# $\ell_1$ minimization without amplitude information

Petros Boufounos
*petrosb@rice.edu*

RICE

# The Big $\ell_1$ Picture

Classical $\ell_1$ reconstruction problems:

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \text{ s.t. } f(\mathbf{x}) = 0$$

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \text{ s.t. } f(\mathbf{x}) \geq 0$$

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 + \lambda f(\mathbf{x})$$

# The Big $\ell_1$ Picture

Classical $\ell_1$ reconstruction problems:

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \text{ s.t. } f(\mathbf{x}) = 0$$

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \text{ s.t. } f(\mathbf{x}) \geq 0$$

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 + \lambda f(\mathbf{x})$$

Sparsity Model

Data Fidelity:

$$f(\mathbf{x}) = \|\mathbf{\Phi x} - \mathbf{y}\|_2$$

$$f(\mathbf{x}) = \epsilon - \|\mathbf{\Phi x} - \mathbf{y}\|_2$$

# Data without amplitude information

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 + \lambda f(\mathbf{x})$$

Q: What if data provide no amplitude information on $\mathbf{x}$?

# Data without amplitude information

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 + \lambda f(\mathbf{x})$$

Q: What if data provide no amplitude information on $\mathbf{x}$?

$$f(\alpha \mathbf{x}) = \alpha f(\mathbf{x})$$

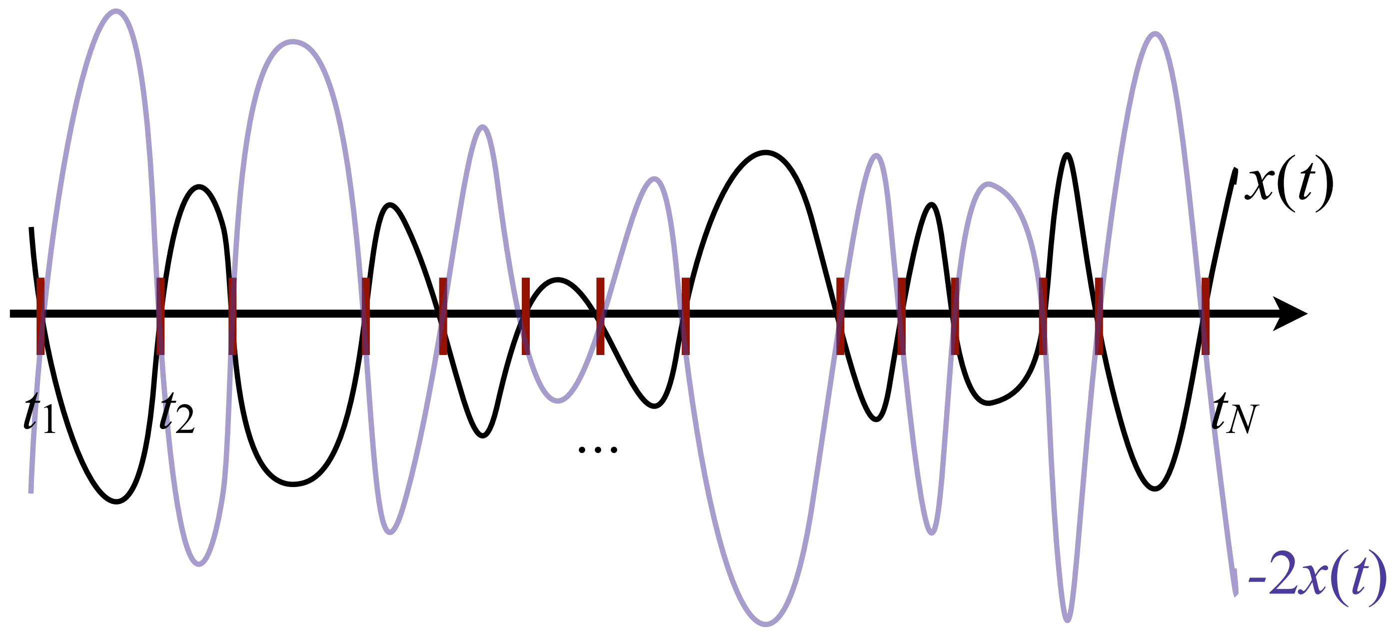
Minimizing solution is degenerate:

$\mathbf{x}=0$

Impose an amplitude constraint.

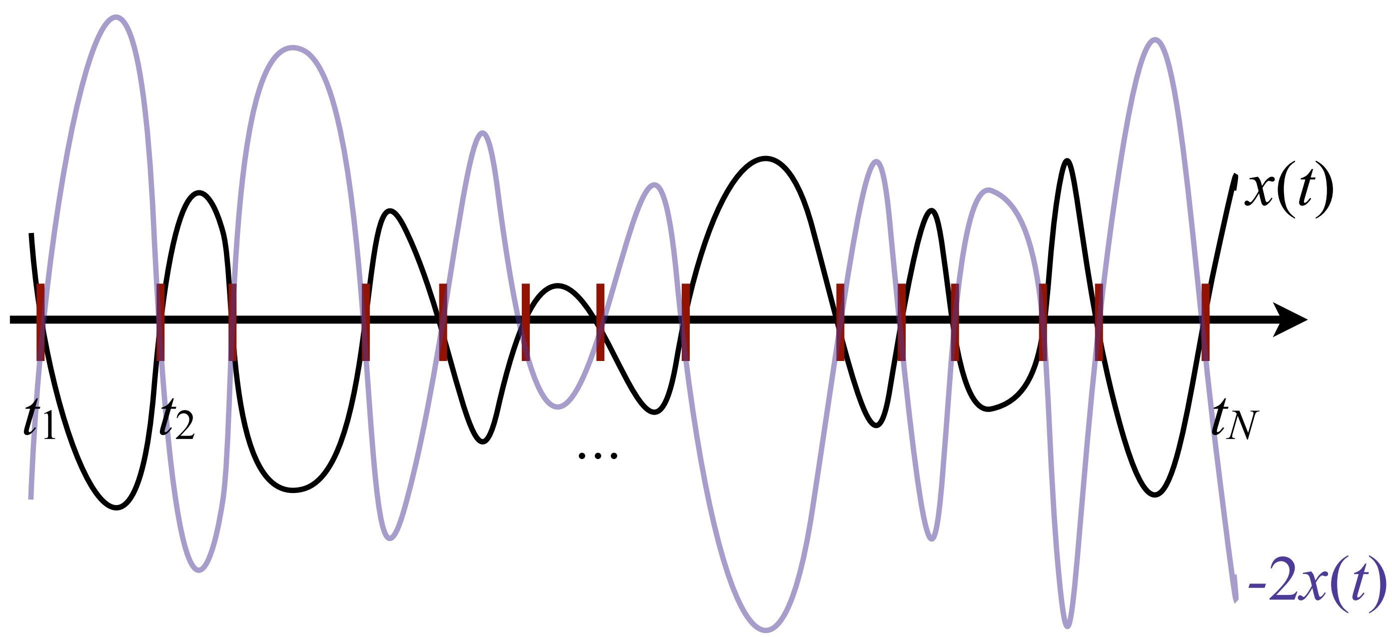
# Case I:
# Sampling the Zero Crossings

# Signal Reconstruction from Zero Crossings

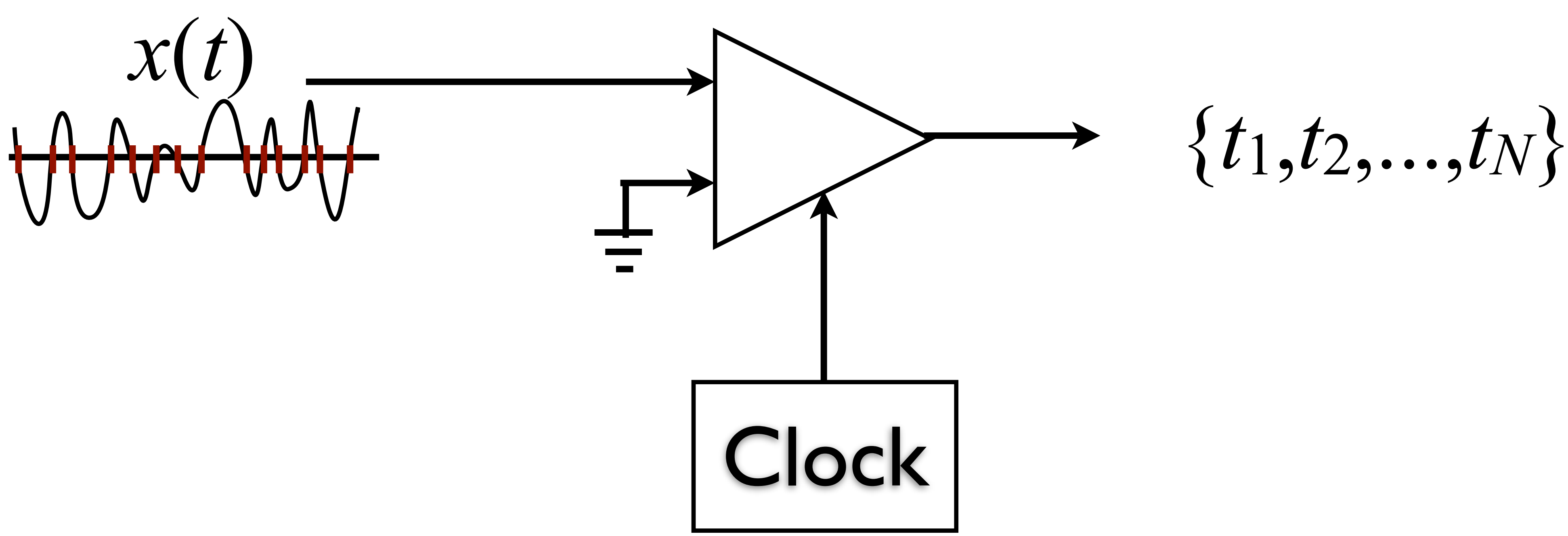



Q: Given only the zero crossings $\{t_1, t_2, \ldots, t_N\}$ of a signal can we reconstruct it?

# Signal Reconstruction from Zero Crossings

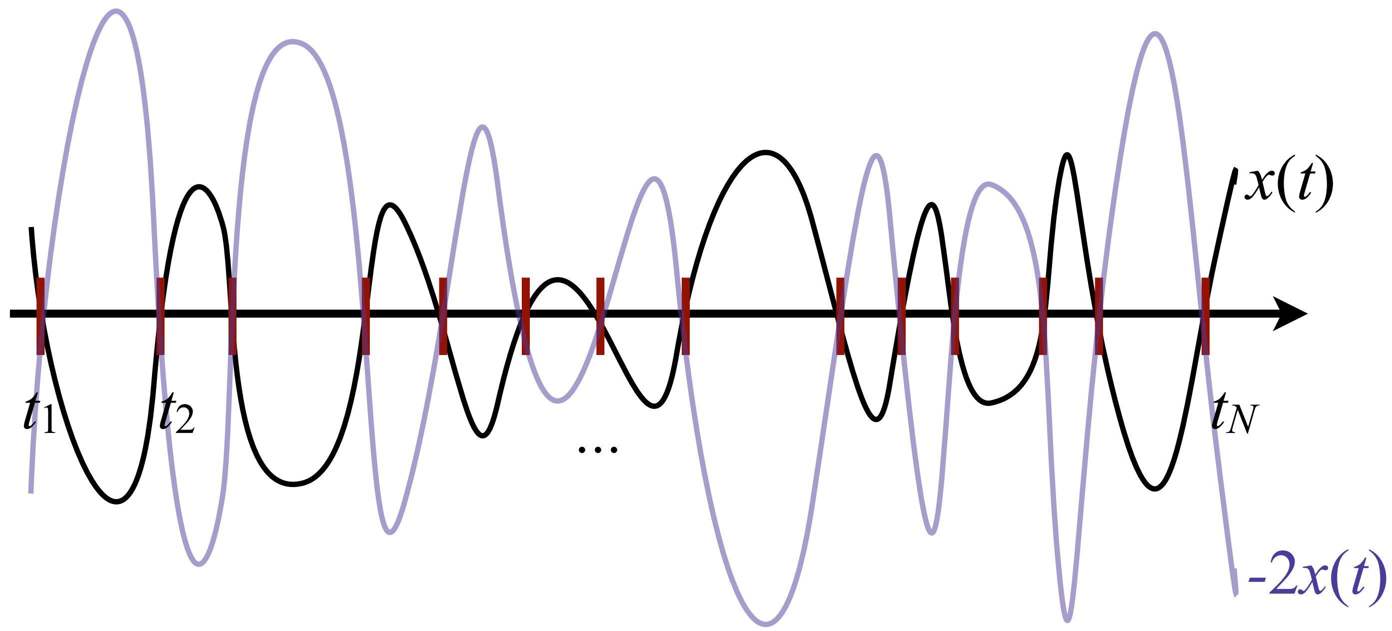



Q: Given only the zero crossings $\{t_1, t_2, \ldots, t_N\}$ of a signal can we reconstruct it?




