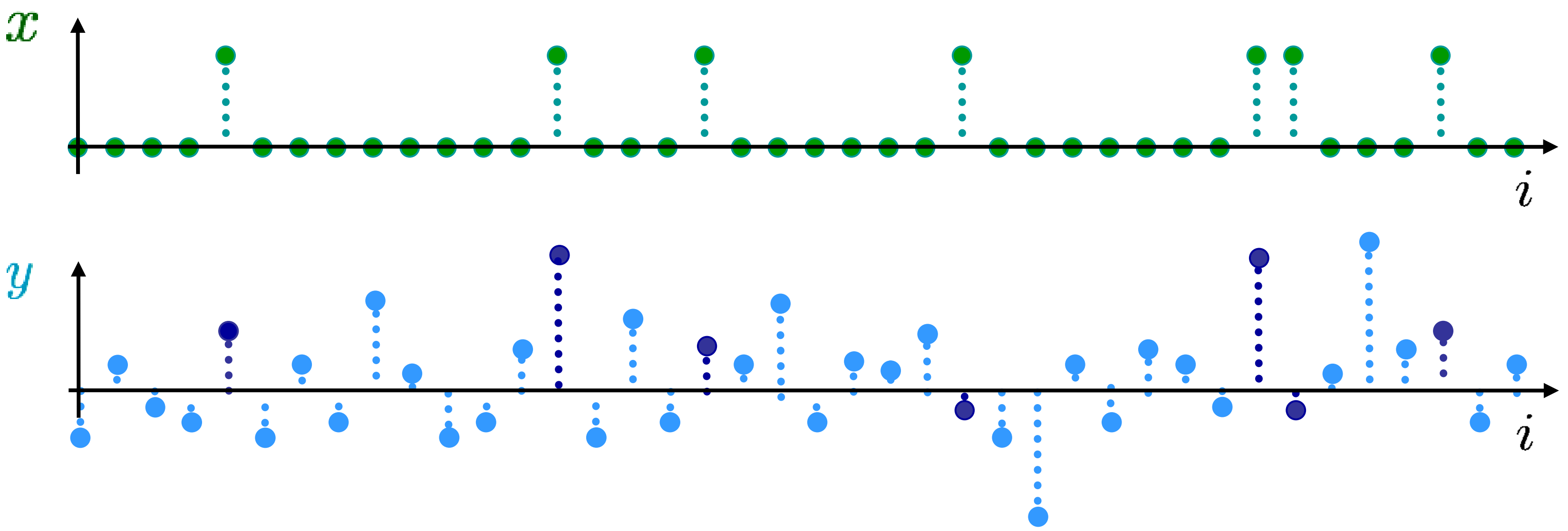
$$\sum_{j=1}^{k} \sum_{i=1}^{n} \gamma_{i,j} \leq R(n)$$

Proportional to total # samples, total time, total energy, cryogen life, etc.

Non-adaptive Sampling

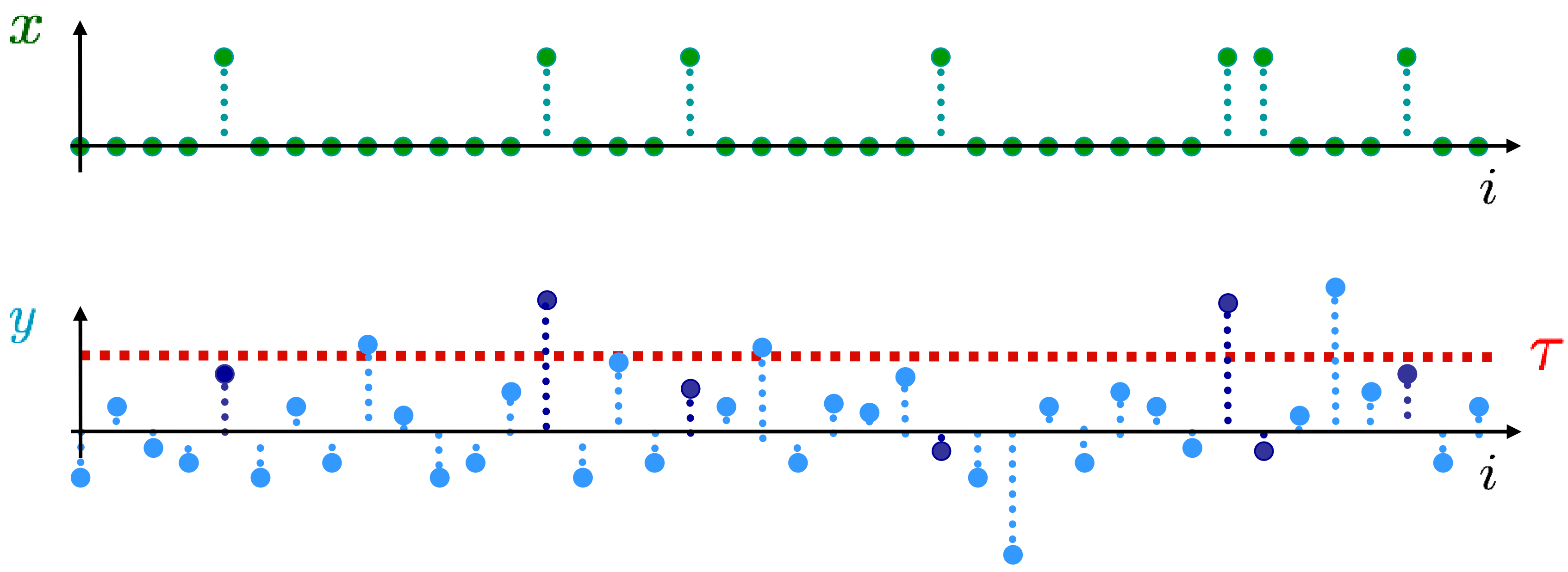
Non-adaptive observations:

$$\begin{array}{l} y_i = x_i + z_i \\ z_i \overset{\text{iid}}{\sim} \mathcal{N}(0,1) \end{array} \quad \Leftrightarrow \quad \left\{ \begin{array}{cc} k = 1 & \\ \gamma_{i,1} = 1, & i = 1, \ldots, n \\ R(n) = n & \end{array} \right\}$$



Support Recovery

Goal: Estimate the signal support set $\mathcal{S} := \{i \in \{1, \ldots, n\} : x_i \neq 0\}$



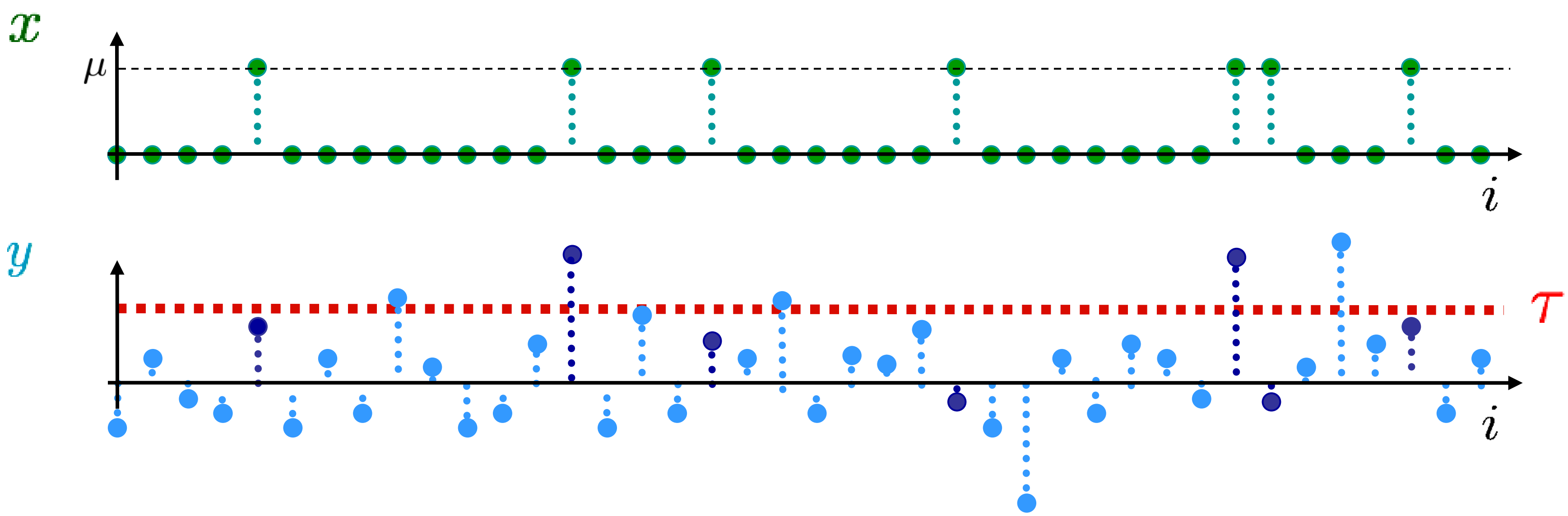
Definition: A *threshold test* is an estimator of the form

$$\widehat{\mathcal{S}}_\tau(y) := \{i \in \{1, \ldots, n\} : y_i > \tau\}$$

A Simple Active Sensing Approach

Baseline: Non-Adaptive Recovery

Goal: Estimate $\mathcal{S}$ with *no* errors (ie, $\widehat{\mathcal{S}} = \mathcal{S}$) from noisy measurements
How large must amplitude μ be?

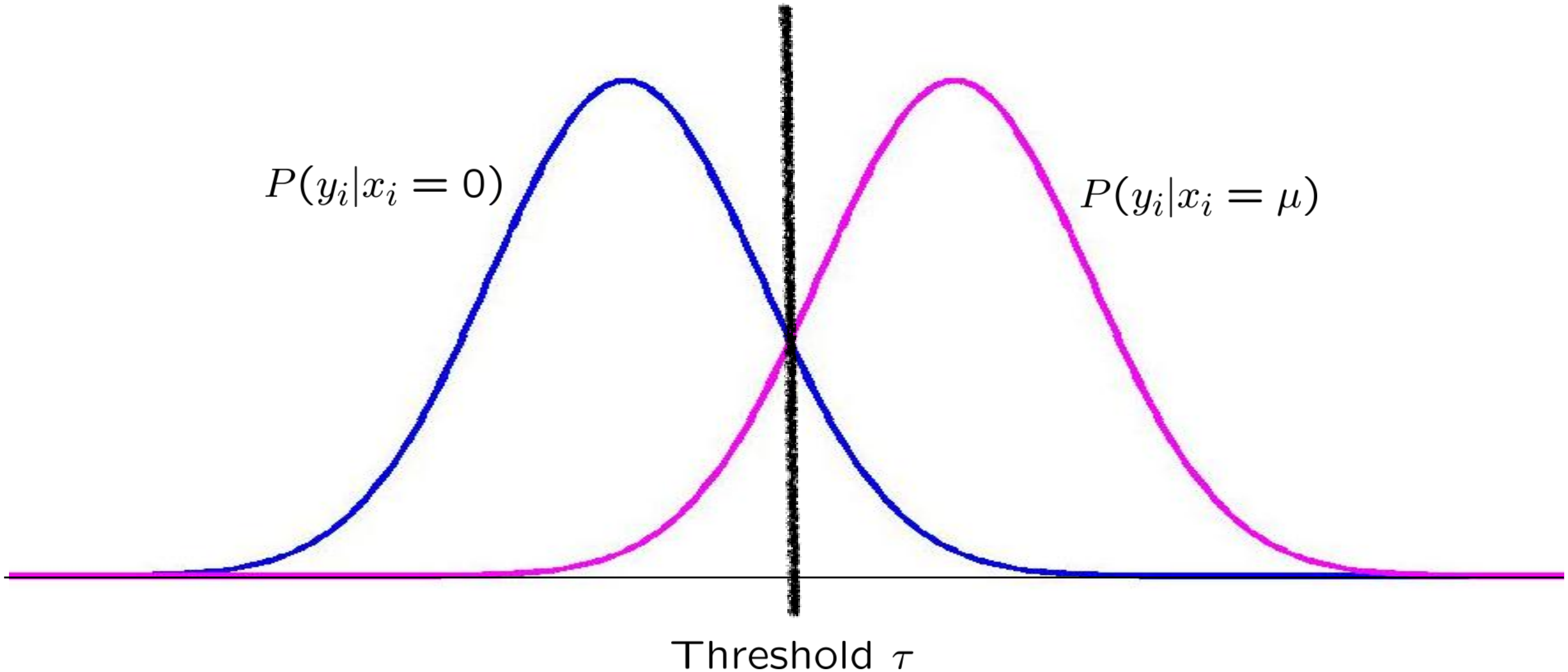


Exact support recovery $\Leftrightarrow \begin{cases} y_i > \tau \text{ for all } i \in \mathcal{S} \\ y_i < \tau \text{ for all } i \in \mathcal{S}^c \end{cases}$

Let $|\mathcal{S}| = s$, then $n - s$ components of x are equal to zero.