Easy implementation: only need a comparator and a clock

# Logan's Theorem

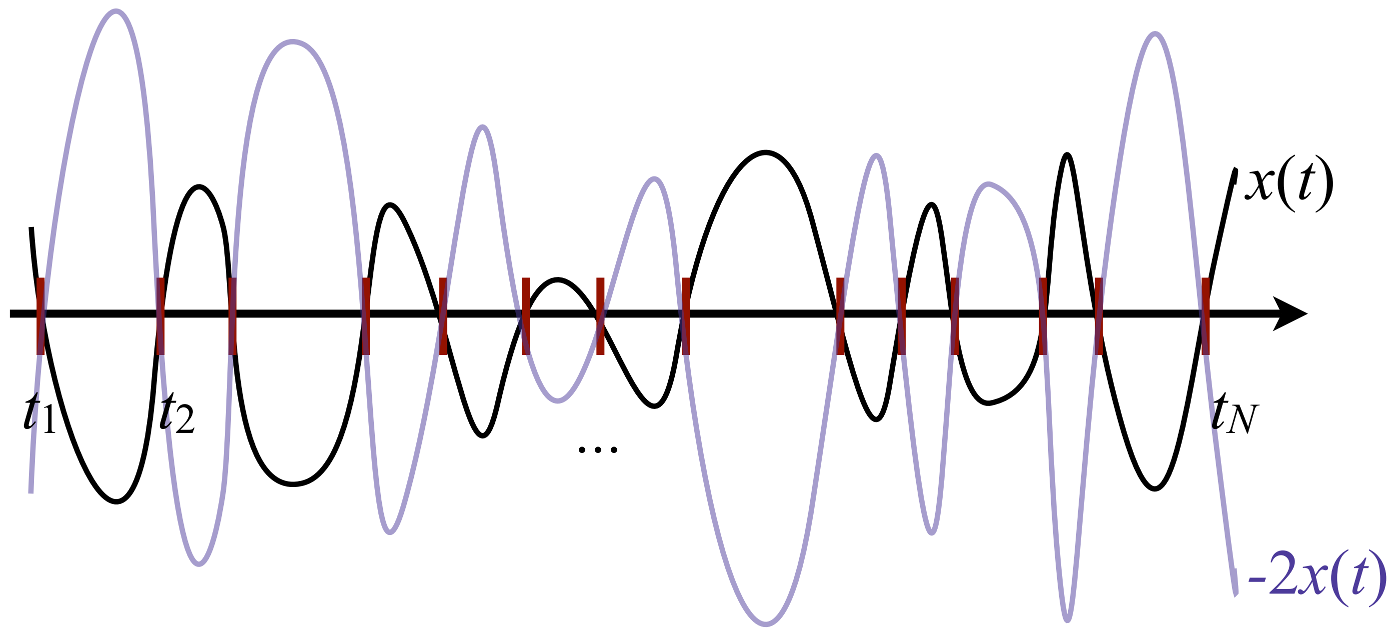



Q: Given only the zero crossings $\{t_1, t_2, \ldots, t_N\}$ of a signal can we reconstruct it?

Logan's Theorem: YES. Signals bandlimited to $[B, 2B)$ are uniquely determined by their zero crossings.

BUT: an arbitrary set of zero crossings might not correspond to a signal bandlimited to $[B, 2B)$. Reconstruction is not robust. There is ambiguity.

# Logan's Theorem




Q: Given only the zero crossings $\{t_1, t_2, \ldots, t_N\}$ of a signal can we reconstruct it?

Logan's Theorem: YES. Signals bandlimited to $[B, 2B)$ are uniquely determined by their zero crossings.

BUT: an arbitrary set of zero crossings might not correspond to a signal bandlimited to $[B, 2B)$.
Reconstruction is not robust. There is ambiguity.

Introduce sparsity to resolve the ambiguity!

# Signal Representation

## Fourier series of $x(t)$:

$$x(t) = \sum_{n \in \mathcal{B}} \left[ a_n \cos(2\pi n t) + b_n \sin(2\pi n t) \right]$$

## Vector of coefficients:

$$\mathbf{x} = \begin{bmatrix} a_{n_1} \\ \vdots \\ a_{n_{N/2}} \\ b_{n_1} \\ \vdots \\ b_{n_{N/2}} \end{bmatrix}$$

# Sampling Operator

Given $\{t_1, t_2, \ldots, t_N\}$,

$$\mathbf{\Phi}_{\{t_k\}} = \begin{bmatrix} \cos(2\pi n_1 t_1) & \ldots & \cos\left(2\pi n_{N/2} t_1\right) & \sin(2\pi n_1 t_1) & \ldots & \sin\left(2\pi n_{N/2} t_1\right) \\ \cos(2\pi n_1 t_2) & \ldots & \cos\left(2\pi n_{N/2} t_2\right) & \sin(2\pi n_1 t_2) & \ldots & \sin\left(2\pi n_{N/2} t_2\right) \\ \vdots & & \vdots & \vdots & & \vdots \\ \cos(2\pi n_1 t_N) & \ldots & \cos\left(2\pi n_{N/2} t_N\right) & \sin(2\pi n_1 t_N) & \ldots & \sin\left(2\pi n_{N/2} t_N\right) \end{bmatrix}$$

Samples the signal at those times:

$$\mathbf{\Phi}_{\{t_k\}} \mathbf{x} = \begin{bmatrix} x(t_1) \\ \vdots \\ x(t_N) \end{bmatrix}$$

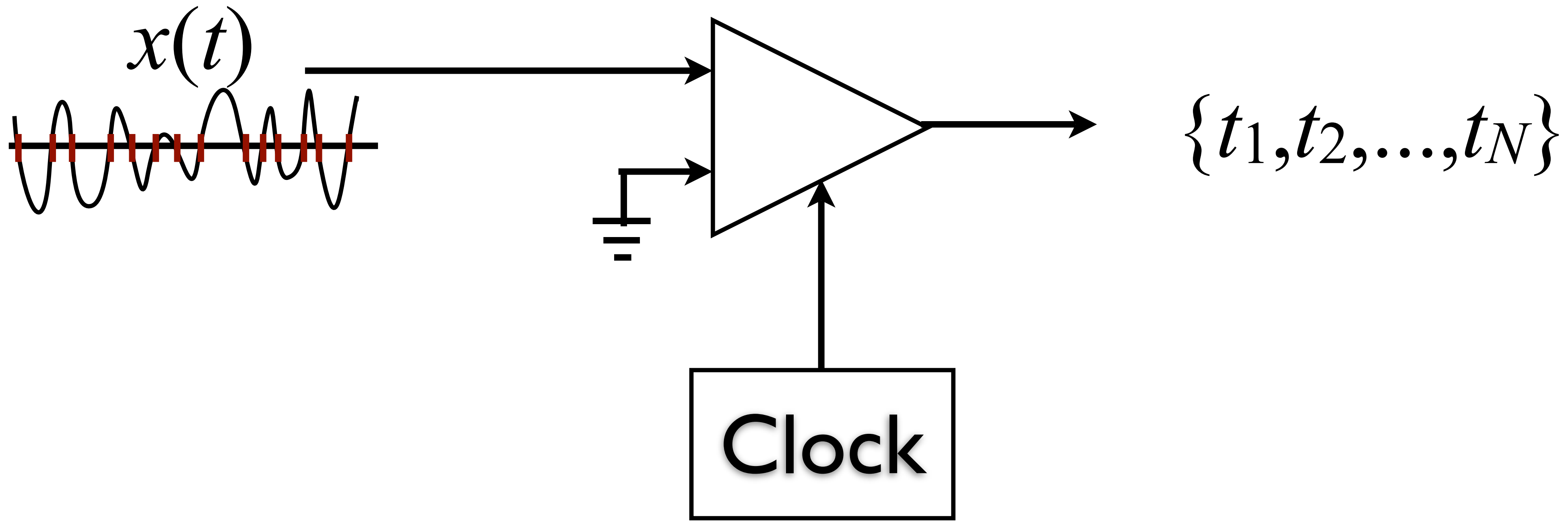
# Reconstruction Problem

$$\mathbf{\Phi}_{\{t_k\}}\mathbf{x} = \begin{bmatrix} x(t_1) \\ \vdots \\ x(t_N) \end{bmatrix}$$

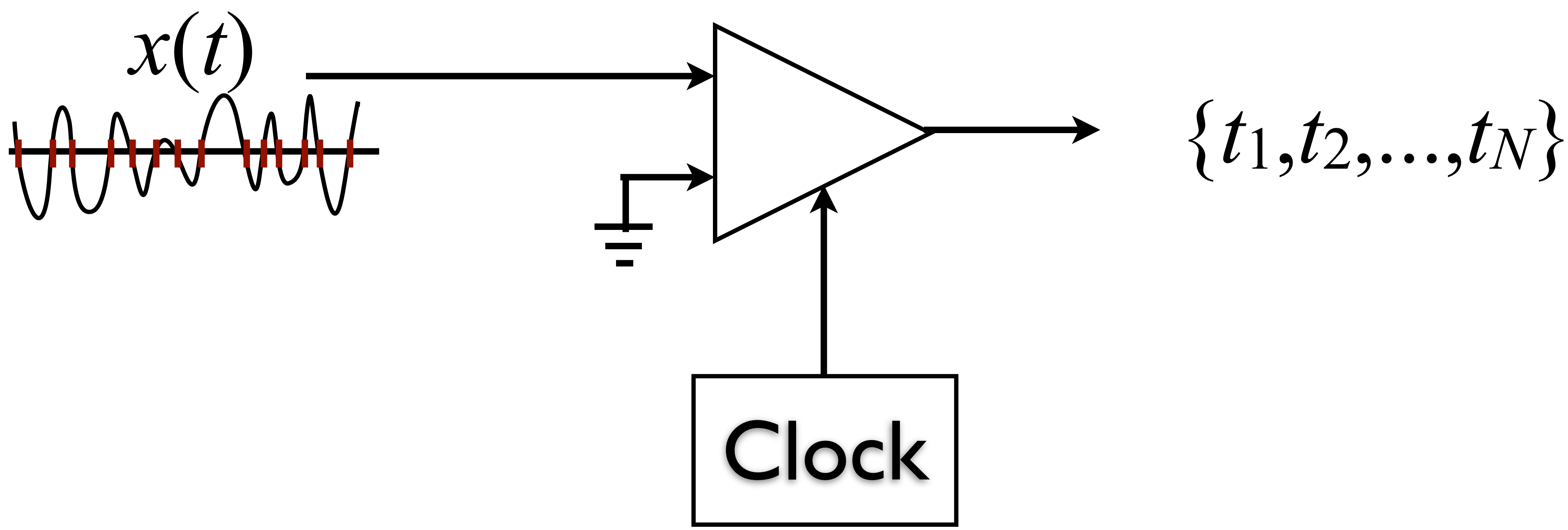
If $T=\{t_1,t_2,\ldots,t_N\}$ are the zero crossings, then the desired signal is in the nullspace of $\Phi$: $\Phi_T\mathbf{x} = 0.$

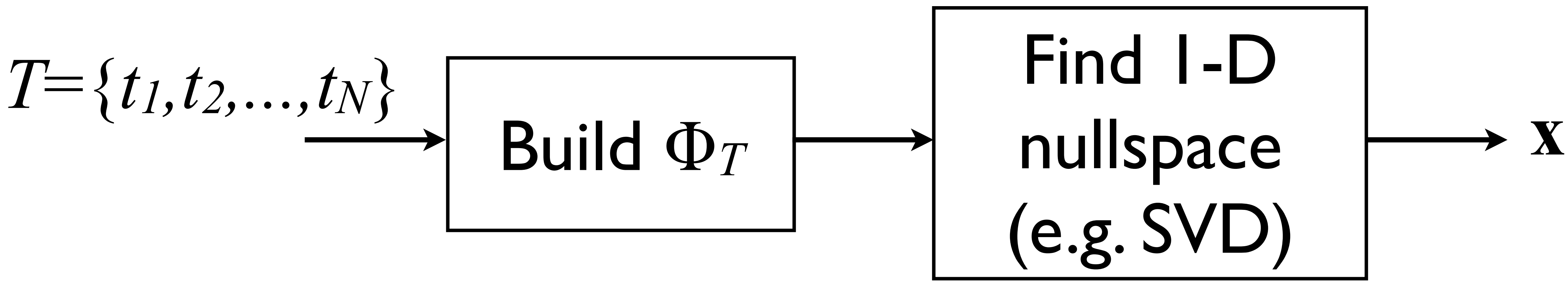
Logan's theorem$\Rightarrow$ $\Phi_T$ has a one-dimensional nullspace.

# Signal Acquisition and Reconstruction

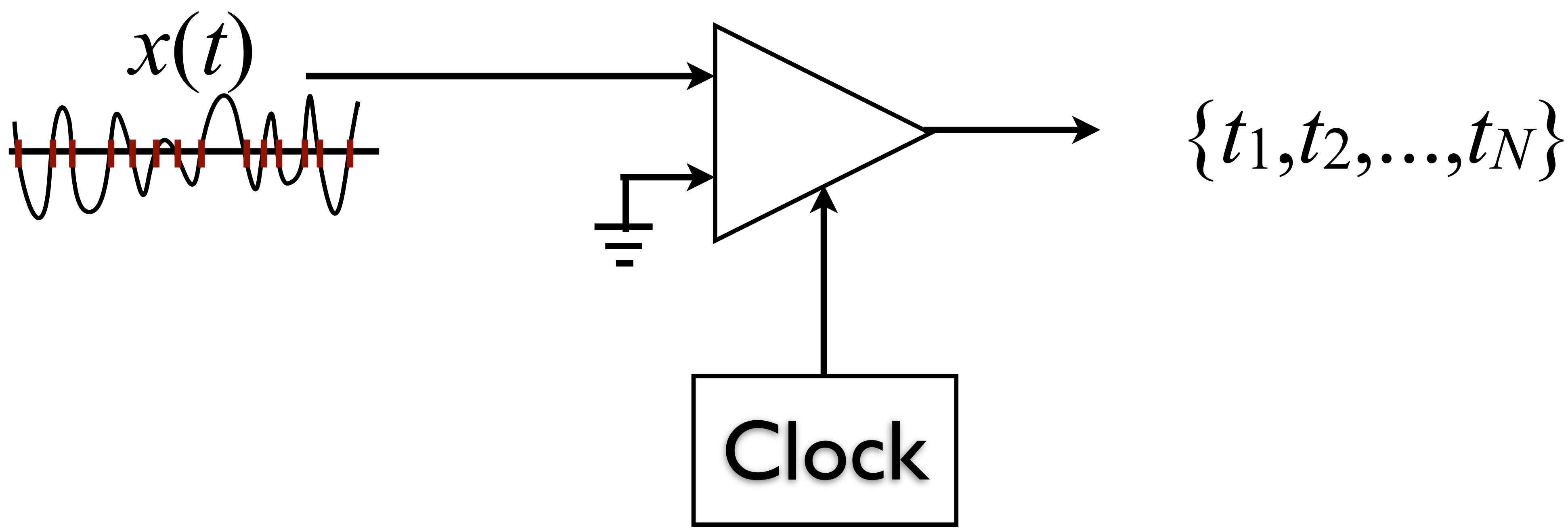


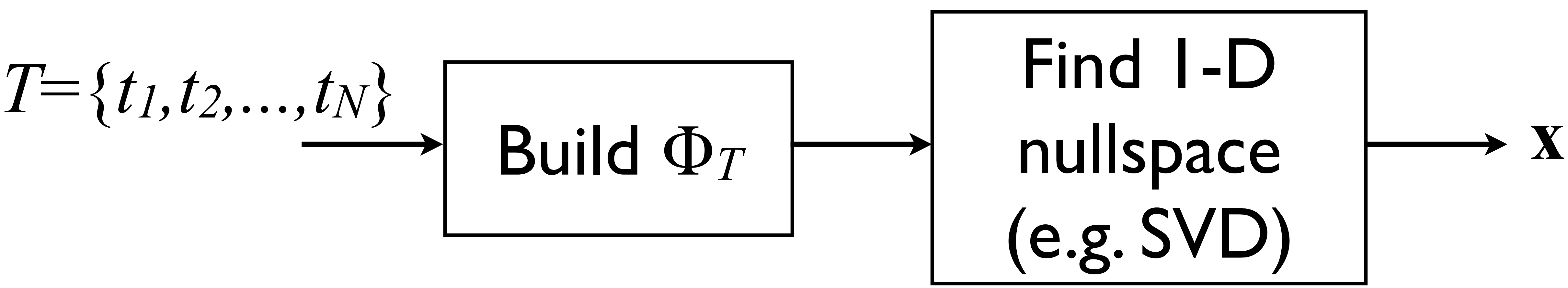
# Signal Acquisition and Reconstruction

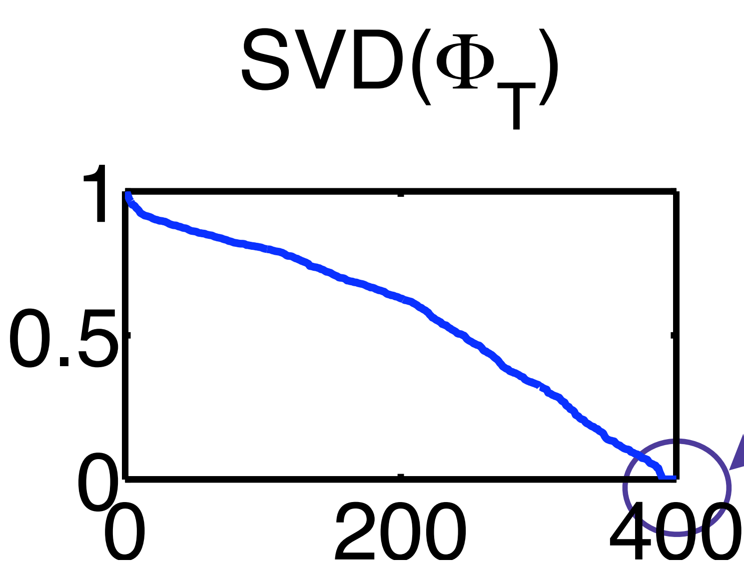


# Signal Acquisition and Reconstruction




In practice: noise and quantization.

No nullspace!

Many small singular values.

Ambiguity!

# Sparse Reconstruction

## $\ell_1$ minimization:

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1$$

subject to $\Phi\mathbf{x} = 0$

# Sparse Reconstruction

## $\ell_1$ minimization:

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1$$

subject to $\Phi\mathbf{x} = 0$

## Relaxation:

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 + \frac{\lambda}{2}\|\Phi\mathbf{x}\|_2^2$$

# Sparse Reconstruction

## $\ell_1$ minimization:

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1$$

$$\text{subject to} \quad \Phi\mathbf{x} = 0$$

## Relaxation:

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 + \frac{\lambda}{2}\|\Phi\mathbf{x}\|_2^2$$

## Unit energy constraint:

$$\begin{aligned} \widehat{\mathbf{x}} &= \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 + \frac{\lambda}{2}\|\Phi\mathbf{x}\|_2^2 \\ &\quad \text{subject to } \|\mathbf{x}\|_2 = 1 \end{aligned}$$

# Fixed Point Equilibrium

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 + \frac{\lambda}{2}\|\Phi\mathbf{x}\|_2^2$$

$$\text{subject to } \|\mathbf{x}\|_2 = 1$$

Unconstrained minimization:

$$\mathrm{Cost}(\mathbf{x}) = g(\mathbf{x}) + \frac{\lambda}{2} f(\Phi\mathbf{x})$$

$$\mathrm{Cost}'(\mathbf{x}) = g'(\mathbf{x}) + \frac{\lambda}{2}\Phi^* f'(\Phi\mathbf{x})$$

where:

$$(g'(\mathbf{x}))_i = \begin{cases} -1 & x_i < 0 \\ [-1, 1] & x_i = 0 \\ +1 & x_i > 0 \end{cases}$$

No change if gradients are projected on unit sphere

# Minimization Algorithm (based on FPC [Hale, Yin, Zhang, '07])

**Big Picture:** Gradient descent until equilibrium.

Initialization parameters: $\widehat{\mathbf{x}}, \ \tau$

1. Compute quadratic gradient: $\mathbf{h} = \Phi^T \Phi \widehat{\mathbf{x}}$
2. Project onto sphere: $\mathbf{h}_p = \mathbf{h} - \langle \widehat{\mathbf{x}}, \mathbf{h} \rangle$
3. Quadratic gradient descent: $\widehat{\mathbf{x}} \leftarrow \widehat{\mathbf{x}} - \tau \mathbf{h}_p$
4. Shrink ($\ell_1$ gradient descent):

$$\widehat{x}_i \quad \leftarrow \quad \operatorname{sign}(\widehat{x}_i) \max \left\{ |\widehat{x}_i| - \frac{\tau}{\lambda}, 0 \right\}$$

5. Normalize: $\widehat{\mathbf{x}} \quad \leftarrow \quad \dfrac{\widehat{\mathbf{x}}}{\|\widehat{\mathbf{x}}\|}$
6. Iterate until equilibrium.

# Optimization on the Sphere

$$\widehat{\mathbf{x}} \quad = \quad \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 + \frac{\lambda}{2}\|\Phi\mathbf{x}\|_2^2$$
$$\text{subject to } \|\mathbf{x}\|_2 = 1$$

Optimization is not convex.

Convergence to global optimum not guaranteed.

# Optimization on the Sphere

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 + \frac{\lambda}{2}\|\Phi\mathbf{x}\|_2^2$$
$$\text{subject to } \|\mathbf{x}\|_2 = 1$$

Optimization is not convex.

Convergence to global optimum not guaranteed.

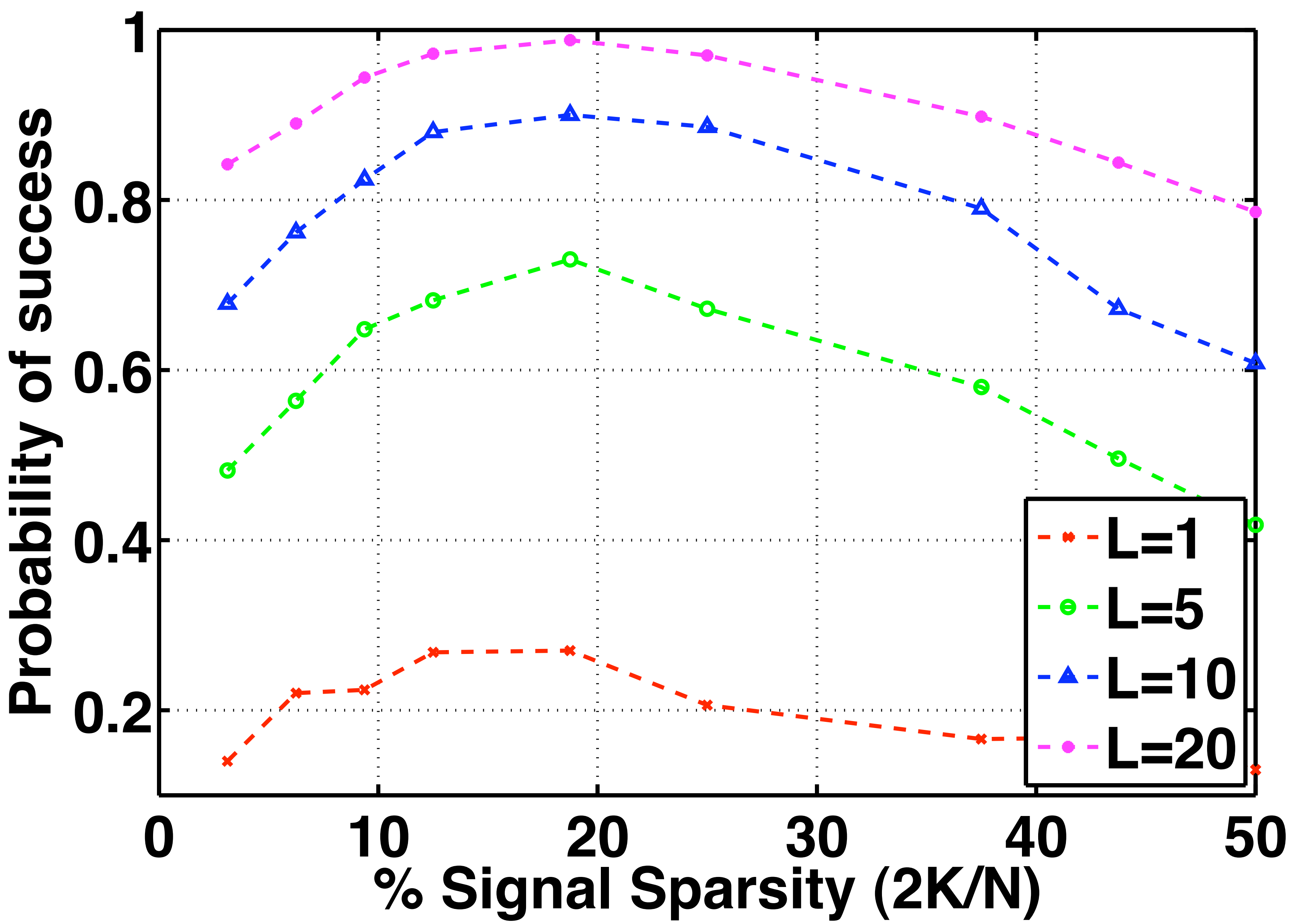
Exploit randomness:

- Execute $L$ times with random initializations.
- Pick best solution.

If $P$=$P$(success for 1 execution), then
$P$(overall success)=1-(1-$P$)$^L$

# Results

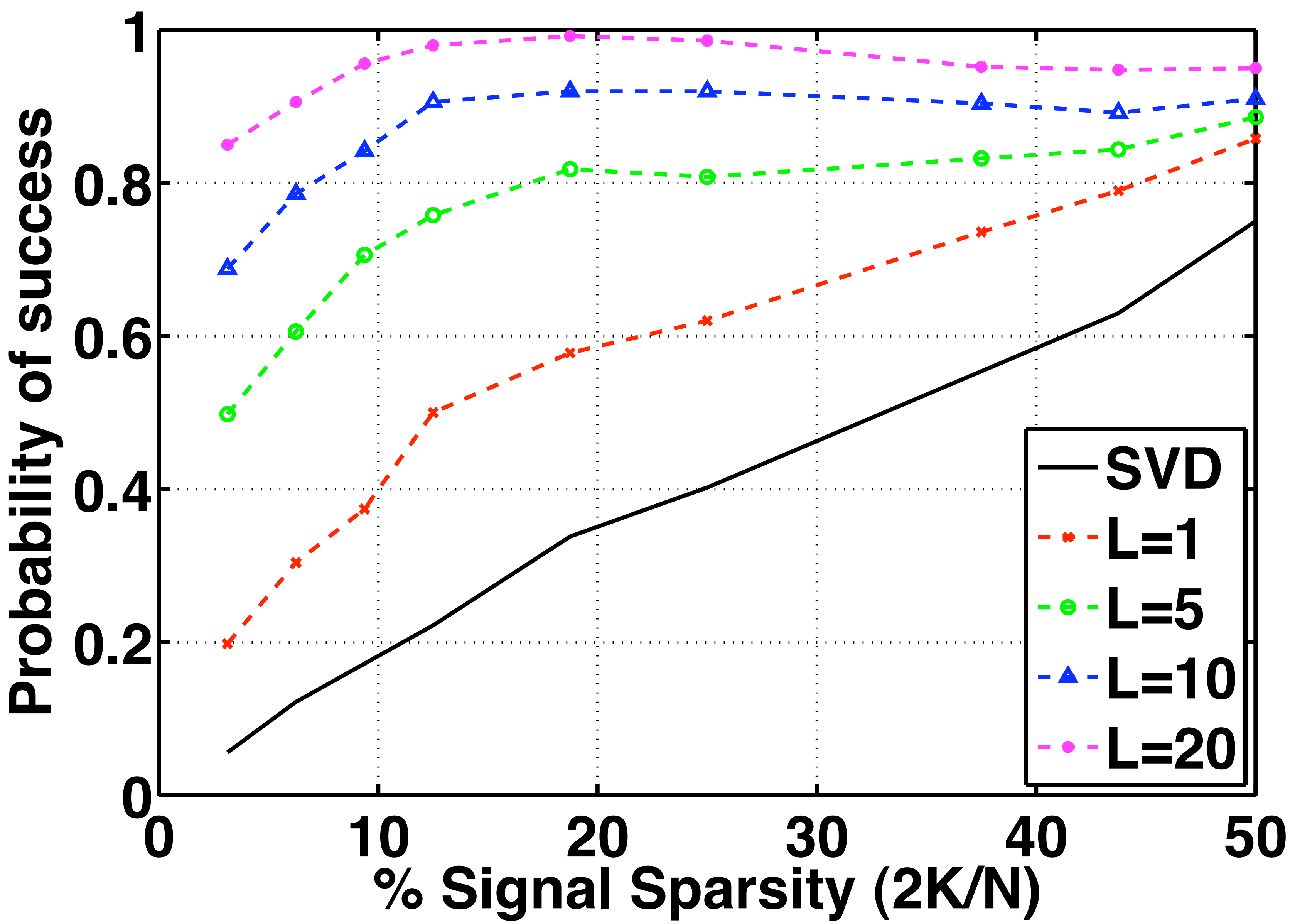
## Probability of Success




L=number of random initializations
N=256 coefficients

# Results

## Probability of Success




L=number of random initializations
N=256 coefficients

# Further Relaxation (w/ Cinmay Hegde)

Optimization on sphere:

$$\widehat{\mathbf{x}} \quad = \quad \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 + \frac{\lambda}{2}\|\Phi\mathbf{x}\|_2^2$$
$$\text{subject to } \|\mathbf{x}\|_2 = 1$$

Relaxation of sphere constraint:

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 + \frac{\lambda_1}{2}\|\Phi\mathbf{x}\|_2^2 + \lambda_2 \left| \|\mathbf{x}\|_2^2 - 1 \right|^2$$

We can now use standard $\ell_1$ algorithms!