Fundamentally a Multiple Hypothesis Test

Test signal present vs. signal absent at each coordinate:



Non-Adaptive Support Recovery

How large must μ be to ensure probability of error tends to zero?

$$\mathbb{P}(\widehat{\mathcal{S}} \neq \mathcal{S}) \leq (n-s)\ \mathbb{P}(y_i > \tau | x_i = 0)\ +\ s\ \mathbb{P}(y_i < \tau | x_i = \mu)$$

$$\leq \underbrace{\frac{n-s}{2} \exp\left(-\frac{\tau^2}{2}\right)} + \underbrace{\frac{s}{2} \exp\left(-\frac{(\mu-\tau)^2}{2}\right)}$$

Want each term to tend to zero

$$\tau \gtrsim \sqrt{2 \log(n-s)} \qquad \mu \gtrsim \tau + \sqrt{2 \log s}$$

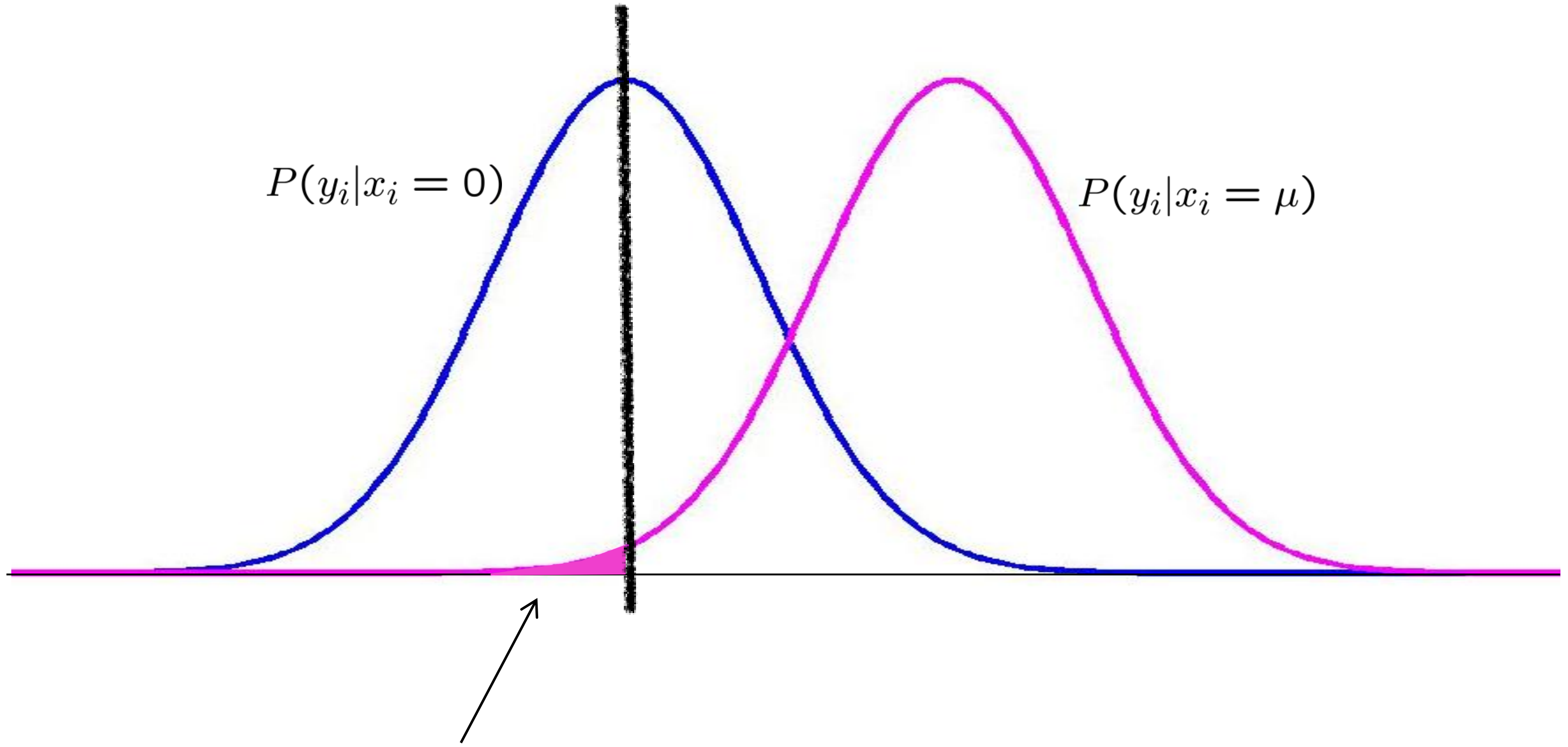
$$\mu \gtrsim \sqrt{2 \log(n-s)} + \sqrt{2 \log s}$$

Necessary condition: $\mu \gtrsim \sqrt{2 \log n}$

Sequential Testing & Refinement

Key Idea: Use a *sequence* of testing and refinement steps

In each step, for each component...

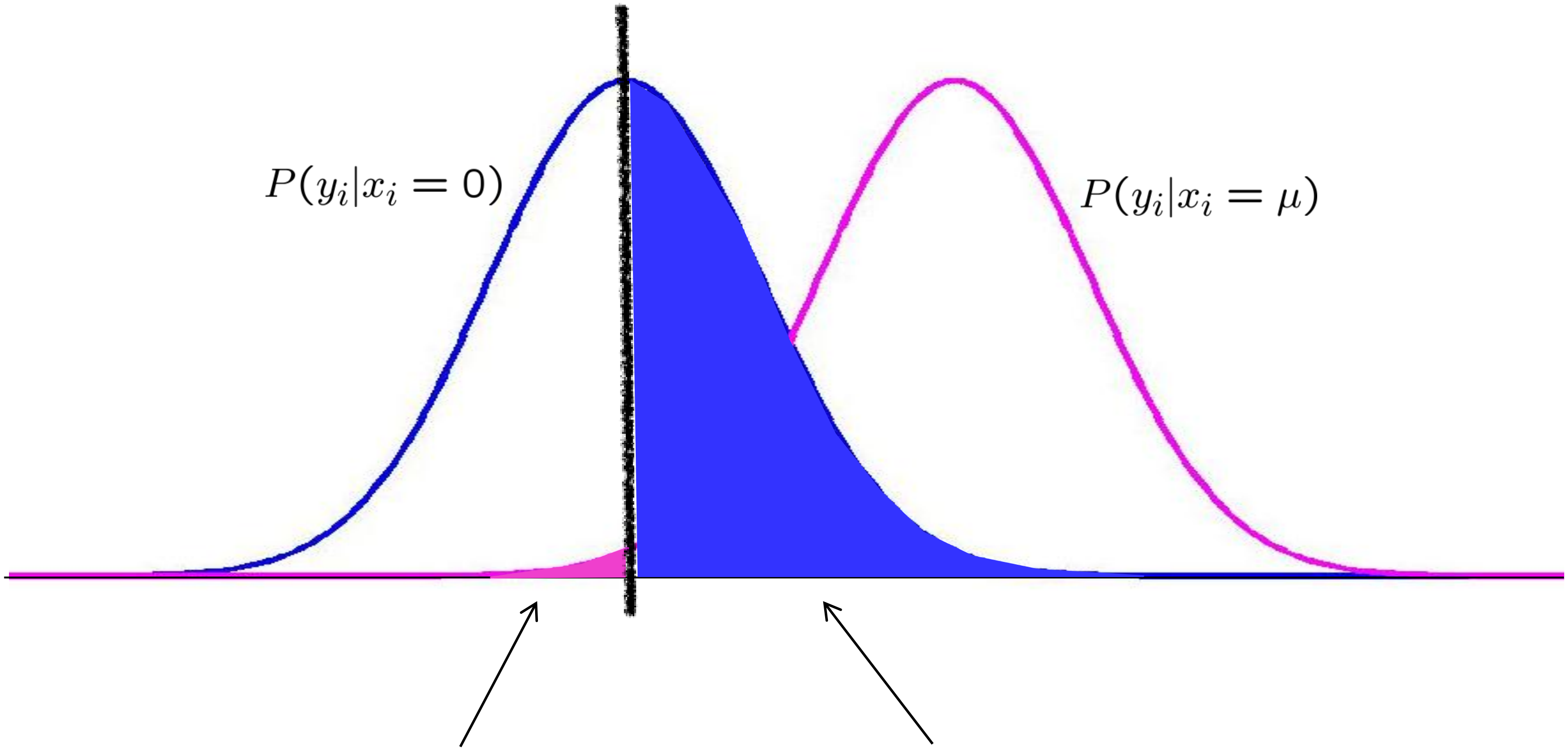


...keep miss probability small...

Sequential Testing & Refinement

Key Idea: Use a *sequence* of testing and refinement steps

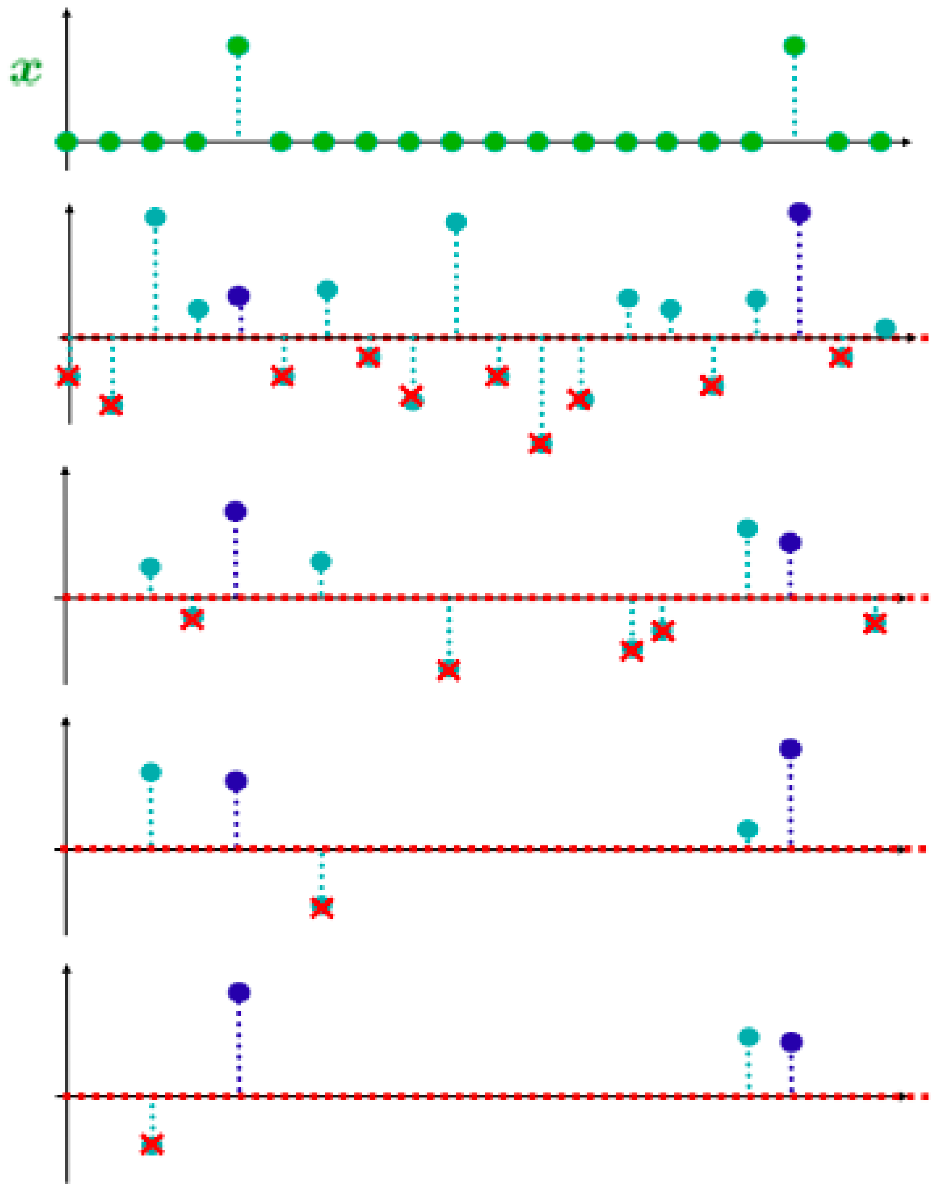
In each step, for each component...



...keep miss probability small...

...but *allow* large false discovery probability

Idealized Example



threshold at zero and re-measure only those components that survive

repeat several times

most of true signal components survive several thresholding steps, almost all of noise components do not

Sequential Thresholding

Sequential Thresholding

initialize: $\mathcal{S}_0 = \{1,\dots,n\}$, $\gamma_{i,j}^{-1} = 2$

for $j = 1,\dots,k$

1) measure: $y_{i,j} \sim \mathcal{N}(x_i, 2)$, $i \in \mathcal{S}_{j-1}$

2) threshold: $\mathcal{S}_j = \{i : y_{i,j} \geq 0\}$

end

output: $\mathcal{S}_k = \{i : y_{i,k} > 0\}$

total precision budget: $\mathbb{E}\left[\sum_{i,j} \gamma_{i,j}\right]$

$$= \frac{1}{2}\sum_{j=1}^{k} \mathbb{E}|\mathcal{S}_{j-1}|$$

$$\leq \frac{1}{2}\sum_{j=1}^{k}\left(\frac{n-s}{2^{j-1}} + s\right)$$

$$\leq n - s + ks \approx n$$

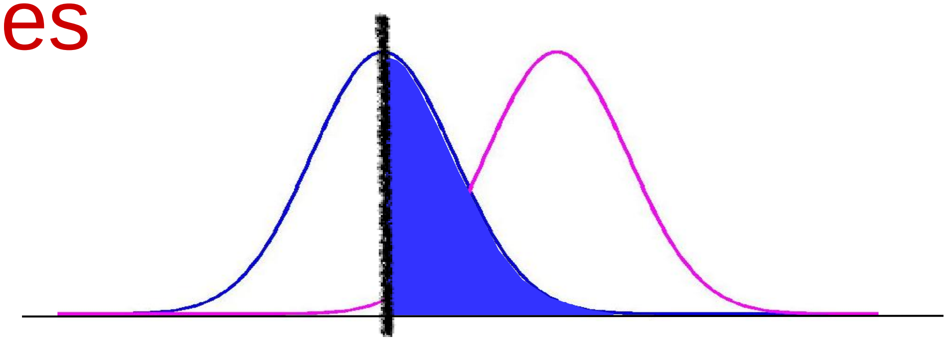
(when $n \gg s$)

probability of error:

$$\mathbb{P}(\mathcal{S}_k \neq \mathcal{S}) = \mathbb{P}\left(\{\mathcal{S}^c \cap \mathcal{S}_k \neq \emptyset\} \cup \{\mathcal{S} \cap \mathcal{S}_k^c \neq \emptyset\}\right)$$

$$\leq \mathbb{P}\left(\mathcal{S}^c \cap \mathcal{S}_k \neq \emptyset\right) + \mathbb{P}\left(\mathcal{S} \cap \mathcal{S}_k^c \neq \emptyset\right)$$

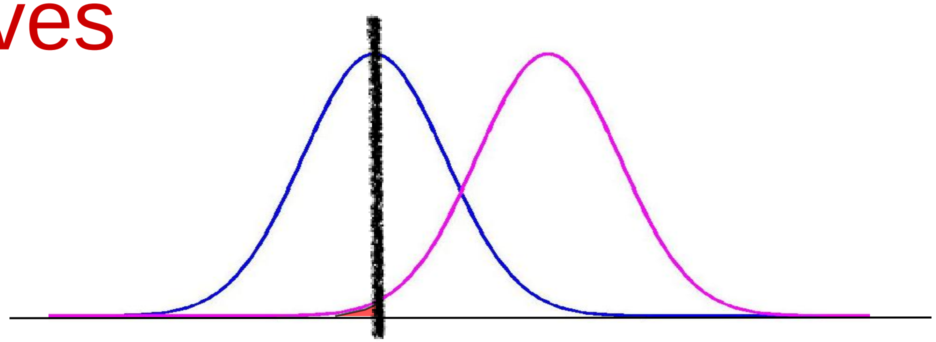
False Positives



$$\mathbb{P}(\mathcal{S}_k \neq \mathcal{S}) \;\leq\; \mathbb{P}(\mathcal{S}^c \cap \mathcal{S}_k \neq \emptyset) + \mathbb{P}(\mathcal{S} \cap \mathcal{S}_k^c \neq \emptyset)$$

$$\begin{aligned}
\mathbb{P}(\mathcal{S}^c \cap \mathcal{S}_k \neq \emptyset) &= \mathbb{P}\left(\bigcup_{i \notin \mathcal{S}} \bigcap_{j=1}^{k} y_{i,j} > 0\right) \\
&\leq \sum_{i \notin \mathcal{S}} \mathbb{P}\left(\bigcap_{j=1}^{k} y_{i,j} > 0\right) \\
&= \sum_{i \notin \mathcal{S}} 2^{-k} = \frac{n-s}{2^k}
\end{aligned}$$

False Negatives



$$\mathbb{P}(\mathcal{S}_k \neq \mathcal{S}) \quad \leq \quad \mathbb{P}\left(\mathcal{S}^c \cap \mathcal{S}_k \neq \emptyset\right) \; + \; \mathbb{P}\left(\mathcal{S} \cap \mathcal{S}_k^c \neq \emptyset\right)$$

$$\begin{aligned} \mathbb{P}\left(\mathcal{S} \cap \mathcal{S}_k^c \neq \emptyset\right) \quad &= \quad \mathbb{P}\left(\bigcup_{j=1}^{k} \bigcup_{i \in \mathcal{S}} y_{i,j} < 0\right) \\ &\leq \quad \frac{ks}{2} \exp\left(-\frac{\mu^2}{4}\right) \end{aligned}$$

Probability of Error Bound

$$\mathbb{P}(\mathcal{S}_k \neq \mathcal{S}) \leq \mathbb{P}(\mathcal{S}^c \cap \mathcal{S}_k \neq \emptyset) + \mathbb{P}(\mathcal{S} \cap \mathcal{S}_k^c \neq \emptyset)$$

$$\leq \frac{n-s}{2^k} + \frac{ks}{2} \exp\left(-\frac{\mu^2}{4}\right)$$

$$= \frac{n-s}{2^k} + \frac{1}{2} \exp\left(-\frac{(\mu^2 - 4\log(ks))}{4}\right)$$

Choose $k = \log_2 n^{1+\epsilon}$ and consider high-dimensional limit...

$$\mathbb{P}(\mathcal{S}_k \neq \mathcal{S}) \leq \cancelto{0}{\frac{n-s}{2^k}} + \frac{1}{2} \exp\left(-\frac{(\mu^2 - 4\log(s(1+\epsilon)\log_2 n))}{4}\right)$$

Probability of error goes to zero if

$$\mu \gtrsim \sqrt{4\log(s(1+\epsilon)\log_2 n)}$$

(Malloy & Nowak, 2011)

Improvements Through Sequential Design

non-sequential: $\mu \gtrsim \sqrt{2\log(n-s)} + \sqrt{2\log s}$ (necessary)

sequential thresholding: (sufficient)

$$\begin{aligned} \mu &\gtrsim \sqrt{4\log(s(1+\epsilon)\log_2 n)} \\ &\cong 2\sqrt{\log s + \log\log_2 n} \end{aligned}$$

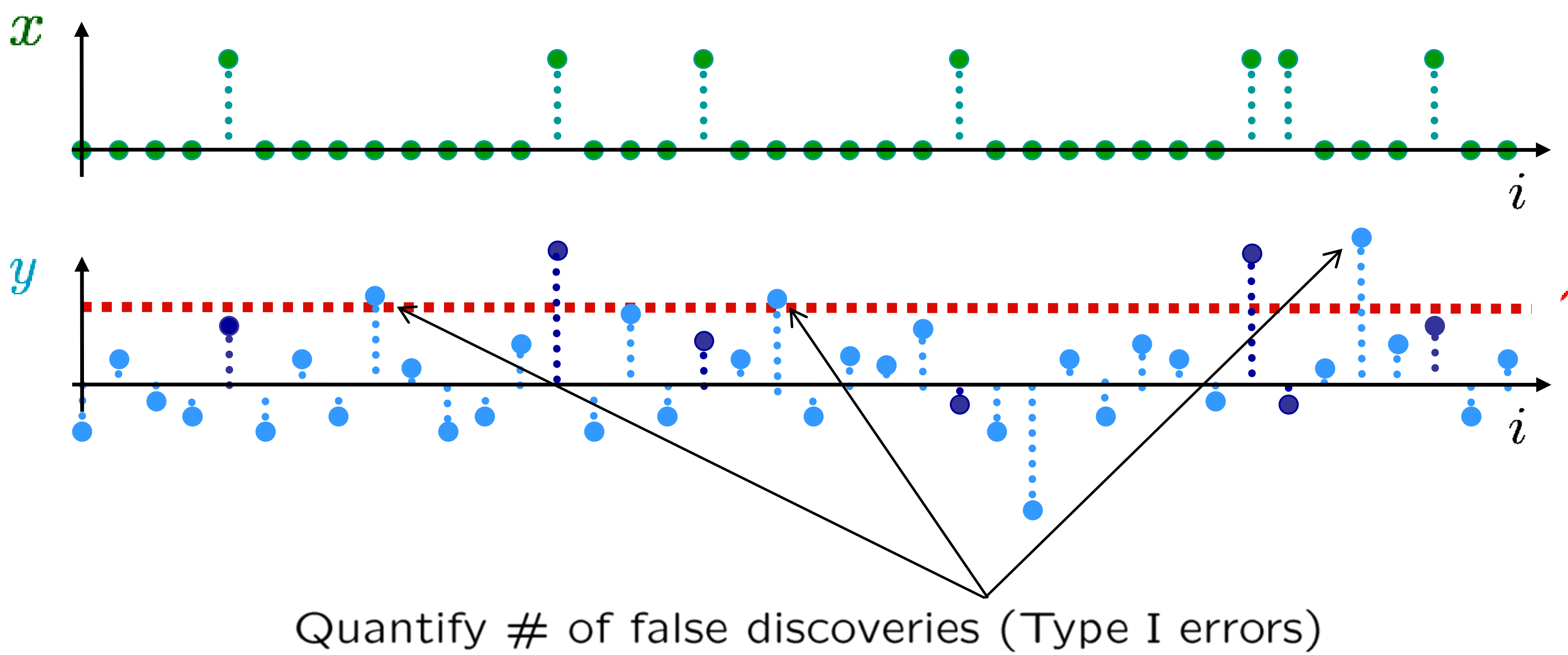
significant gains when $s \ll n$

greater sensitivity for same precision budget or lower experimental requirements for equivalent sensitivity