# $\ell_1$ minimization formulation

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 + \frac{\lambda_1}{2}\|\Phi\mathbf{x}\|_2^2 + \lambda_2 \left|\|\mathbf{x}\|_2^2 - 1\right|^2$$

Let:

$$\widetilde{\Phi} = \left[\begin{array}{c} c\mathbf{x} \\ \Phi \end{array}\right], \; c = \left(\frac{\lambda_2}{\lambda_1}\right)^2$$

At equilibrium:

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 + \frac{\lambda_1}{2}\left\|\widetilde{\Phi}\mathbf{x} - \left[\begin{array}{c} c \\ \mathbf{0} \end{array}\right]\right\|_2^2$$

# Reweighted FPC algorithm

Initialization parameters: $\widehat{\mathbf{x}},\ \lambda_1,\ \lambda_2$

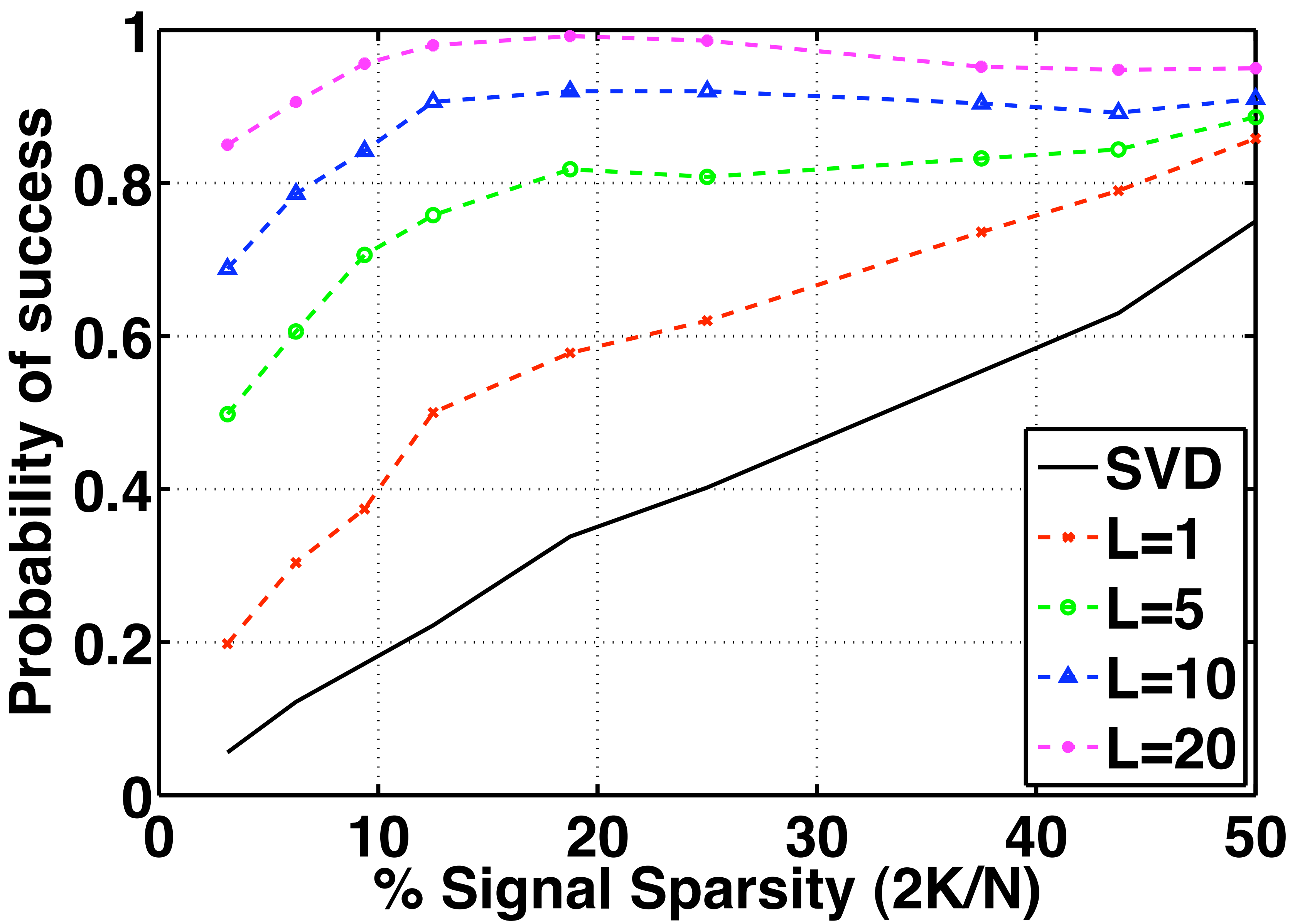
1. Build $\widetilde{\Phi} = \begin{bmatrix} c\widehat{\mathbf{x}} \\ \Phi \end{bmatrix},\ c = \left(\frac{\lambda_2}{\lambda_1}\right)^2$

2. Estimate using FPC:

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 + \frac{\lambda_1}{2} \left\| \widetilde{\Phi}\mathbf{x} - \begin{bmatrix} c \\ \mathbf{0} \end{bmatrix} \right\|_2^2$$

3. Iterate until equilibrium.

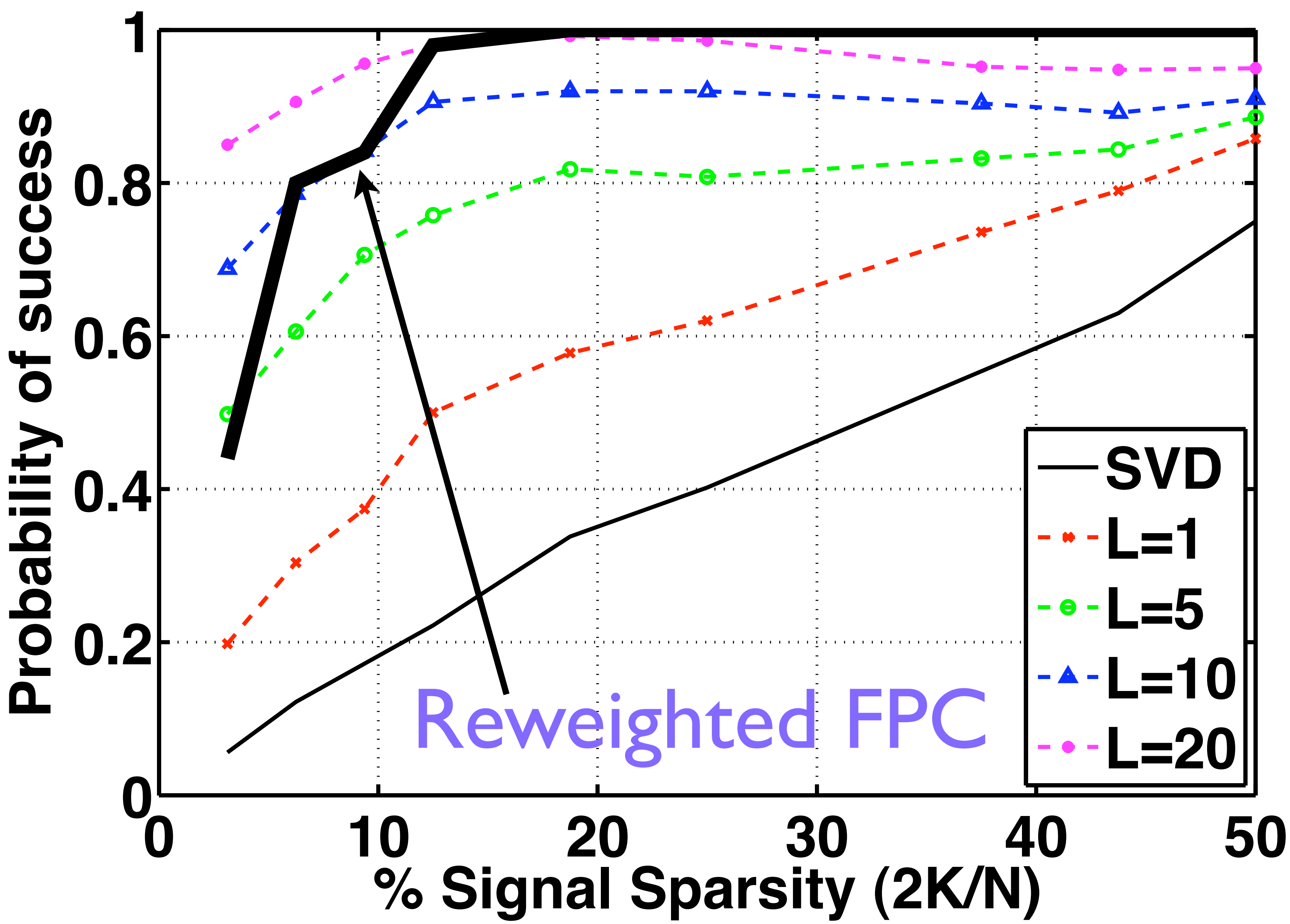
Probability of Success




L=number of random initializations
N=256 coefficients

# Results

## Probability of Success




L=number of random initializations
N=256 coefficients

# Case II:
# 1-bit Compressive Sensing

# 1-Bit Compressive Sensing

Q: Can we quantize measurements to 1-bit:

$$\mathbf{y} = \mathrm{sign}(\Phi \mathbf{x})$$

$$y_i = \mathrm{sign}(\langle \phi_i, \mathbf{x} \rangle)$$

and recover the signal (within a positive scaling factor)?

# 1-Bit Compressive Sensing

Q: Can we quantize measurements to 1-bit:

$$\mathbf{y} = \text{sign}(\Phi \mathbf{x})$$
$$y_i = \text{sign}(\langle \phi_i, \mathbf{x} \rangle)$$

and recover the signal (within a positive scaling factor)?

1-bit measurements are inexpensive.

Focus on bits rather than measurements.

Exact recovery is not possible.

# Reconstruction from 1-bit Measurements

Sign information from 1-bit measurements:

$$y_i = \mathrm{sign}(\Phi \mathbf{x})_i \Leftrightarrow y_i \cdot (\Phi \mathbf{x})_i \geq 0$$

Reconstruction should enforce model.

Reconstruction should be consistent with measurements.

$$\begin{aligned} \widehat{\mathbf{x}} \quad &= \quad \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 \\ \text{subject to} \quad &\quad y_i \cdot (\Phi \mathbf{x})_i \geq 0 \end{aligned}$$

# Reconstruction from 1-bit Measurements

Sign information from 1-bit measurements:

$$y_i = \mathrm{sign}(\Phi \mathbf{x})_i \Leftrightarrow y_i \cdot (\Phi \mathbf{x})_i \geq 0$$

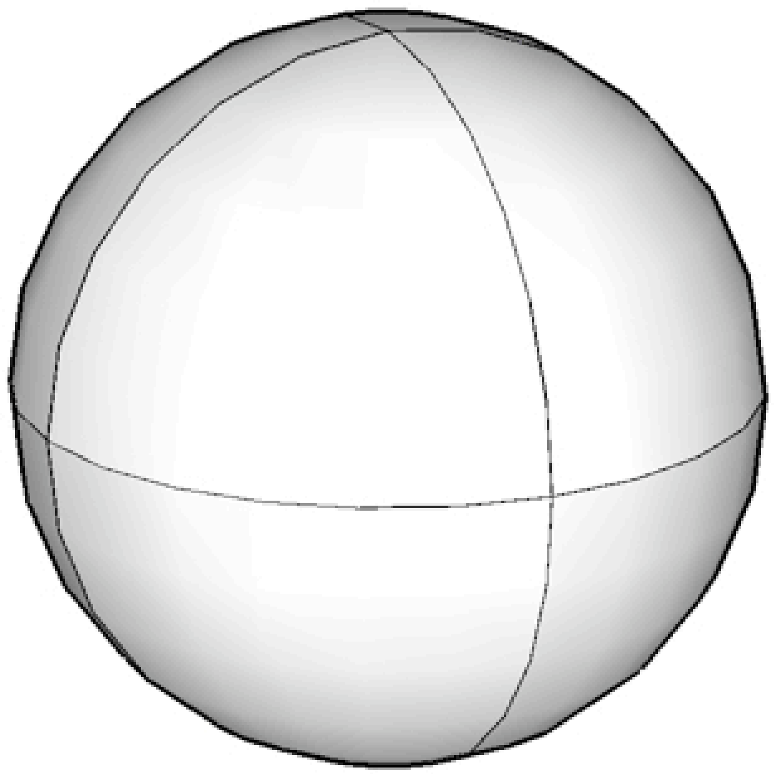
Reconstruction should enforce model.