Active Sensing for Sparse Recovery

-- A Relaxed Error Criteria --

Measuring Error: False Discoveries

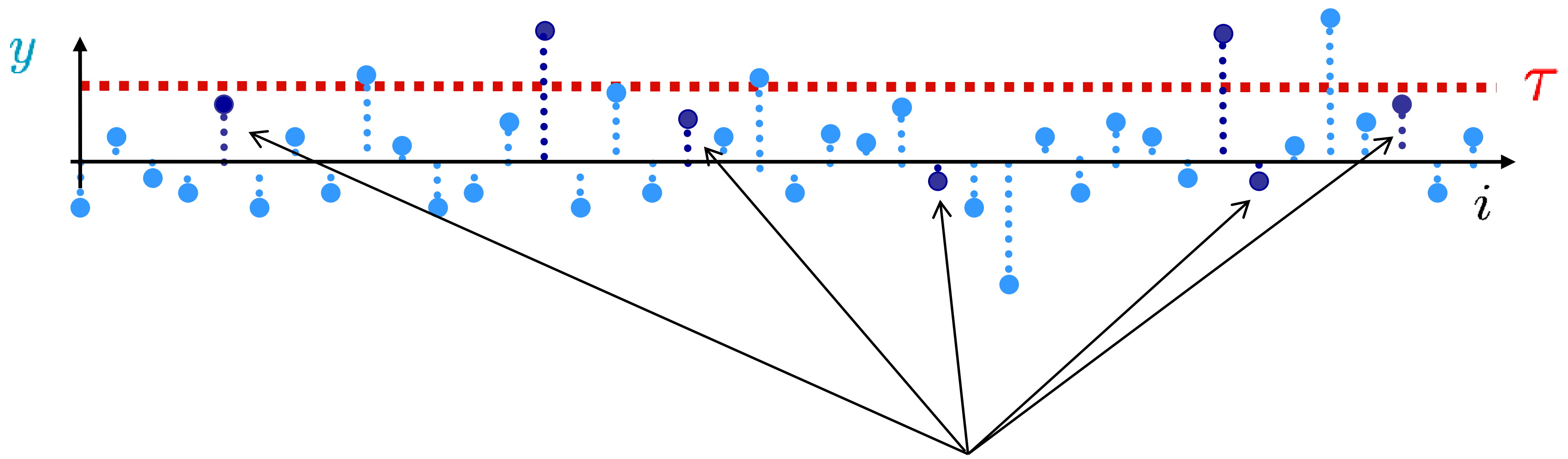


Definition: The *False Discovery Proportion* (FDP) of an estimator $\widehat{\mathcal{S}}$ is

$$\mathrm{FDP}(\widehat{\mathcal{S}}) := \frac{|\widehat{\mathcal{S}} \setminus \mathcal{S}|}{|\widehat{\mathcal{S}}|} = \frac{\text{\# falsely discovered components}}{\text{total \# discovered components}}$$

Here, FDP = **3/5**

Measuring Error: Non-Discoveries



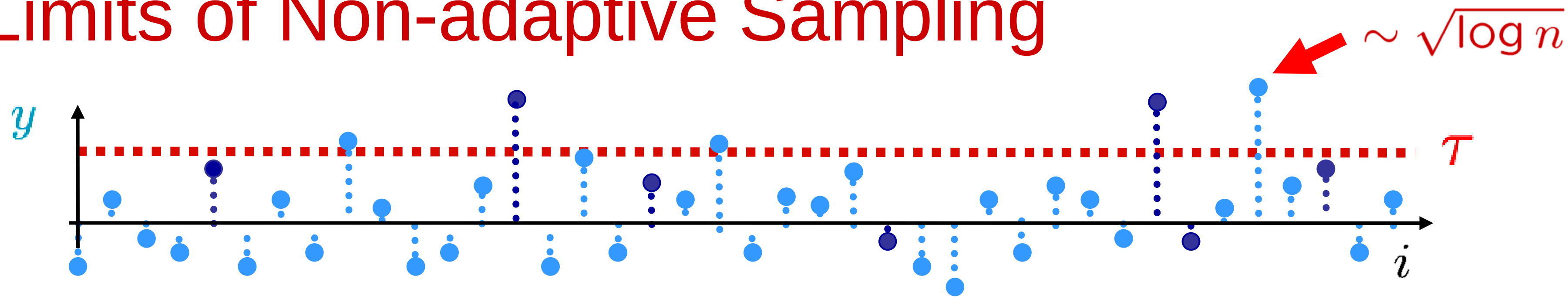
Also interested in quantifying false negatives (Type II errors)

Definition: The *Non-Discovery Proportion* (NDP) of $\widehat{\mathcal{S}}$ is

$$\mathrm{NDP}(\widehat{\mathcal{S}}) := \frac{|\mathcal{S} \setminus \widehat{\mathcal{S}}|}{|\mathcal{S}|} = \frac{\#\ \text{signal components missed}}{\text{total}\ \#\ \text{signal components}}$$

Here, NDP = **5/7**

Limits of Non-adaptive Sampling



To determine performance in high-dimensional settings (large n), we consider *asymptotic* behavior of FDP and NDP

Assume sublinear sparsity: $|\mathcal{S}| = n^{1-\beta}$ for some fixed $0 < \beta < 1$

e.g., $\beta = 3/4 \Rightarrow$ $n = 10000 \rightarrow |\mathcal{S}| = 10$
$n = 1000000 \rightarrow |\mathcal{S}| = 32$

Theorem: (Donoho & Jin, 2003)
Assume x has $n^{1-\beta}$, $\beta \in (0,1)$, nonzero components of amplitude $\sqrt{2r \log n}$, $r > 0$. If $r > \beta$, there exists a threshold test that yields an estimator $\widehat{\mathcal{S}} = \widehat{\mathcal{S}}(y)$ for which

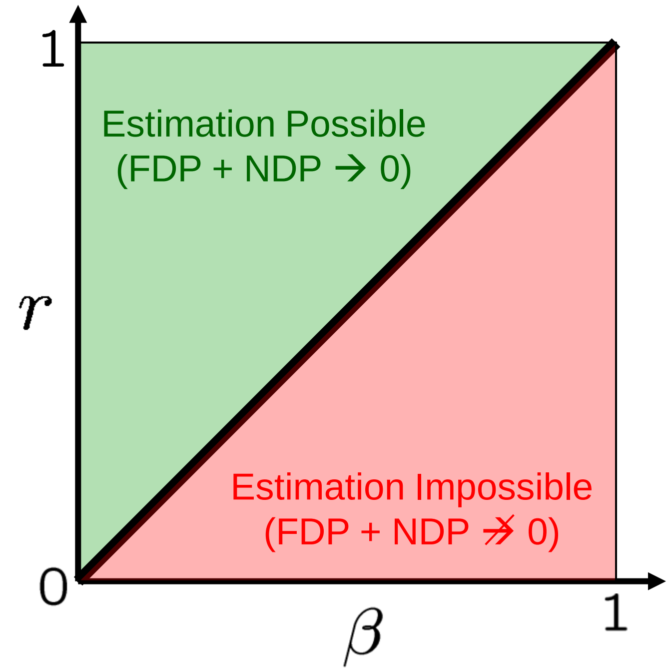
$$\text{FDP}(\widehat{\mathcal{S}}) \xrightarrow{P} 0, \ \text{NDP}(\widehat{\mathcal{S}}) \xrightarrow{P} 0, \ \text{as } n \rightarrow \infty$$

where $\xrightarrow{P}$ denotes convergence in probability. Further, if $r < \beta$, there does not exist a threshold test that can guarantee that both NDP and FDP tend to zero as $n \rightarrow \infty$.

Sharp Delineation in “Parameter Space”

Signal Strength

$\mu = \sqrt{2r \log n}$



Sparsity

$\#$ signal components $= n^{1-\beta}$

What if no signal component amplitudes exceed $\sqrt{2\beta \log n}$?

Distilled Sensing (DS)

Input: Number of observation steps: k
Precision per step: R_j s.t. $\sum_{j=1}^{k} R_j \leq R(n)$

Initialize: Index set $I_1 = \{1, 2, \ldots, n\}$

Loop: For each step $j = 1$ to k

1) Allocate precision uniformly over I_j: $\gamma_{i,j} = R_j / |I_j|, \; i \in I_j$
2) Collect observations $y_{i,j}$ for $i \in I_j$
3) Refinement/distillation: $I_{j+1} = \{i \in I_j : y_{i,j} > 0\}$

Output: Final observations: $y_{\text{DS}} := y_{i,k}, \; i \in I_k$

*To recover negative components, replace $y_{i,j}$ by $-y_{i,j}$ in distillation step

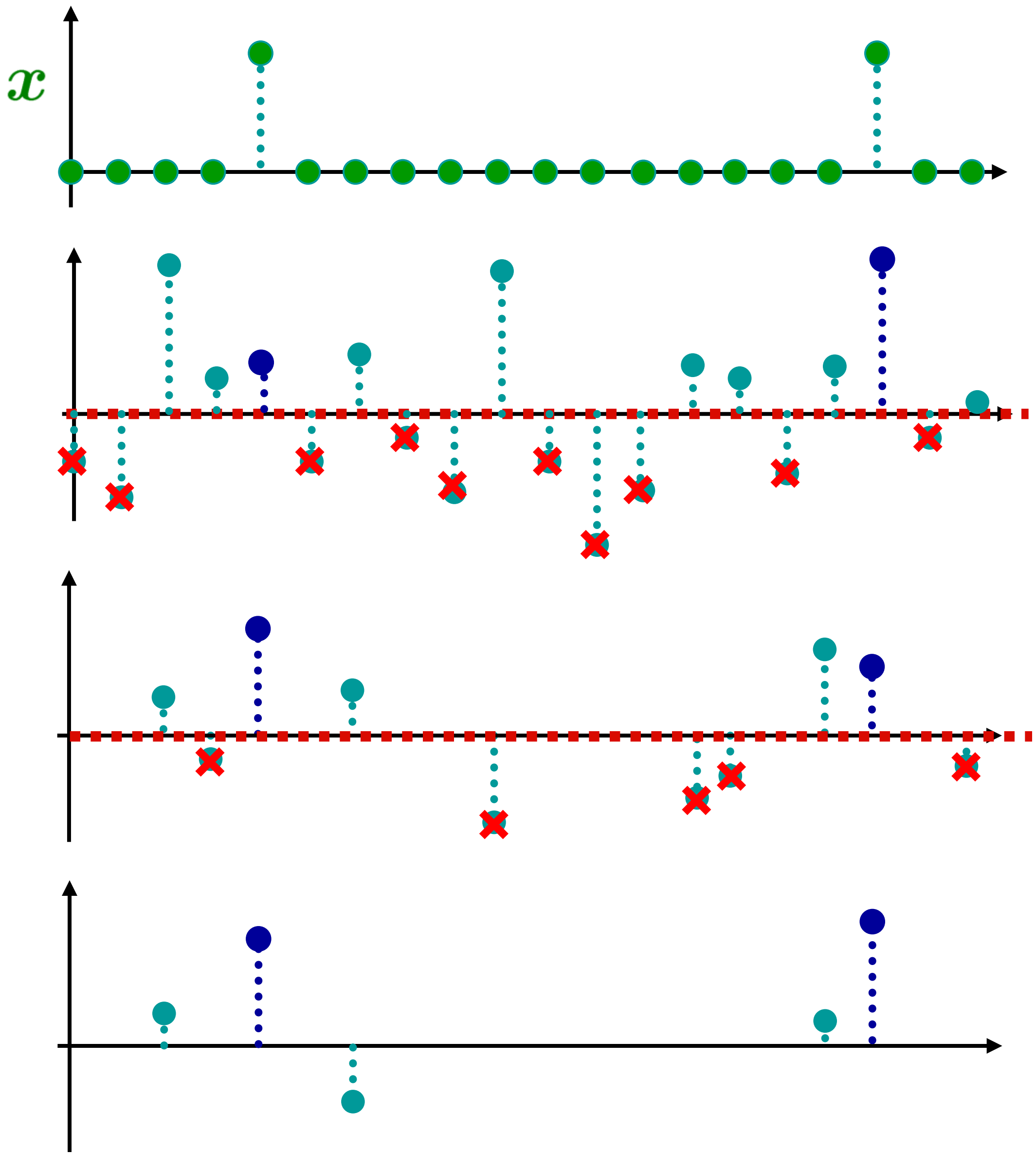
Key Idea: $|I_{j+1}| \approx |I_j|/2$ when x is sparse

– Assume $R_{j+1}/R_j = \rho > 1/2$, then

$$\gamma_{i,j+1} = \frac{R_{j+1}}{|I_{j+1}|} \approx 2\rho \frac{R_j}{|I_j|} = 2\rho \; \gamma_{i,j} \quad (\text{for } i \in I_j \cap I_{j+1})$$

SNR improvement by $2\rho > 1$

Idealized Example



Let $k = 3$

$R_j = R(n)/3, \ \ j = 1, 2, 3$

$$y_{i,1} = x_i + \mathcal{N}(0, 3)$$

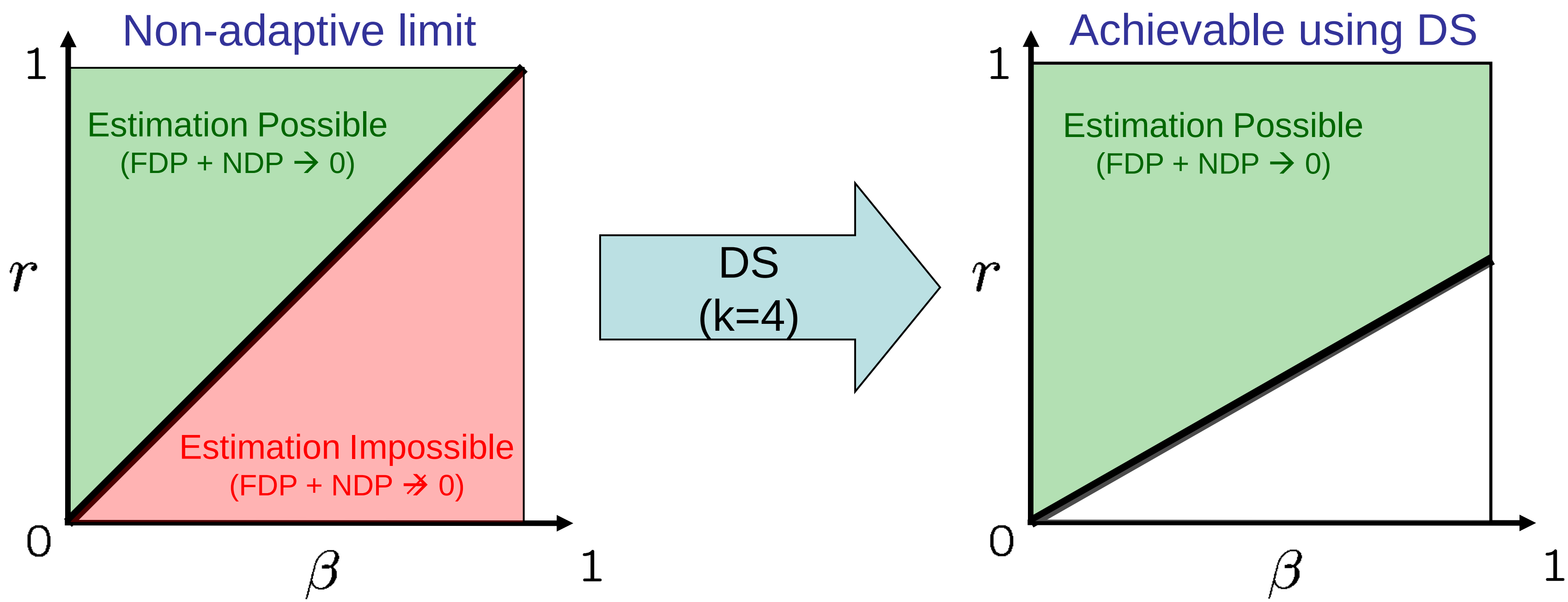
$$y_{i,2} = x_i + \mathcal{N}\left(0, \frac{3}{2}\right)$$

$$y_{i,3} = x_i + \mathcal{N}\left(0, \frac{3}{4}\right)$$

Equal Allocation of Sensing Resources

Theorem: (JH, R. Castro, and R. Nowak, 2008)
Assume x has $n^{1-\beta}$, $\beta \in (0,1)$, nonzero components, and let $R_j = n/k$ (equal precision allocation) for a fixed $k \in \mathbb{N}$. If the signal component amplitudes exceed $\sqrt{2\beta \frac{k}{2^{k-1}} \log n}$, then there exists a threshold test that yields an estimator $\widehat{\mathcal{S}} = \widehat{\mathcal{S}}(y_{\mathrm{DS}})$ for which

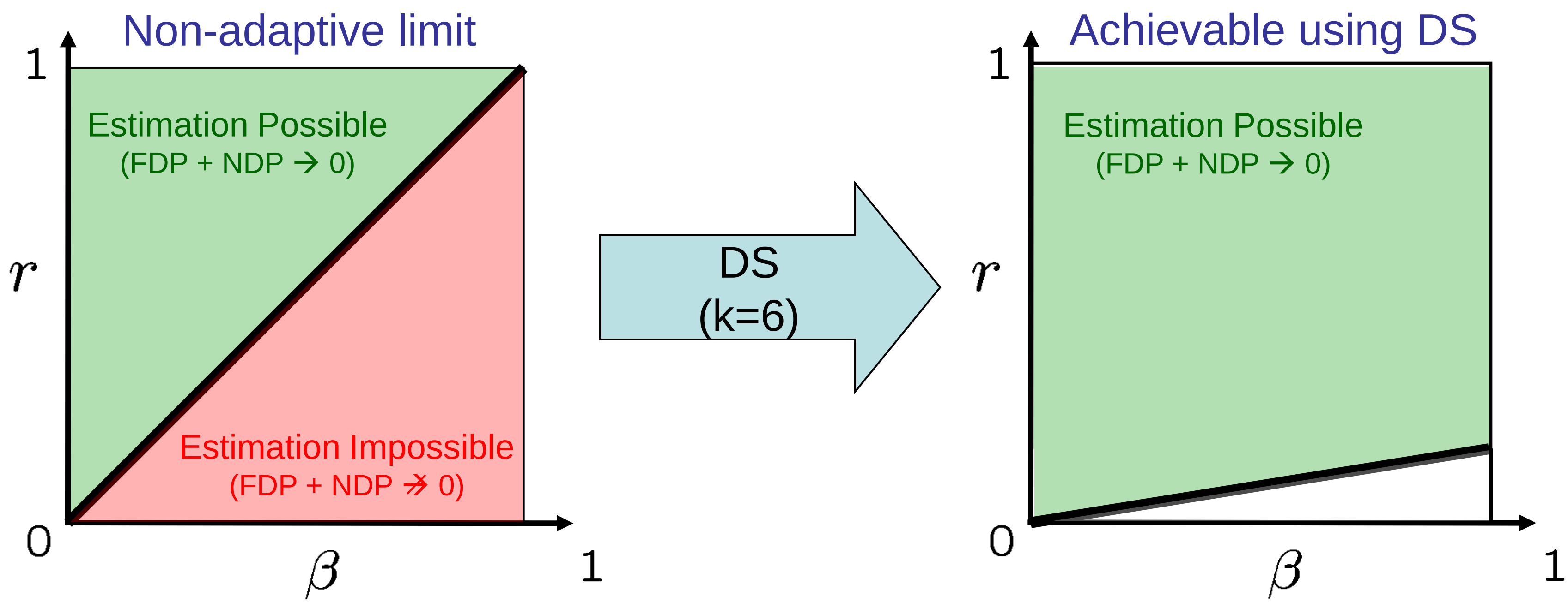
$$\mathrm{FDP}(\widehat{\mathcal{S}}) \xrightarrow{P} 0, \;\; \mathrm{NDP}(\widehat{\mathcal{S}}) \xrightarrow{P} 0, \;\; \text{as } n \to \infty.$$



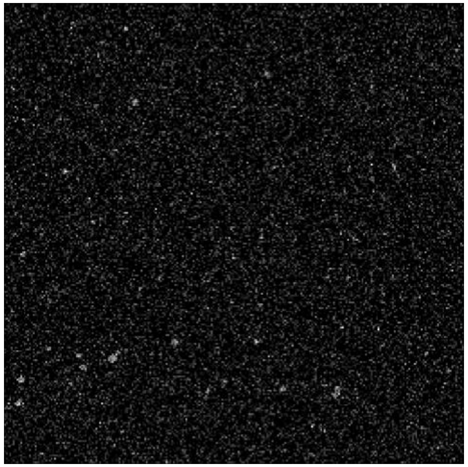
Equal Allocation of Sensing Resources

Theorem: (JH, R. Castro, and R. Nowak, 2008)
Assume x has $n^{1-\beta}$, $\beta \in (0,1)$, nonzero components, and let $R_j = n/k$ (equal precision allocation) for a fixed $k \in \mathbb{N}$. If the signal component amplitudes exceed $\sqrt{2\beta \frac{k}{2^{k-1}} \log n}$, then there exists a threshold test that yields an estimator $\widehat{\mathcal{S}} = \widehat{\mathcal{S}}(y_{\mathrm{DS}})$ for which

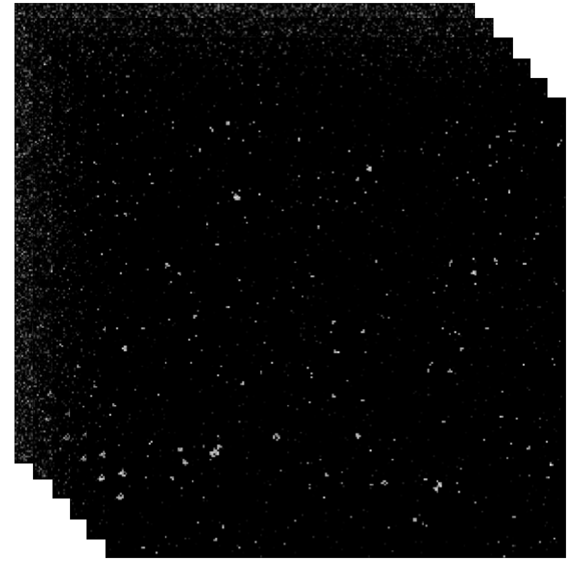
$$\mathrm{FDP}(\widehat{\mathcal{S}}) \xrightarrow{P} 0, \ \mathrm{NDP}(\widehat{\mathcal{S}}) \xrightarrow{P} 0, \ \text{as } n \to \infty.$$