Reconstruction should be consistent with measurements.

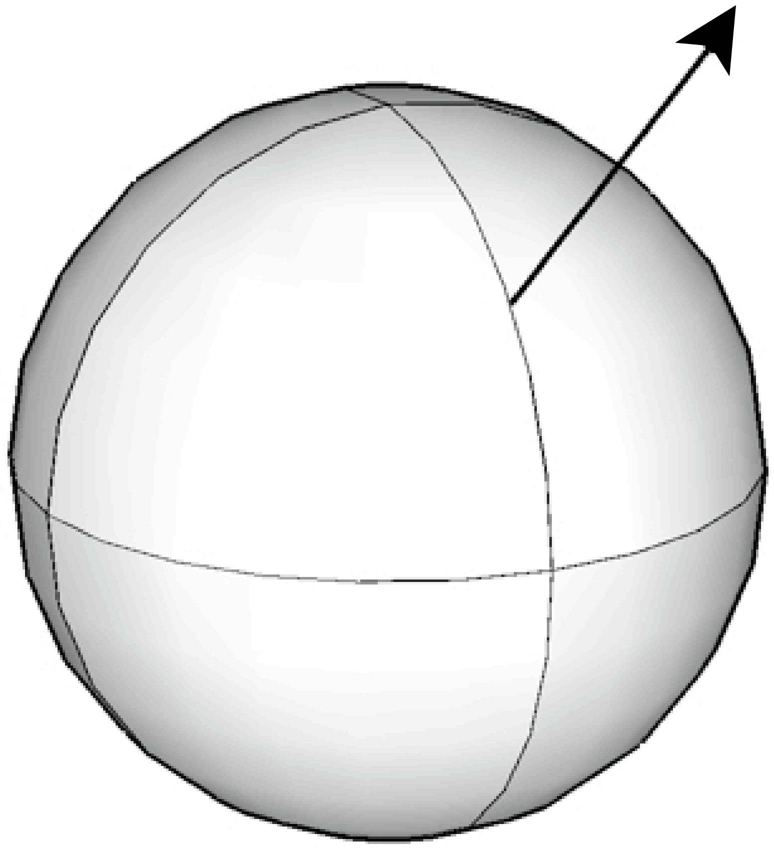
Reconstruction should enforce a non-trivial solution.

$$\begin{aligned} \widehat{\mathbf{x}} \quad &= \quad \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 \\ \text{subject to} \quad &\quad y_i \cdot (\Phi \mathbf{x})_i \geq 0 \\ \text{and} \quad &\quad \|\mathbf{x}\|_2 = 1 \end{aligned}$$