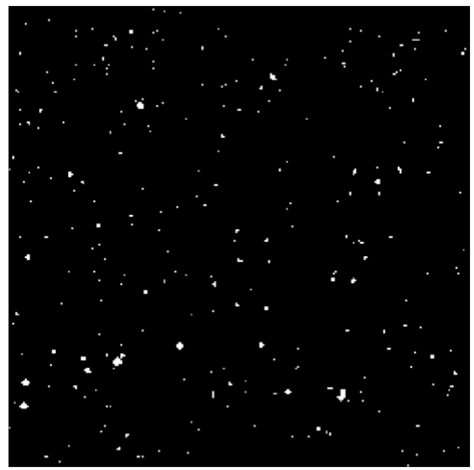
Recall: Noisy Astronomical Imaging



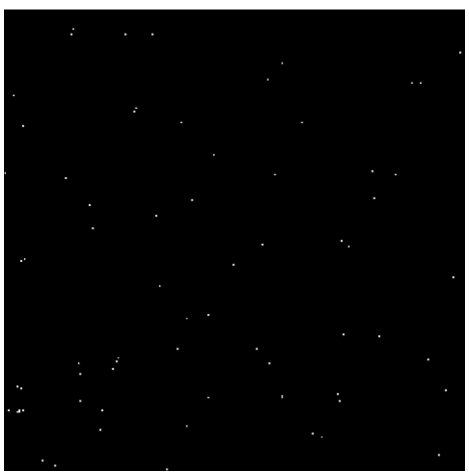
Non-adaptive observations



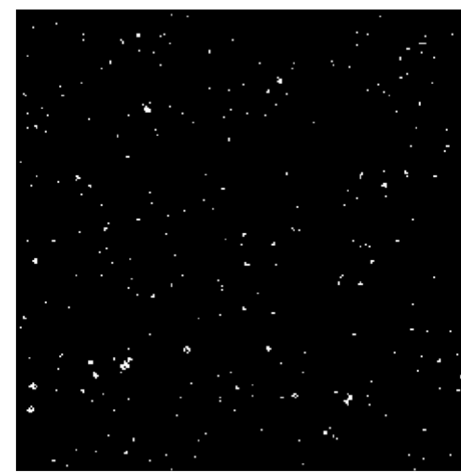
DS observations (k=6)



original signal
(~0.8% non-zero components)

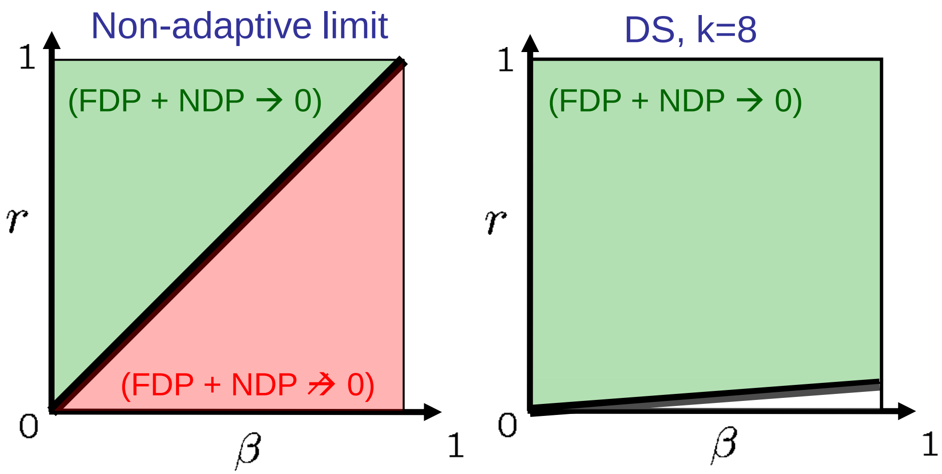


Non-adaptive recovery
(FDP = 0.05)



Adaptive recovery
(FDP = 0.05)

Geometric Allocation of Sensing Resources



What about *very weak* signals?
(amplitude $\ll \sqrt{\log n}$)

Theorem: (JH, R. Castro, and R. Nowak, 2009)
Assume x has $n^{1-\beta}$, $\beta \in (0,1)$, nonzero components having amplitude at least $\mu(n)$. Choose $k = \lceil \log_2 \log n \rceil + 2$, and precision budget allocated over observation steps such that $\sum_{j=1}^{k} R_j \leq n$, $R_{j+1}/R_j = \rho > 1/2$ for $j = 1, \ldots, k-2$, $R_1 = c_1 n$, and $R_k = c_k n$ for $c_1, c_k \in (0,1)$. From the output of the DS procedure, construct the estimate

$$\widehat{\mathcal{S}}_{\mathrm{DS}} := \left\{ i \in I_k : y_{i,k} > \sqrt{2/c_k} \right\}$$

If $\mu(n)$ is *any* arbitrarily slowly growing function of n, then

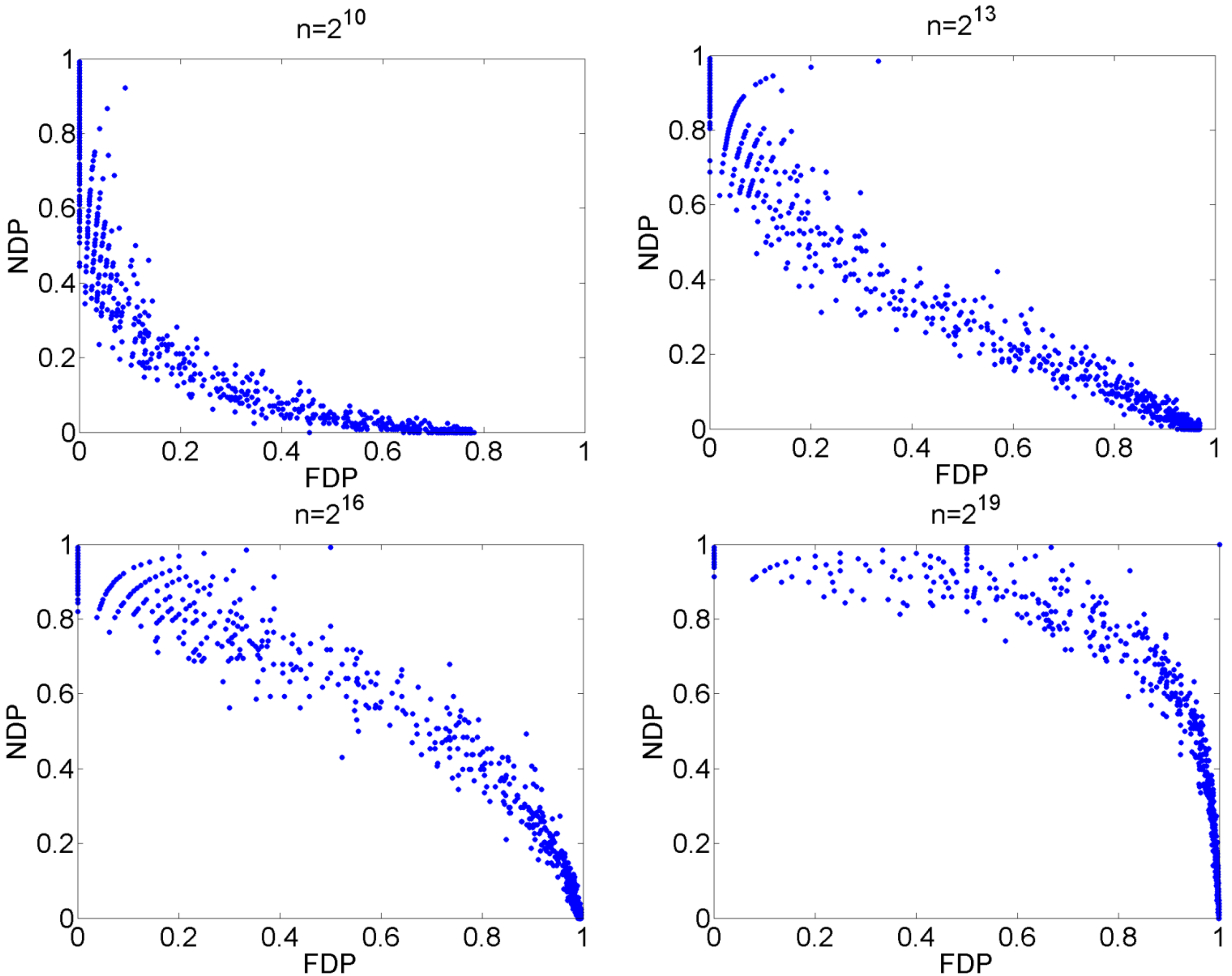
$$\mathrm{FDP}(\widehat{\mathcal{S}}_{\mathrm{DS}}) \xrightarrow{P} 0, \ \mathrm{NDP}(\widehat{\mathcal{S}}_{\mathrm{DS}}) \xrightarrow{P} 0, \ \text{as } n \to \infty.$$

Adaptivity can provide $\sim \log n$ improvement in SNR
and mitigate (or nearly *eliminate*) the curse of dimensionality!

The Curse of Dimensionality…

Non-adaptive vs. DS

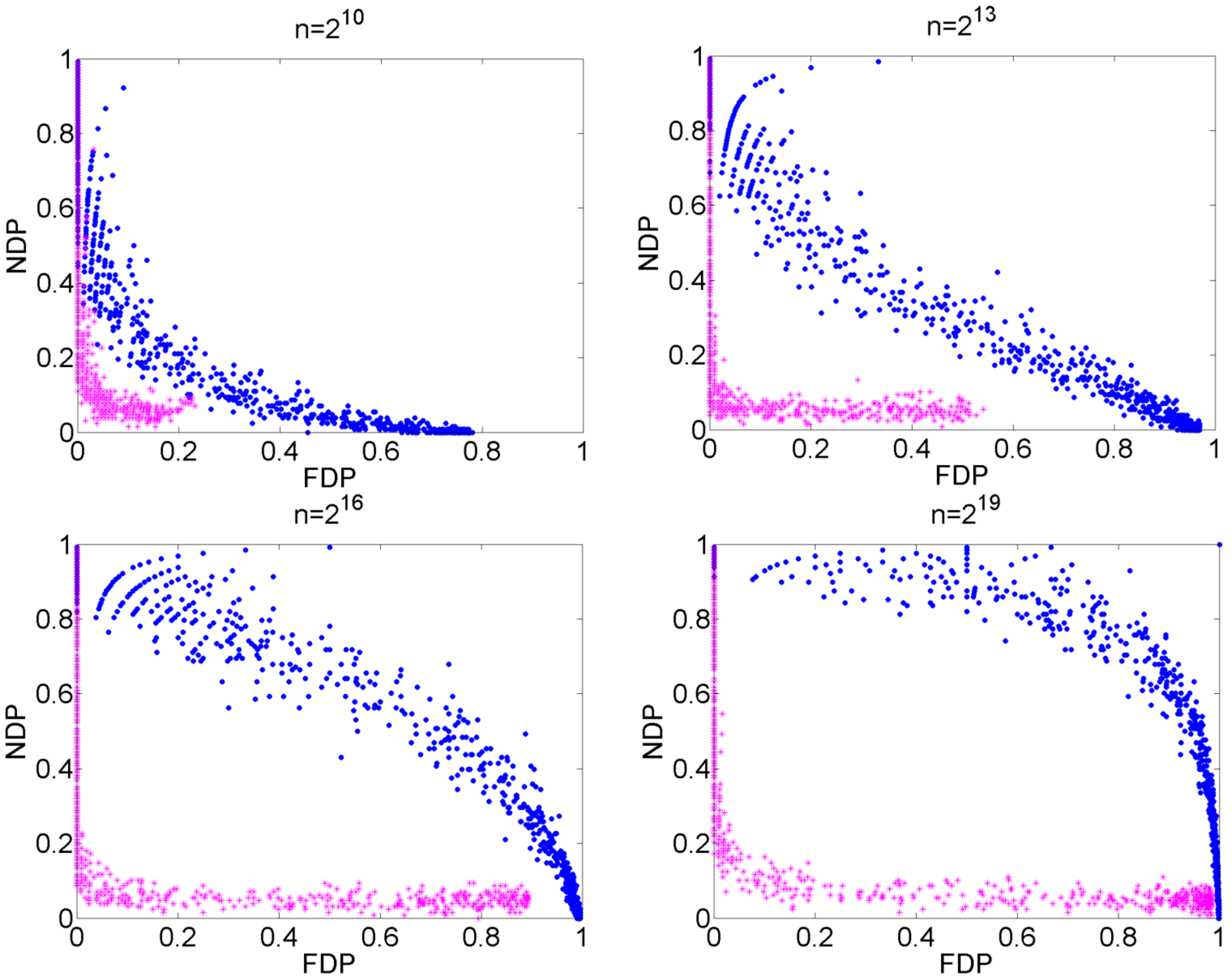
SNR$= \mu^2 = 8$ $|\mathcal{S}| = 128$



…and the Blessing of Adaptivity!

Non-adaptive vs. DS

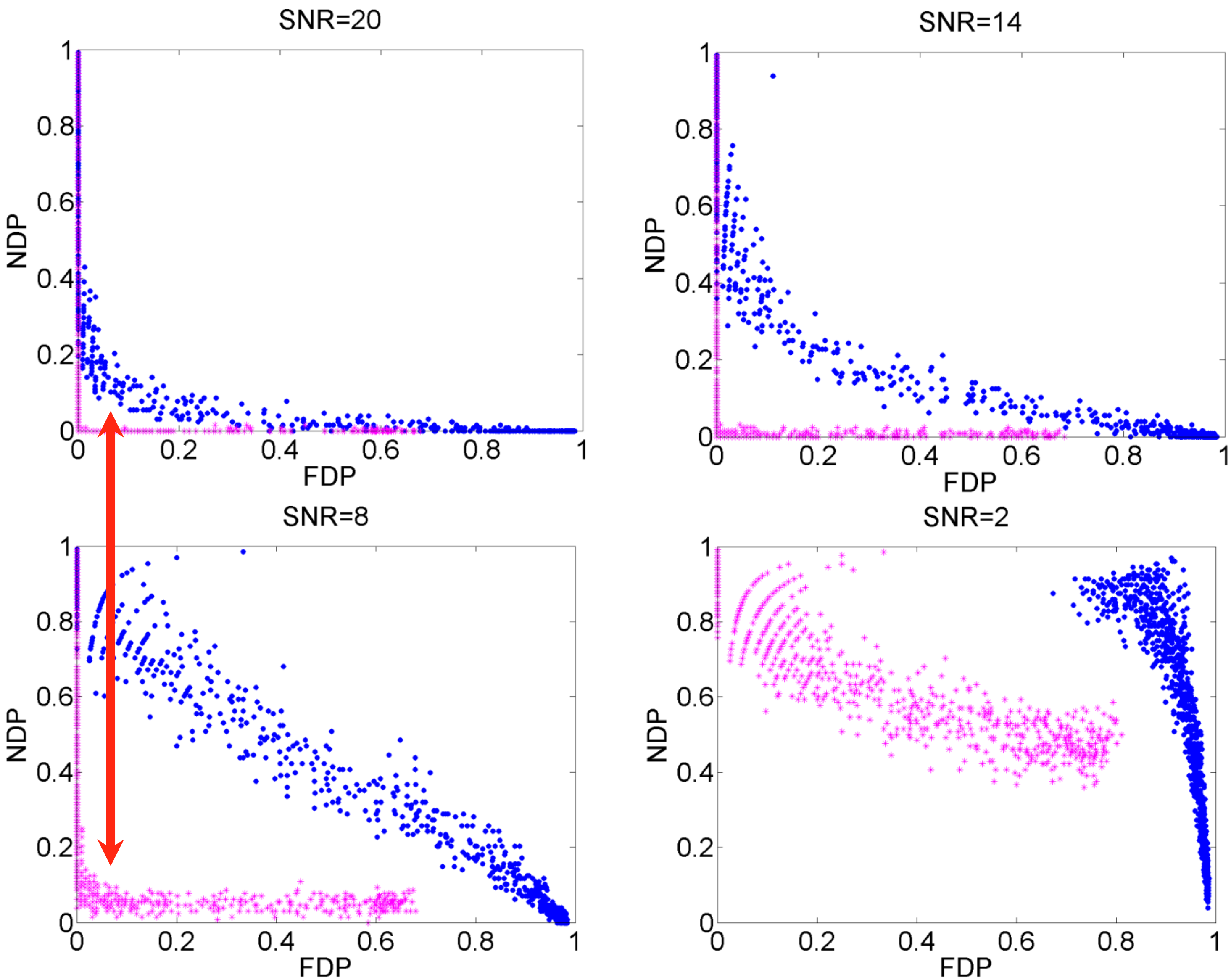
SNR$= \mu^2 = 8 \quad |\mathcal{S}| = 128$



Performance Comparison: Varying SNR

Non-adaptive vs. DS

$$n = 2^{14} = 16384 \qquad |\mathcal{S}| = \sqrt{n} = 128$$



Active Sensing for Sparse Recovery

-- Adaptive Compressive Sampling --

Improvements w.r.t. Other Resources?

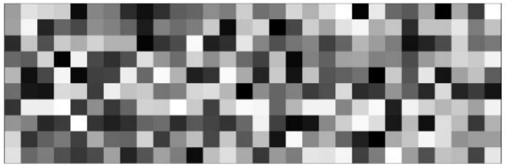
Note that DS requires about $2n$ total measurements:

- n for first step
- about $n/2$ for second step
- about $n/4$ for third step...

Can we achieve noise-resilience benefits of DS using a reduced # of samples?

Noisy Compressive Sensing (CS) Observation Model

$$\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} =$$



$$\mathbb{E}\left[\|A\|_F^2\right] = n$$

$$\times \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} z_1 \\ \vdots \\ z_m \end{bmatrix}$$

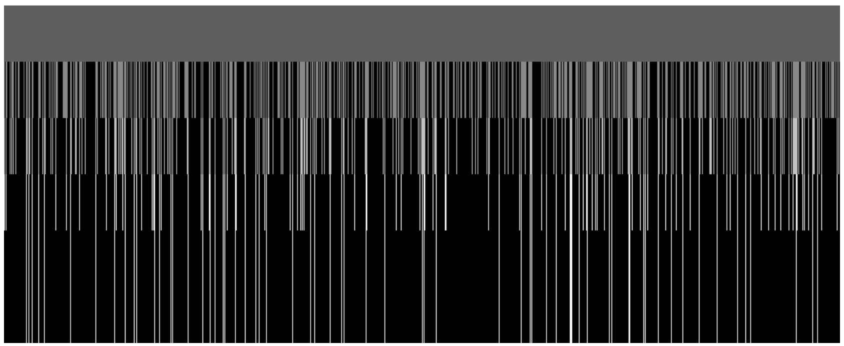
$$z \sim \mathcal{N}(0, I_{m \times m})$$

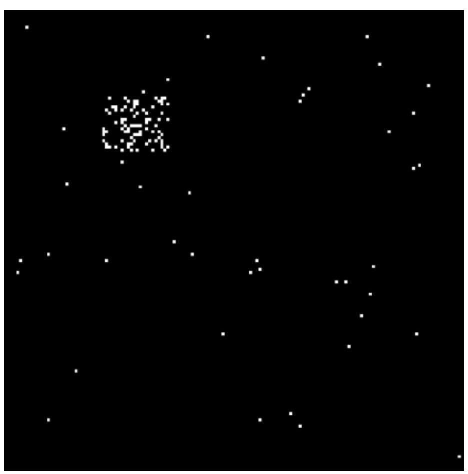
Compressive DS

Theorem: (JH, R. Baraniuk, R. Castro, and R. Nowak, 2009)
Assume x has $|\mathcal{S}| = n^{1-\beta}$ nonzero components of amplitude μ. Collect $O(|\mathcal{S}| \log n)$ adaptive random compressive measurements. When $\mu \succeq \sqrt{\log \log \log n}$, there exists a (tractable) recovery procedure for which

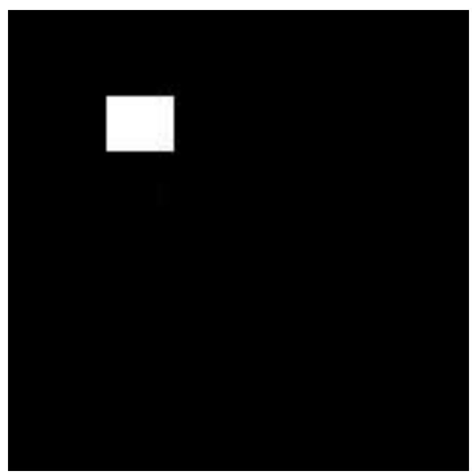
$$\mathrm{FDP}(\widehat{\mathcal{S}}_{\mathrm{CDS}}) \xrightarrow{P} 0, \ \mathrm{NDP}(\widehat{\mathcal{S}}_{\mathrm{CDS}}) \xrightarrow{P} 0, \ \text{as } n \to \infty.$$

Support of random measurement matrix

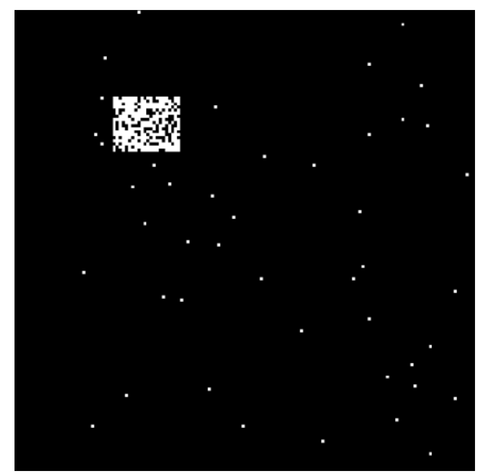




Non-adaptive CS Recovery
(∼ 25% measurements)



Original Signal

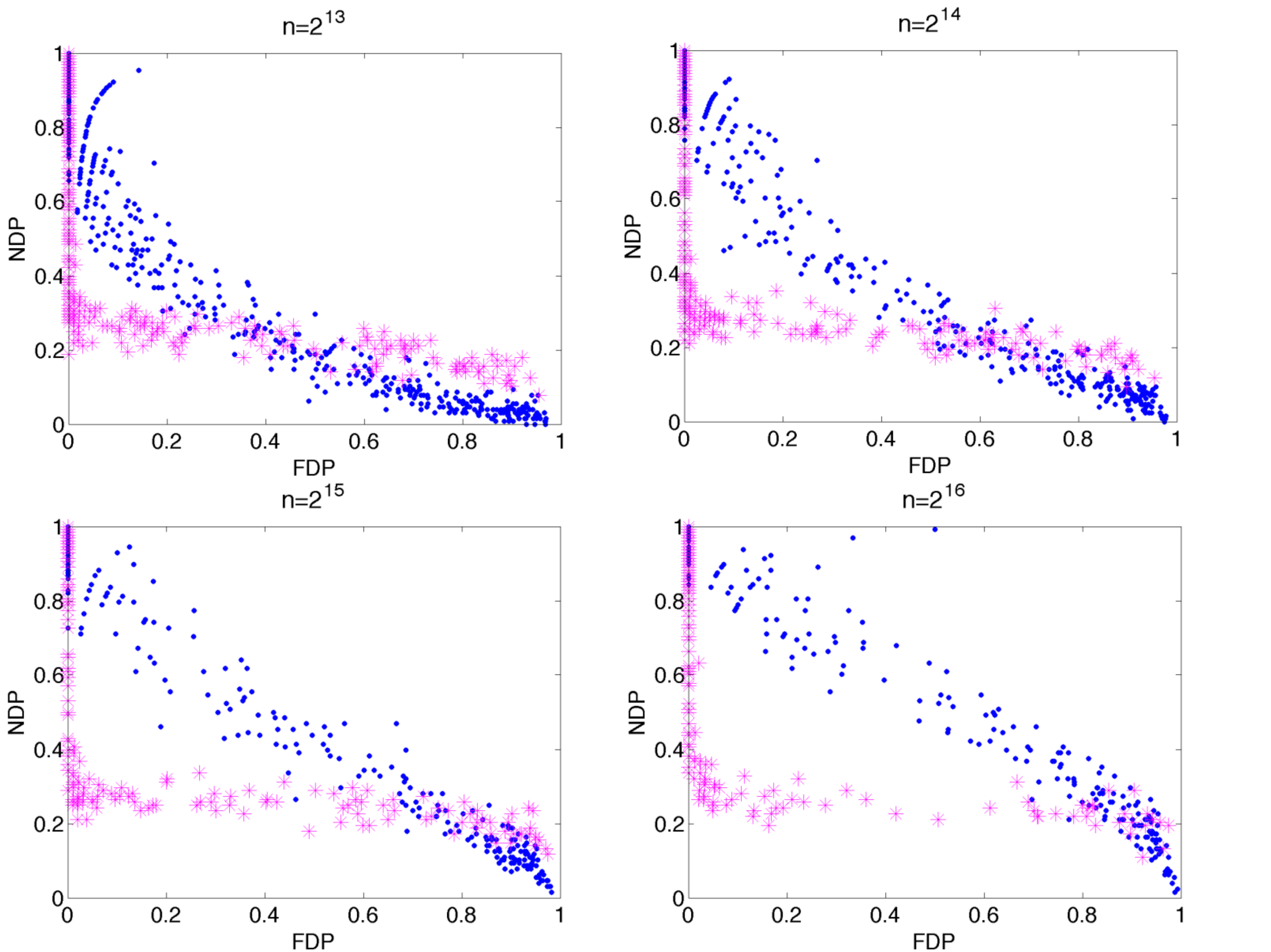


Adaptive CS Recovery
(∼ 25% measurements)