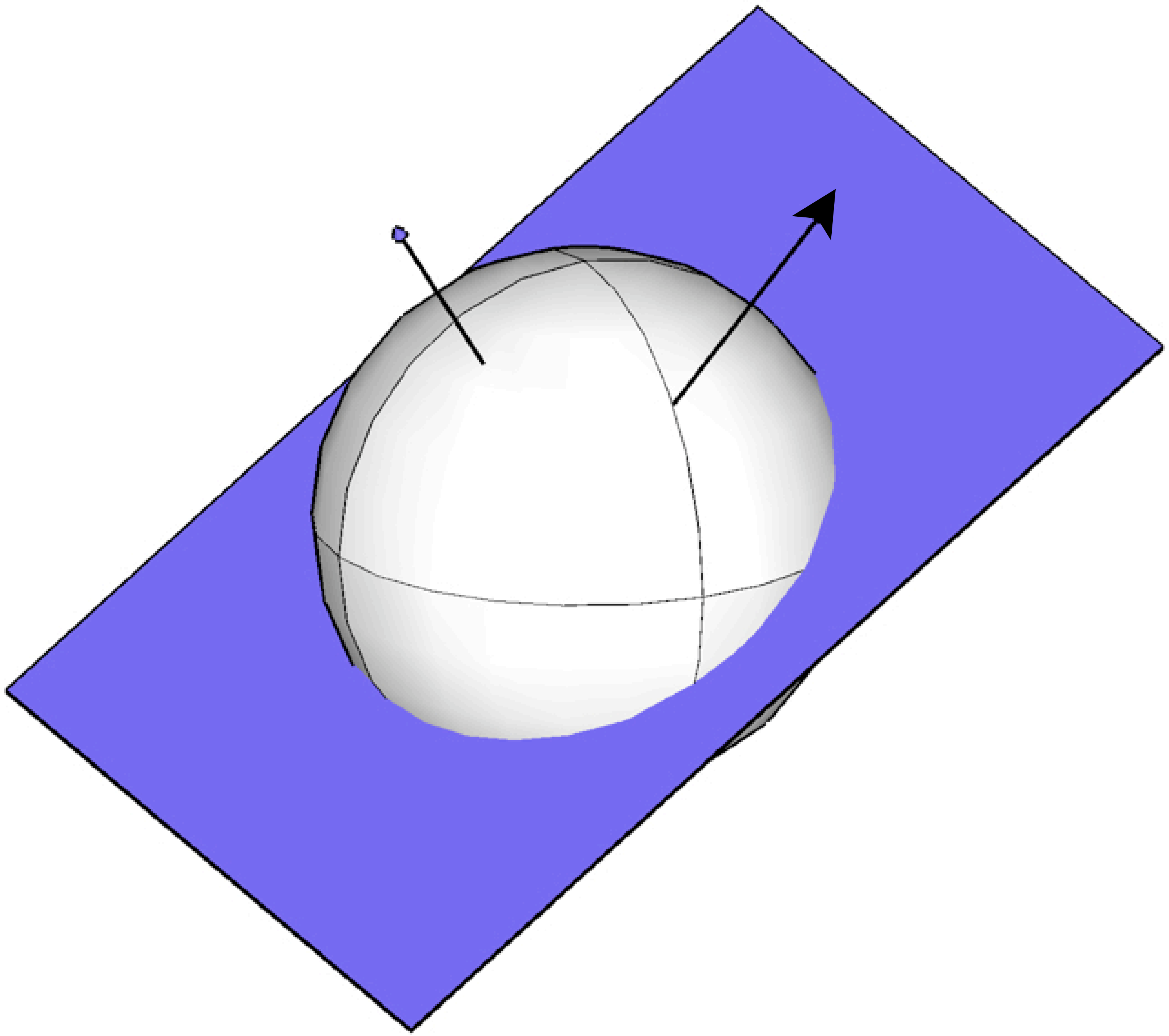
# Information in 1-bit Measurements



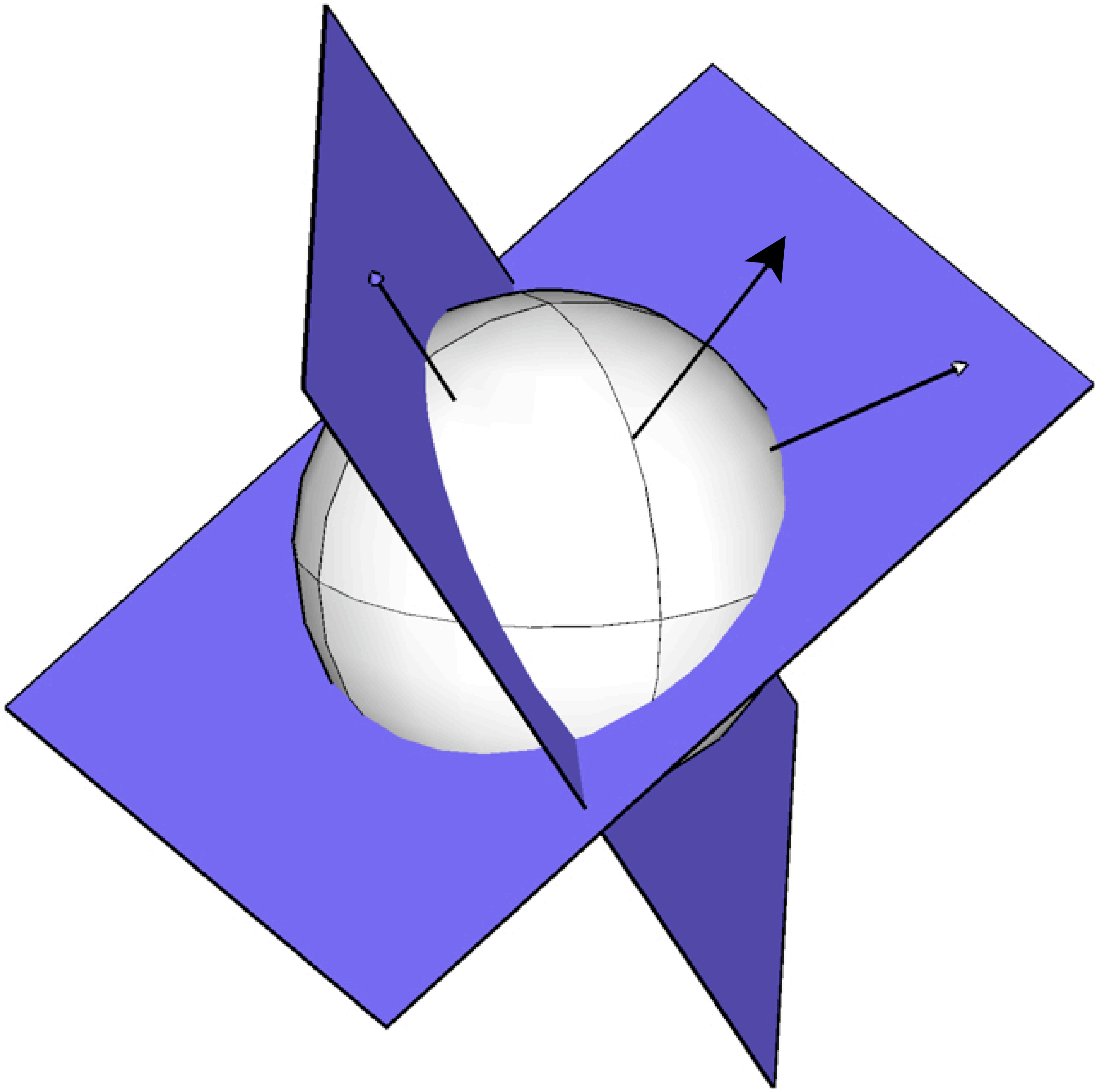
# Information in 1-bit Measurements



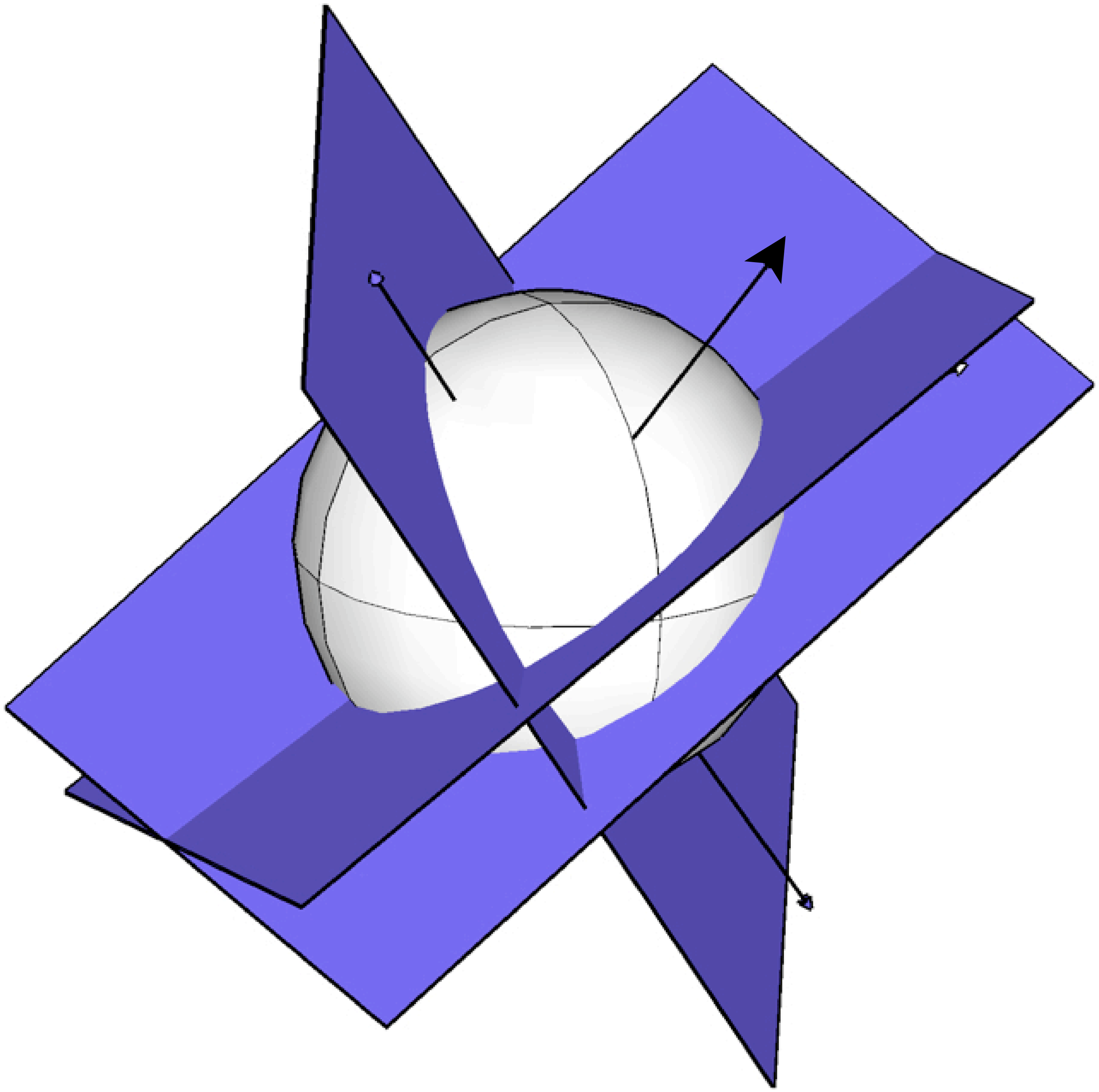
# Information in 1-bit Measurements



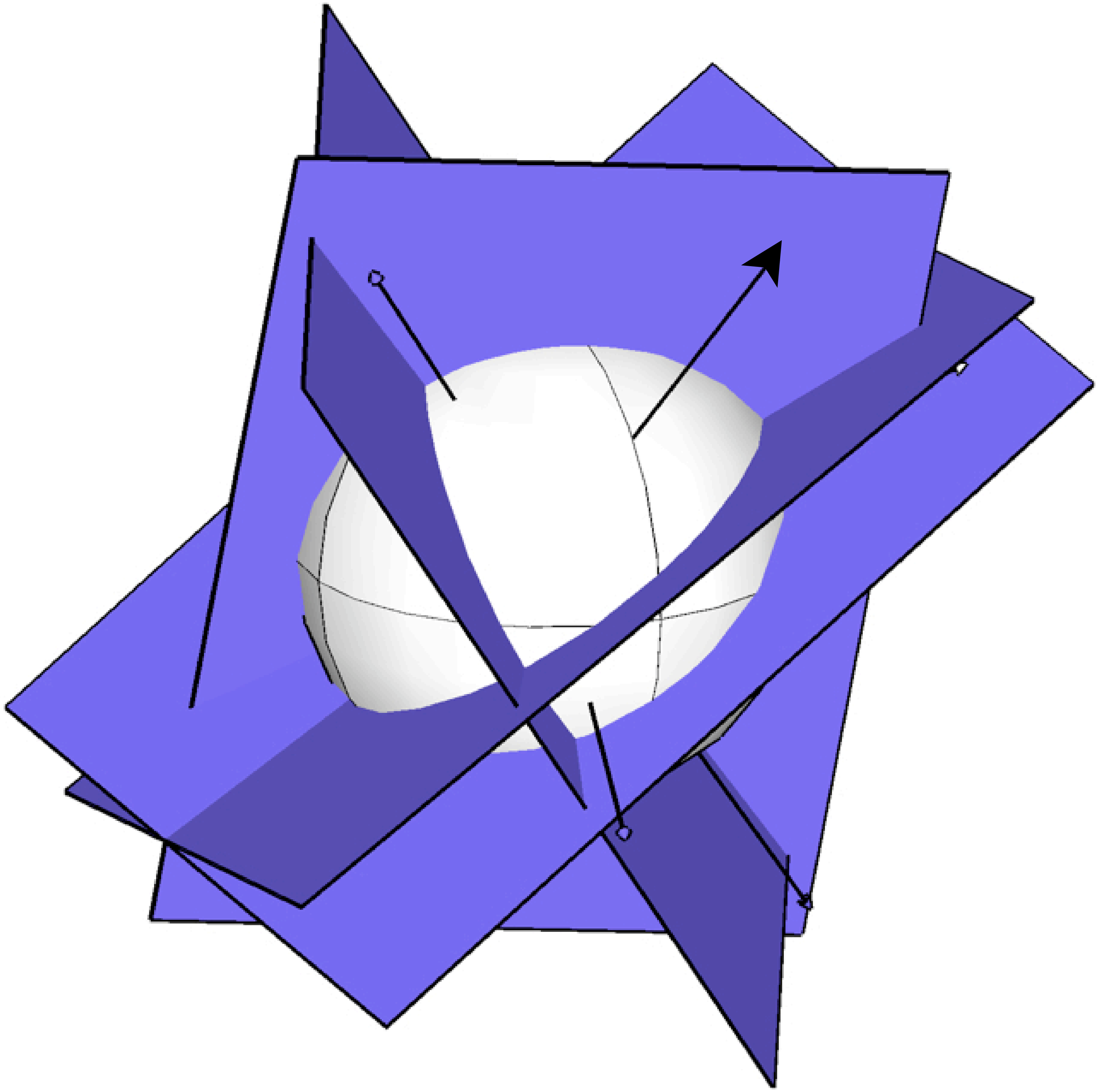
# Information in 1-bit Measurements



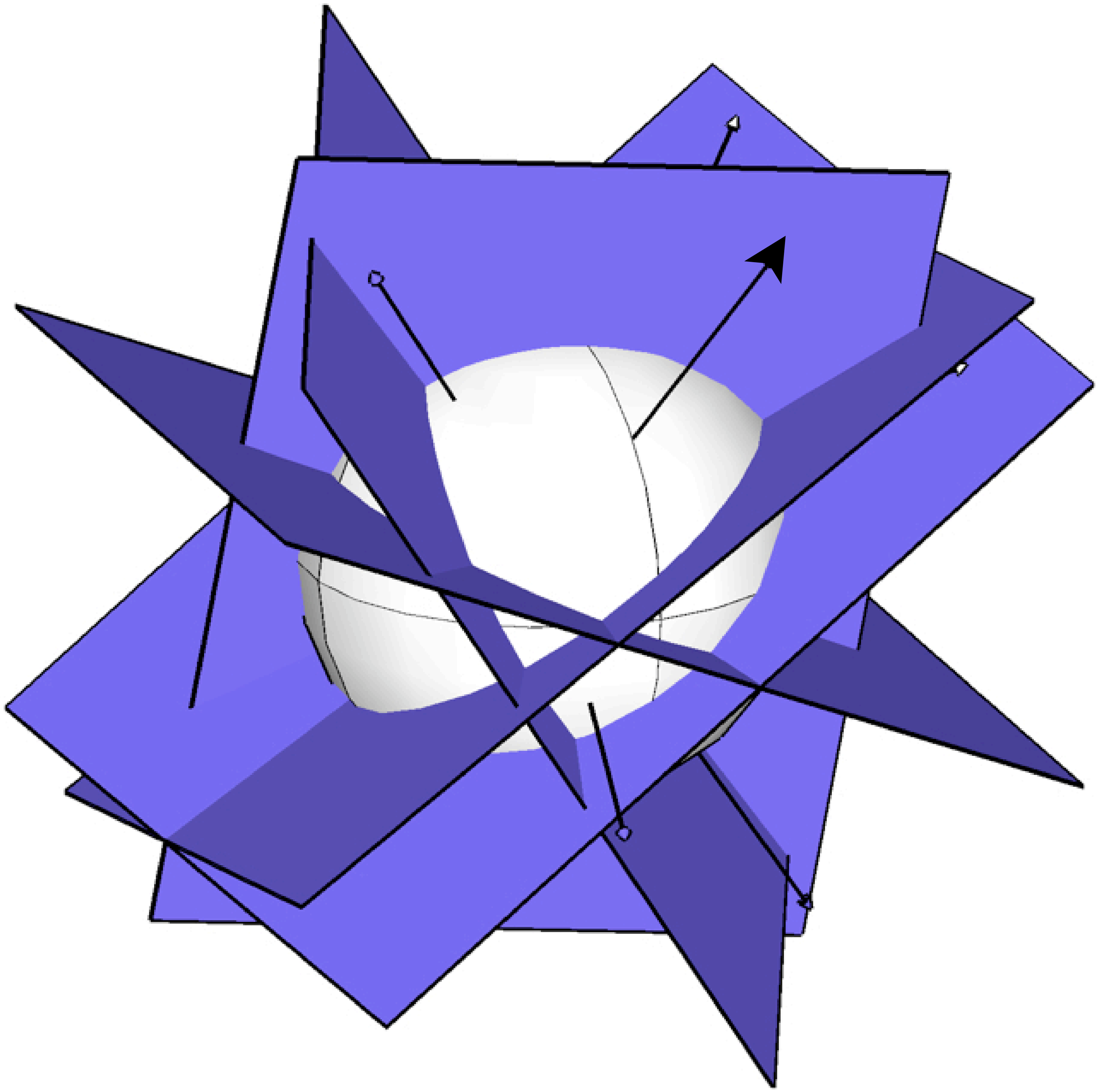
# Information in 1-bit Measurements



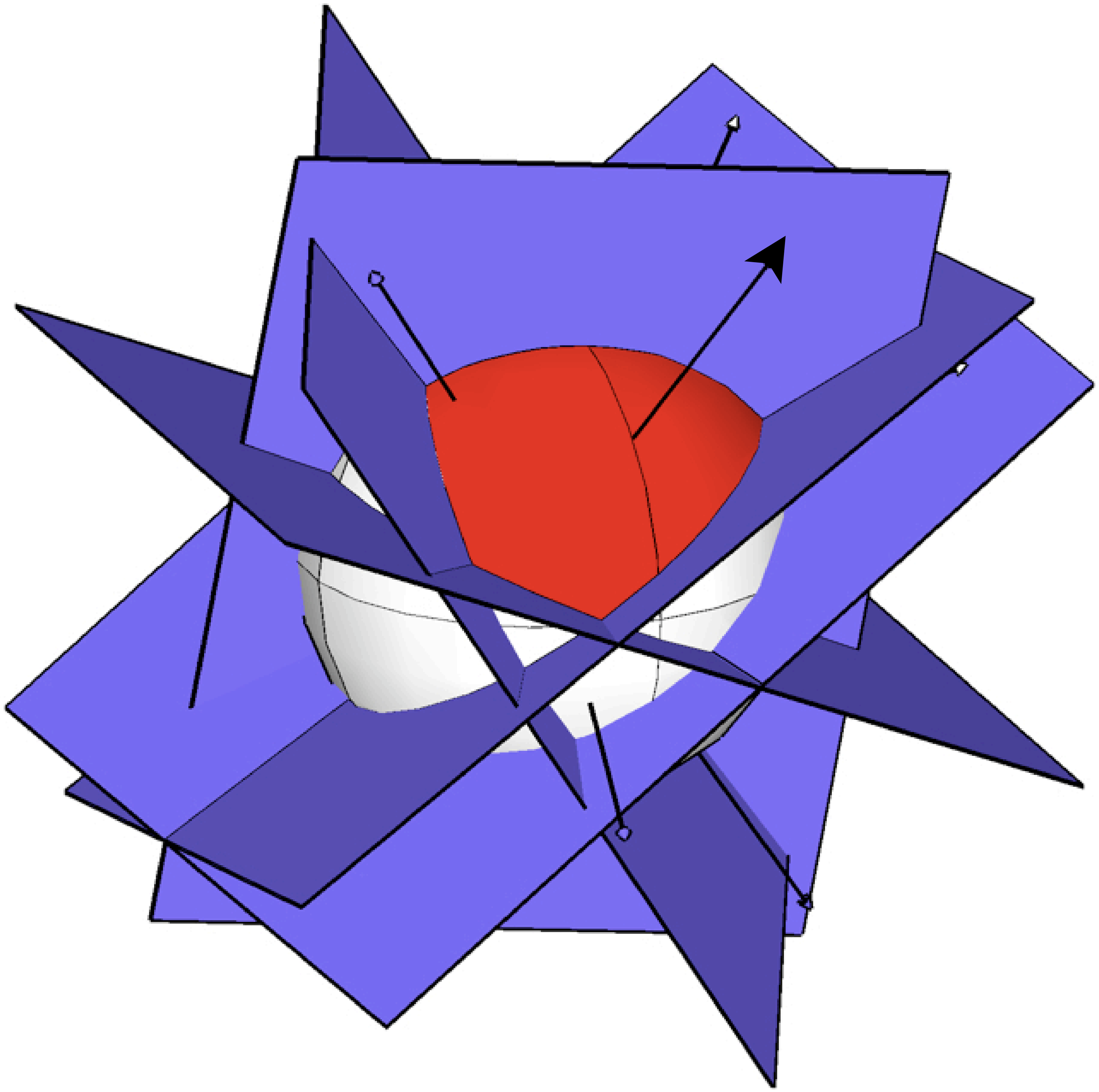
# Information in 1-bit Measurements



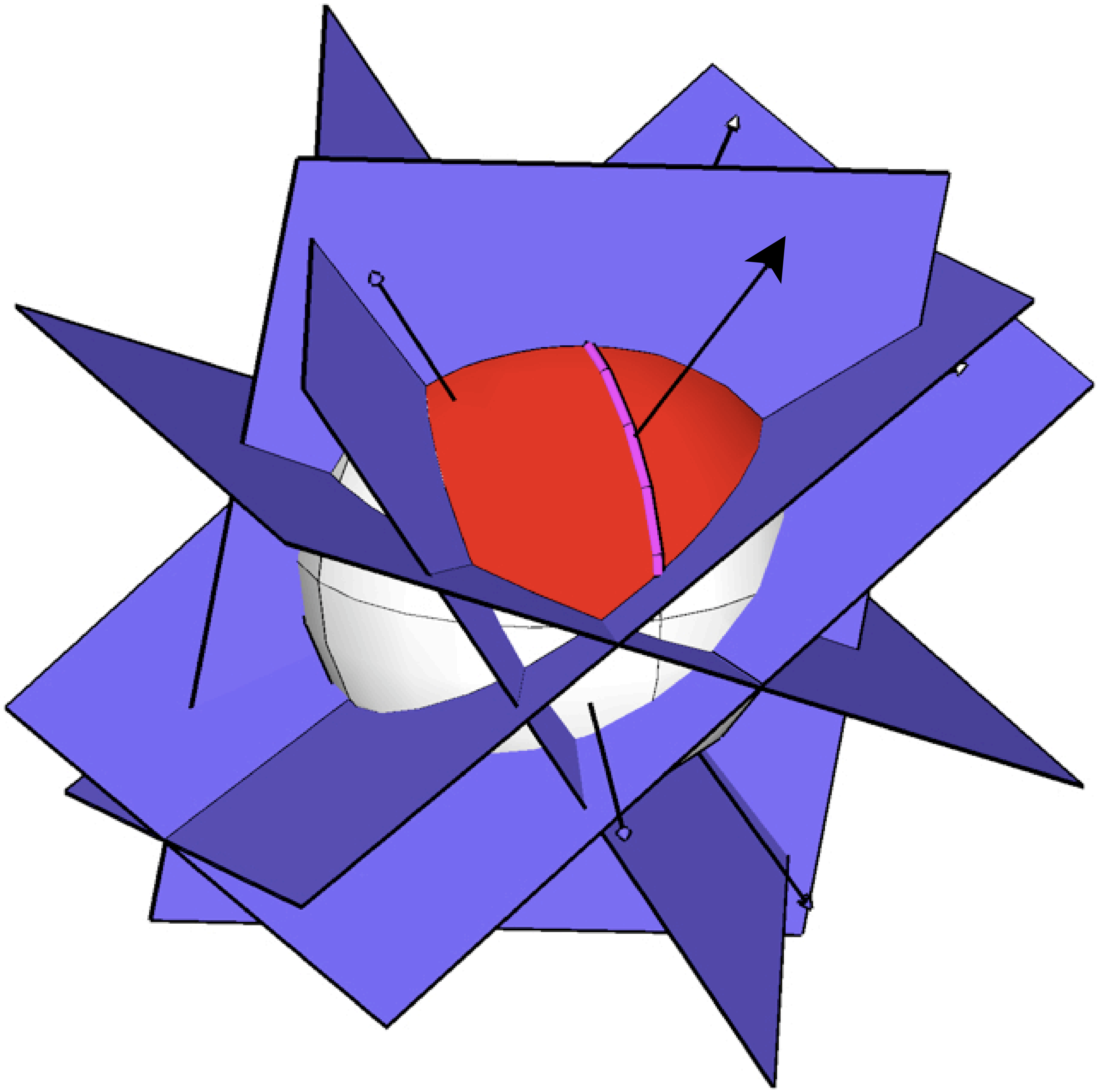
# Information in 1-bit Measurements



# Information in 1-bit Measurements



# Information in 1-bit Measurements



# Constraint Relaxation

$$\begin{aligned} \widehat{\mathbf{x}} \quad &= \quad \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 \\ \text{subject to} \quad &\quad y_i \cdot (\Phi \mathbf{x})_i \geq 0 \\ \text{and} \quad &\quad \|\mathbf{x}\|_2 = 1 \end{aligned}$$

# Constraint Relaxation

$$\begin{aligned} \widehat{\mathbf{x}} &= \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 \\ \text{subject to} \quad & y_i \cdot (\Phi \mathbf{x})_i \geq 0 \\ \text{and} \quad & \|\mathbf{x}\|_2 = 1 \end{aligned}$$

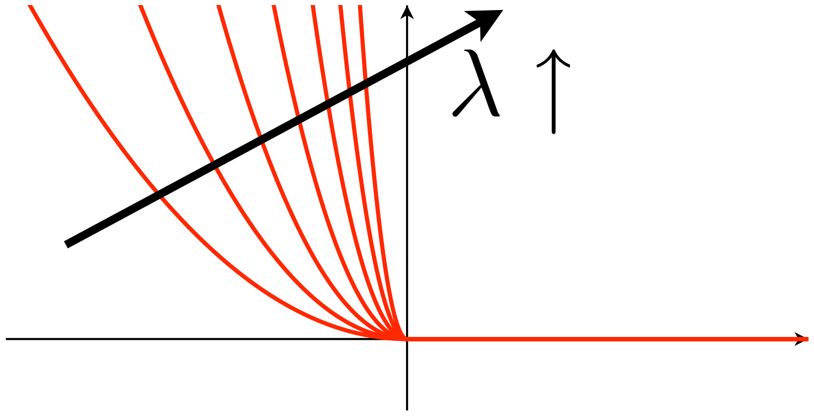
We relax the inequality constraints:

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 + \frac{\lambda}{2} \sum_i f\left(y_i \cdot (\Phi \mathbf{x})\right)$$

$$\text{subject to } \|\mathbf{x}\|_2 = 1$$

where $f(x)$ is a one sided quadratic:

$$f(x) = \begin{cases} x^2 & x \leq 0 \\ 0 & x > 0 \end{cases}$$



# Fixed point equilibrium

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 + \frac{\lambda}{2}\sum_i f\left(y_i \cdot (\Phi\mathbf{x})\right)$$

$$\text{subject to } \|\mathbf{x}\|_2 = 1$$