FDP & NDP for Compressive Distilled Sensing

Non-adaptive CS vs. CDS

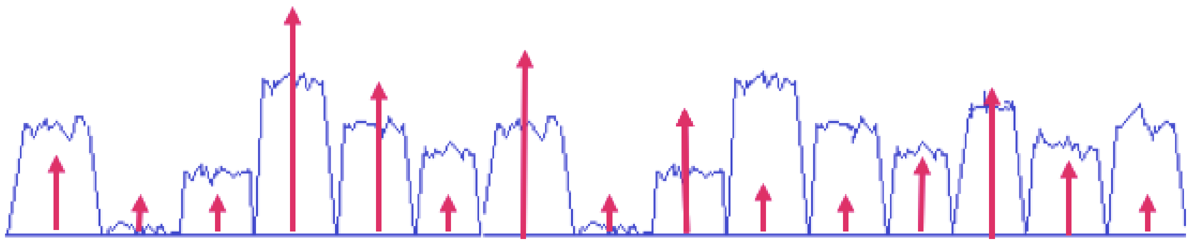
SNR= 12 $\quad |\mathcal{S}| = 128$



Generalization: Beyond Gaussian Models

-- Sequential Thresholding --

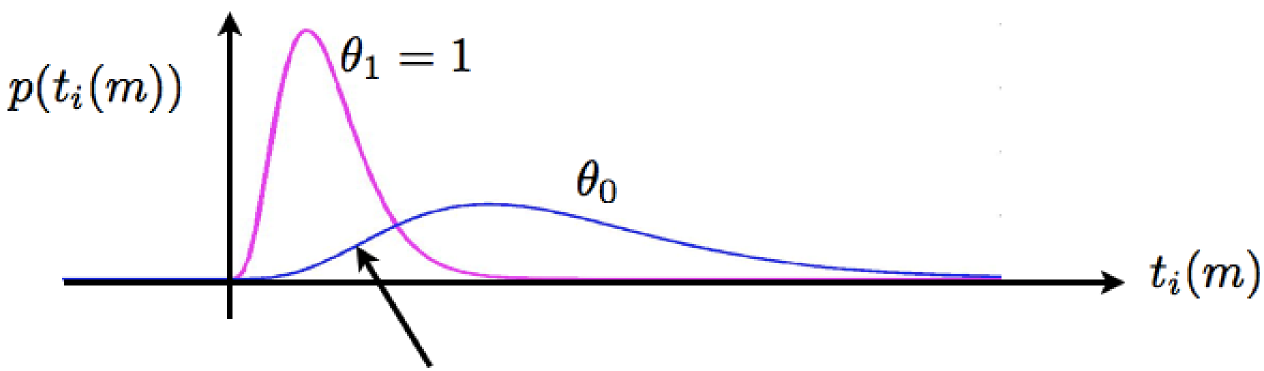
Spectrum Sensing



goal: find open channel(s) as quickly as possible

channel samples: $y_{i,j} \overset{\text{iid}}{\sim} \mathcal{CN}(0,\theta)\ , \quad \theta_0 > \theta_1 = 1$

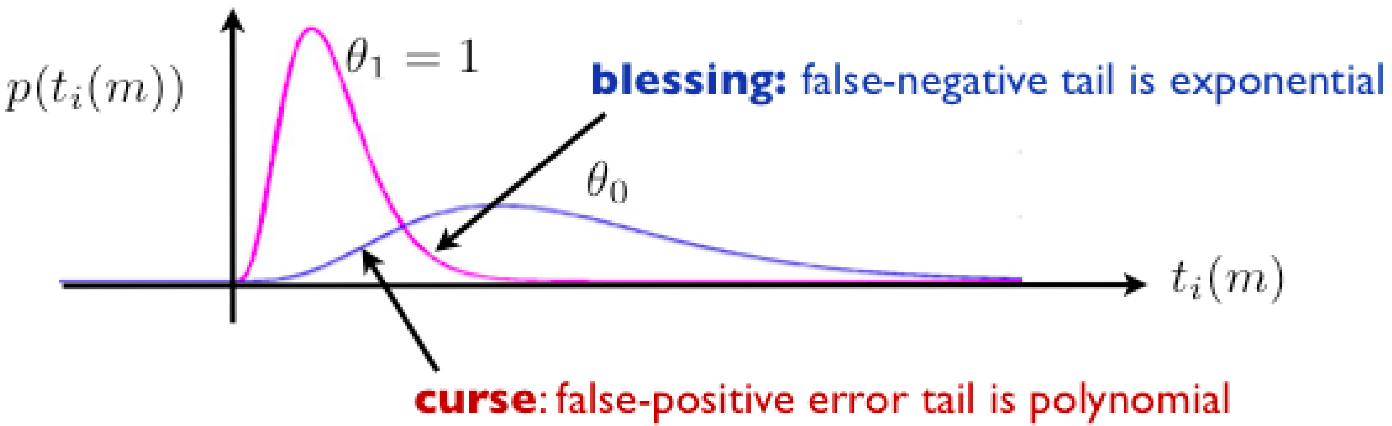
test statistic: $t_i(m) \ = \ \sum_{j=1}^{m} |y_{i,j}|^2 \ \sim \begin{cases} \Gamma(m,\theta_0)\,, & i \notin \mathcal{S} \\ \Gamma(m,1)\,, & i \in \mathcal{S} \end{cases}$



curse: false-positive error tail is polynomial

Spectrum Sensing Application

test statistic: $t_i(m) = \sum_{j=1}^{m} |y_{i,j}|^2 \sim \begin{cases} \Gamma(m, \theta_0), & i \notin \mathcal{S} \\ \Gamma(m, 1), & i \in \mathcal{S} \end{cases}$



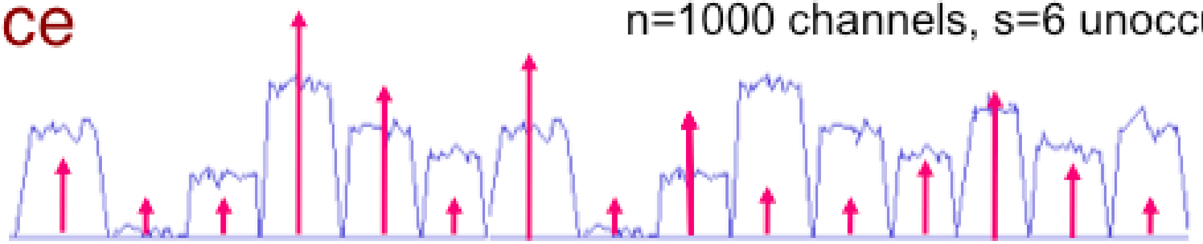
non-sequential: $\theta_0 \geq 2(m-1)(n-s)^{1/2m} \sim n^{1/2m}$ (necessary)

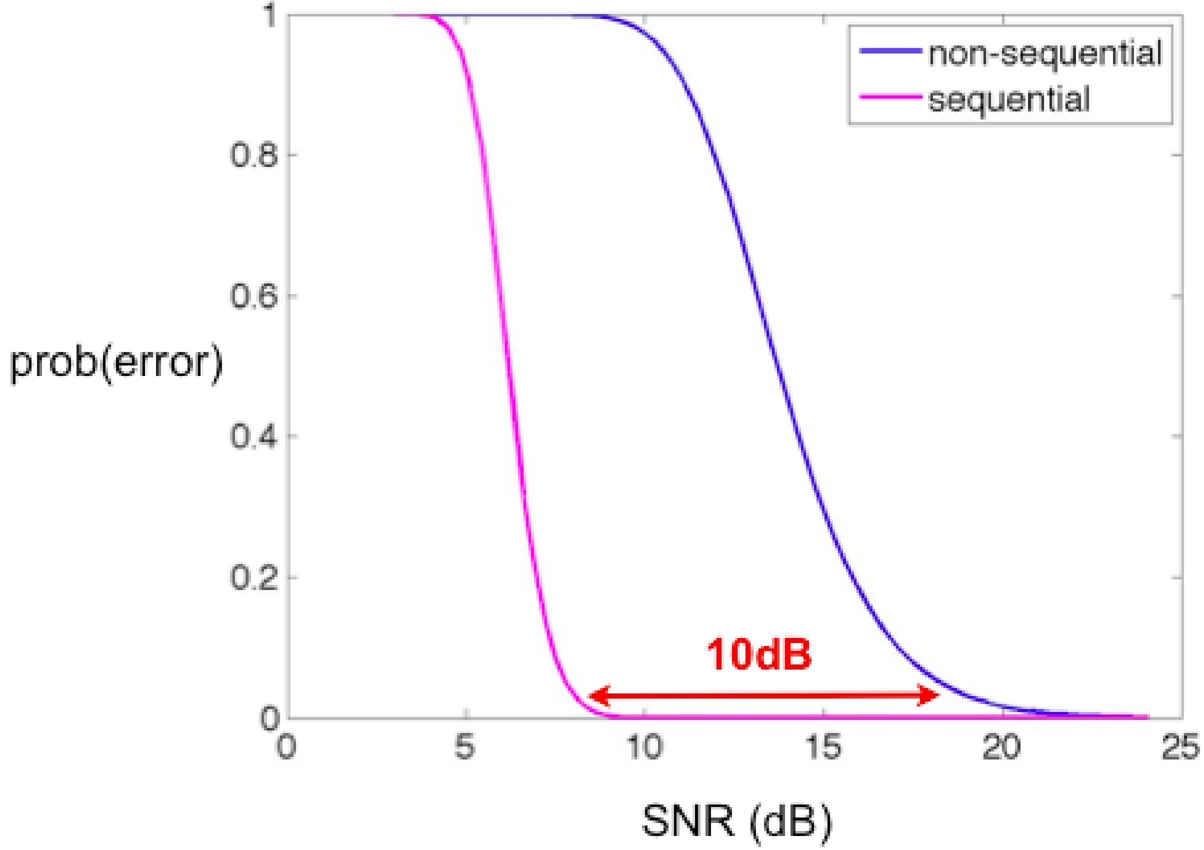
SPRT: $\theta_0 \gtrsim \frac{1}{m} \log s$ minimum requirement for any testing scheme with expected sample budget nm

sequential thresholding: $\theta_0 \geq \frac{1}{2m} \log(s \log_2 n)$ (sufficient)

Performance

n=1000 channels, s=6 unoccupied





sequential thresholding is about 10 times more sensitive
(for equal scan time) or scans 3-4 times faster (for same reliability)