## Unconstrained minimization:

$$\begin{aligned} \mathrm{Y} &\equiv \mathrm{diag}(\mathbf{y}) \\ \mathrm{Cost}(\mathbf{x}) &= g(\mathbf{x}) + \frac{\lambda}{2} f(\mathrm{Y}\Phi\mathbf{x}) \\ \mathrm{Cost}'(\mathbf{x}) &= g'(x) + \frac{\lambda}{2} (\mathrm{Y}\Phi)^T f(\mathrm{Y}\Phi\mathbf{x}) \end{aligned}$$

$$(g'(\mathbf{x}))_i = \begin{cases} -1 & x_i < 0 \\ [-1,1] & x_i = 0 \\ +1 & x_i > 0 \end{cases} \text{ and } \left(\frac{f'(\mathbf{x})}{2}\right)_i = \begin{cases} -x_i & x_i \leq 0 \\ 0 & x_i > 0 \end{cases}$$

No change if gradients are projected on unit sphere.

# Minimization Algorithm

**Big Picture: Gradient descent until equilibrium.**

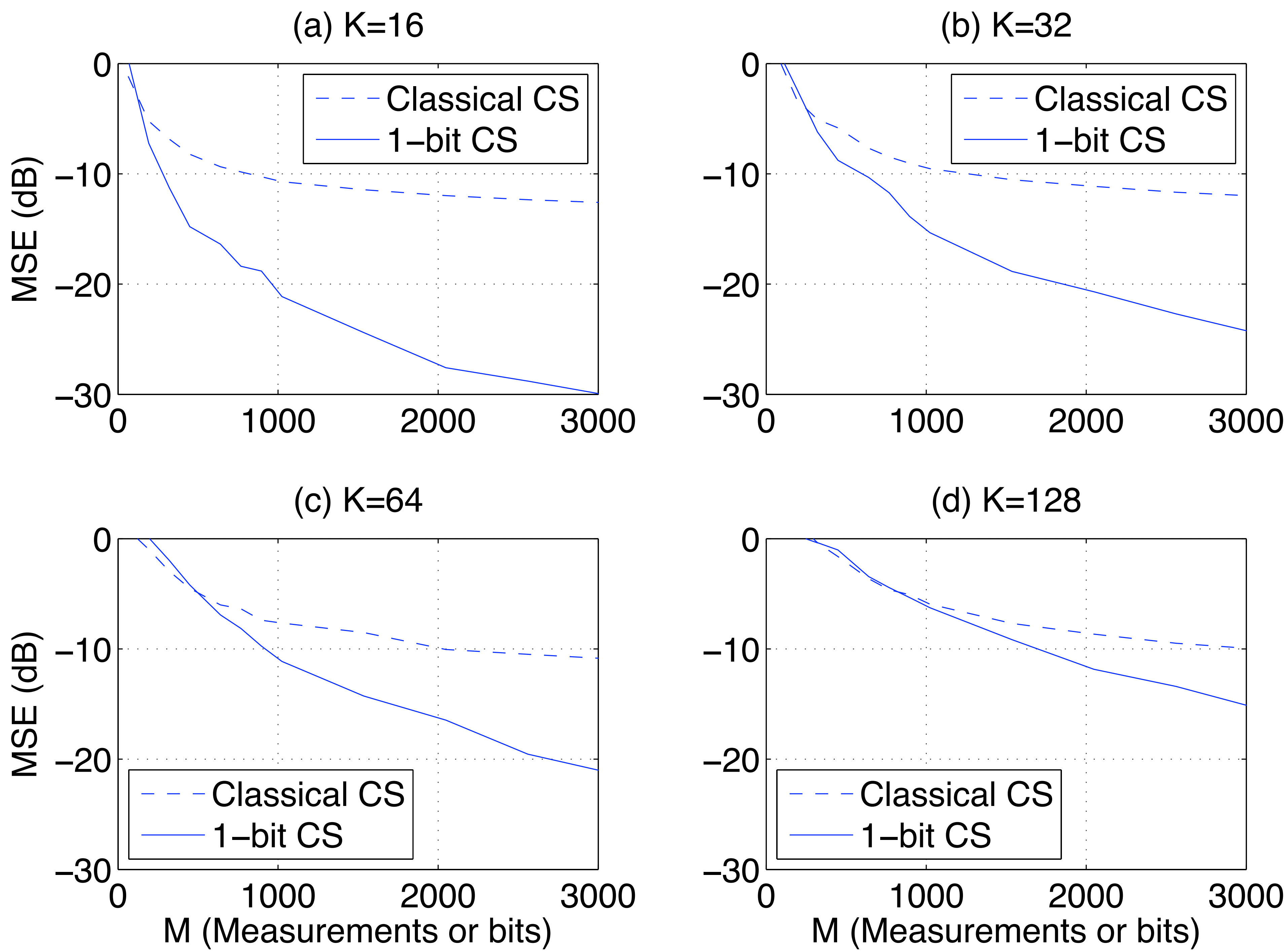
Initialization parameters: $\widehat{\mathbf{x}},\ \tau$

1. Compute quadratic gradient: $\mathbf{h} = (\mathrm{Y}\Phi)^T f'(\mathrm{Y}\Phi\mathbf{x})$
2. Project onto sphere: $\mathbf{h}_p = \mathbf{h} - \langle \widehat{\mathbf{x}}, \mathbf{h} \rangle$
3. Quadratic gradient descent: $\widehat{\mathbf{x}} \leftarrow \widehat{\mathbf{x}} - \tau \mathbf{h}_p$
4. Shrink ($\ell_1$ gradient descent):

$$\widehat{x}_i \quad \leftarrow \quad \mathrm{sign}(\widehat{x}_i) \max\left\{ |\widehat{x}_i| - \frac{\tau}{\lambda}, 0 \right\}$$

5. Normalize: $\widehat{\mathbf{x}} \quad \leftarrow \quad \dfrac{\widehat{\mathbf{x}}}{\|\widehat{\mathbf{x}}\|}$
6. Iterate until equilibrium.

# Reconstruction Error ($N$=512)



# 1-bit Sampling of Images

If the signal is an image, we have more information!
(i.e., a better signal model)

Images are sparse in wavelets and positive:

$$\begin{aligned} \mathbf{x} &= \mathbf{W}\alpha \\ x_i &\geq 0 \\ \text{and } \alpha &\text{ is sparse} \end{aligned}$$

Incorporate better model in the reconstruction:

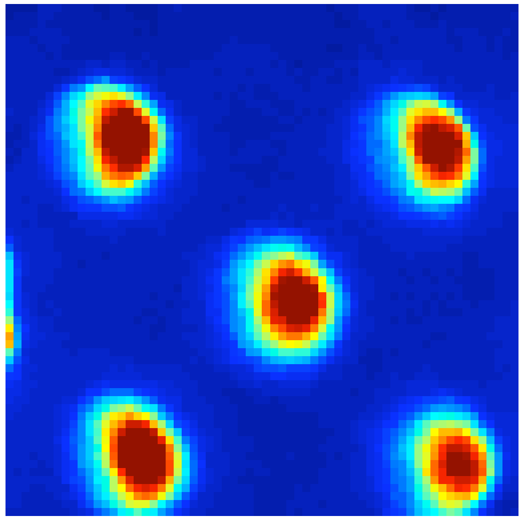
$$\hat{\alpha} = \arg\min_{\alpha} \|\alpha\|_1$$

$$\text{subject to } y_i \times (\Phi\mathbf{W})_i \geq 0$$

$$\text{and } (\mathbf{W}\alpha)_i \geq 0$$

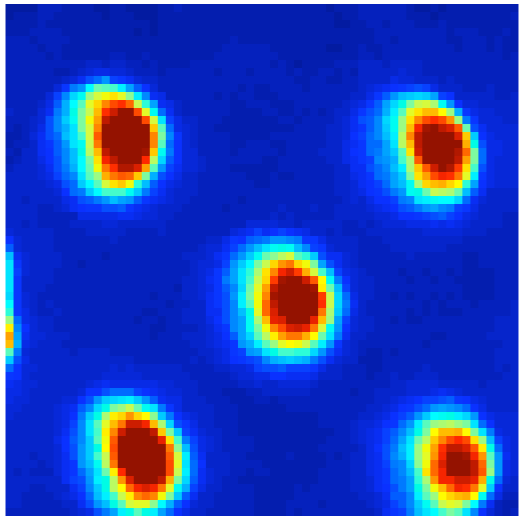
$$\text{and } \|\alpha\|_2 = 1$$

# Results



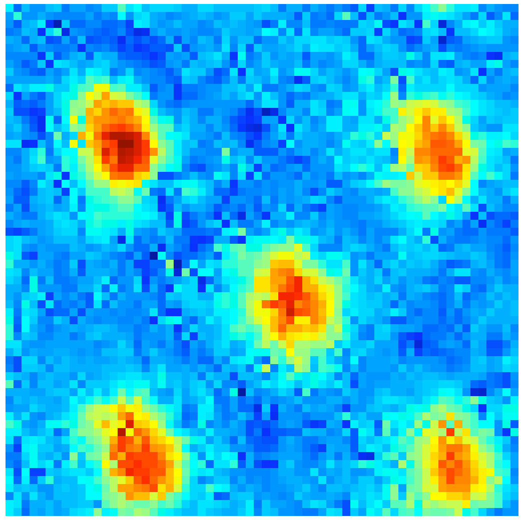
Original Image
4096 pixels
256 levels

# Results

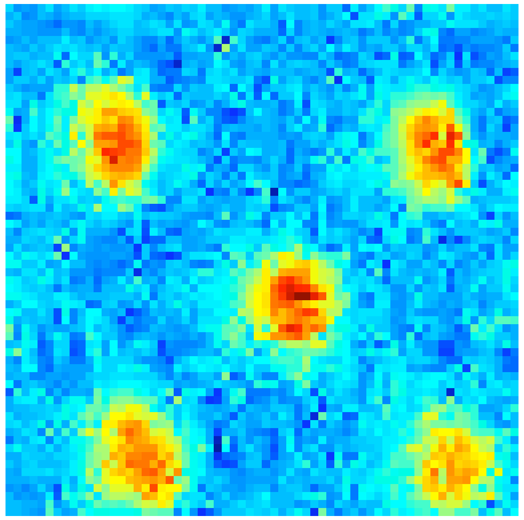


Original Image
4096 pixels
256 levels

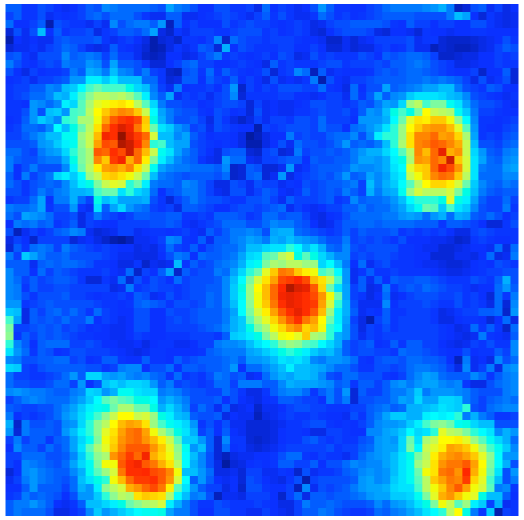
## Classical Compressive Sensing, 1 bit per pixel



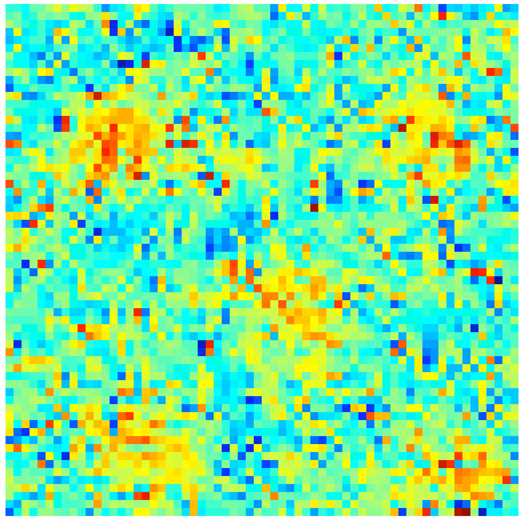
4096 measurements
1 bit per measurement



2048 measurements
2 bits per measurement

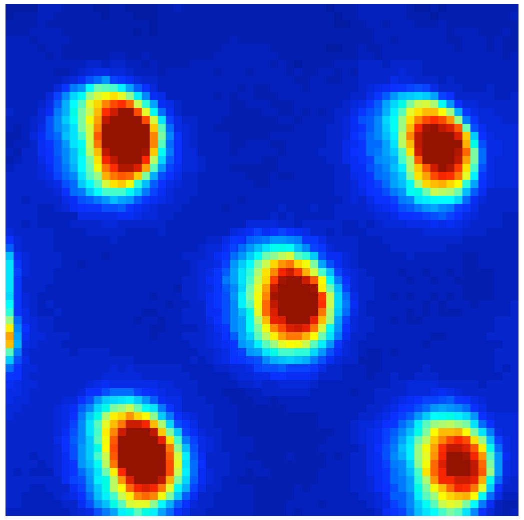


1024 measurements
4 bits per measurement



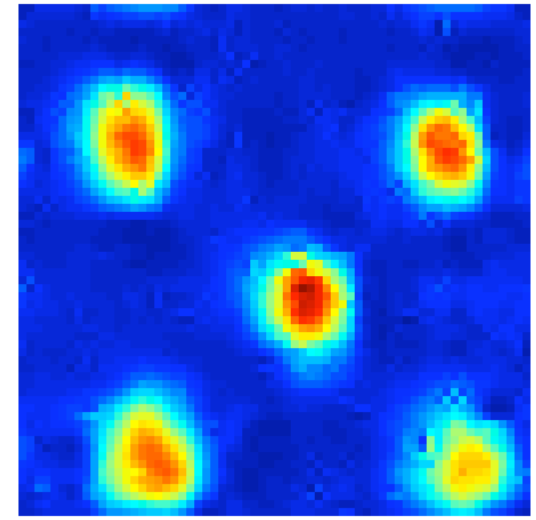
512 measurements
8 bits per measurement

# Results



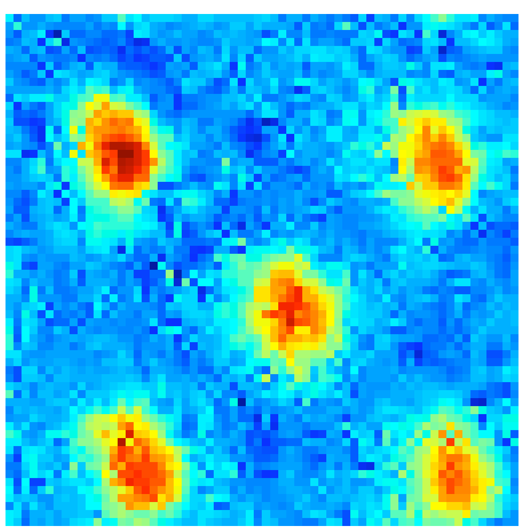
Original Image
4096 pixels
256 levels

## Reconstruction on unit sphere 1 bit per pixel

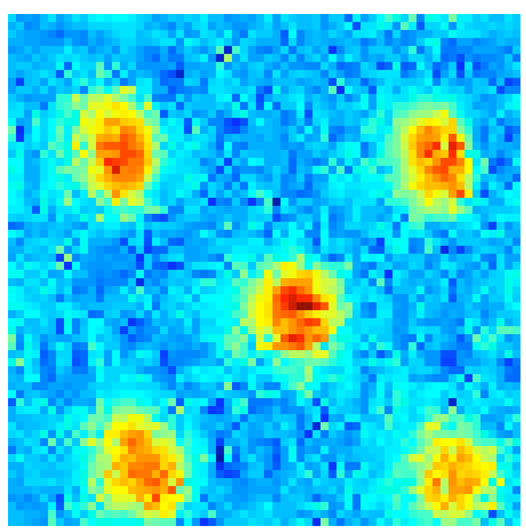


4096 measurements
1 bit per measurement

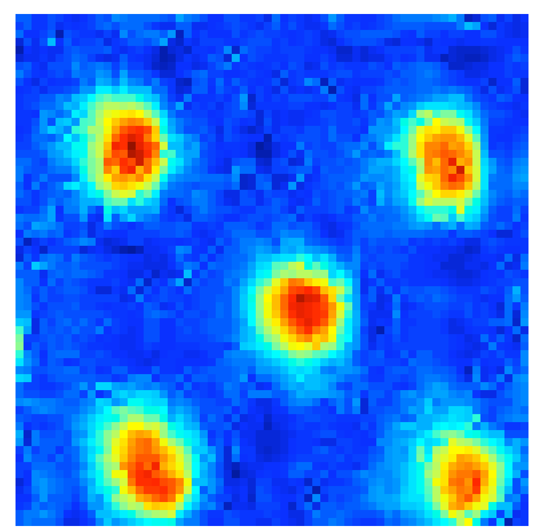
## Classical Compressive Sensing, 1 bit per pixel



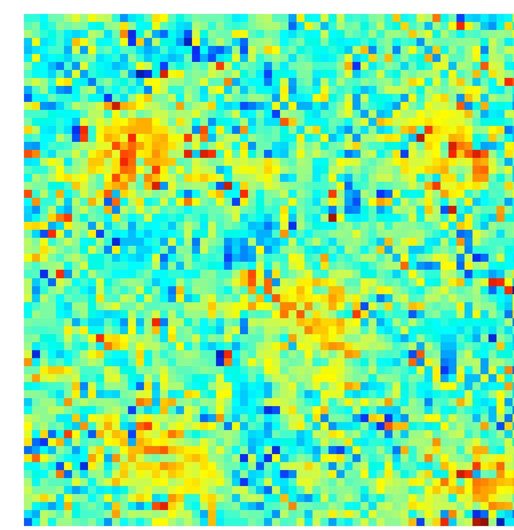
4096 measurements
1 bit per measurement



2048 measurements
2 bits per measurement



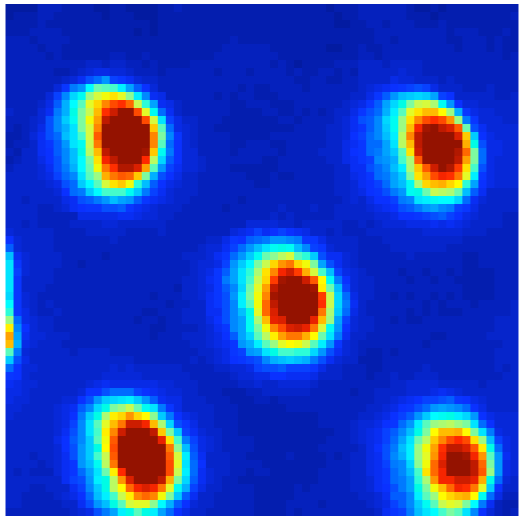
1024 measurements
4 bits per measurement



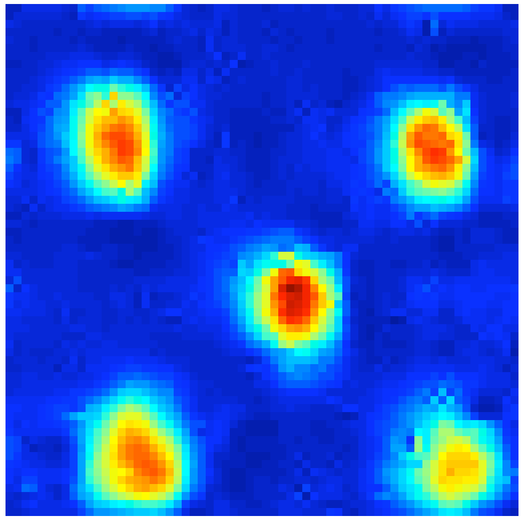
512 measurements
8 bits per measurement

# Results

Reconstruction on unit sphere

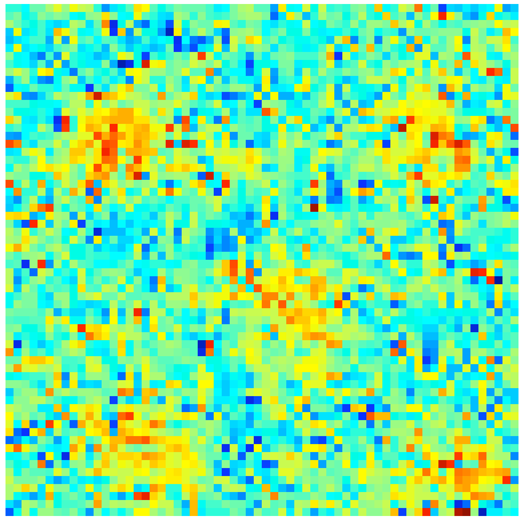


Original Image
4096 pixels
256 levels



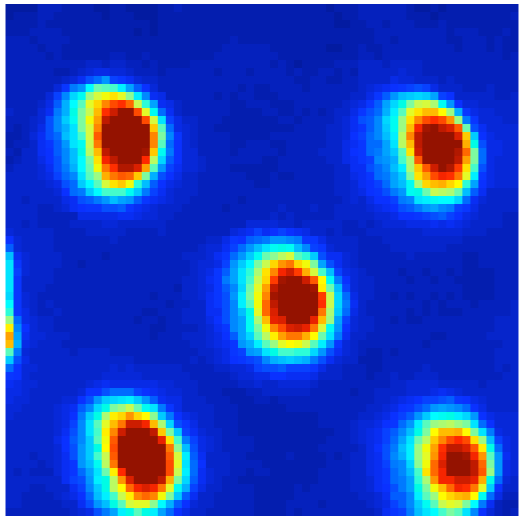
4096 measurements
1 bit per measurement
4096 bits (1 bit per pixel)

Classical Compressive Sensing



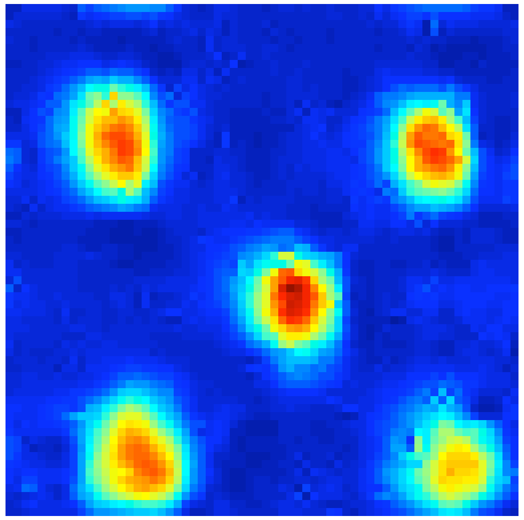
512 measurements
8 bits per measurement

# Results



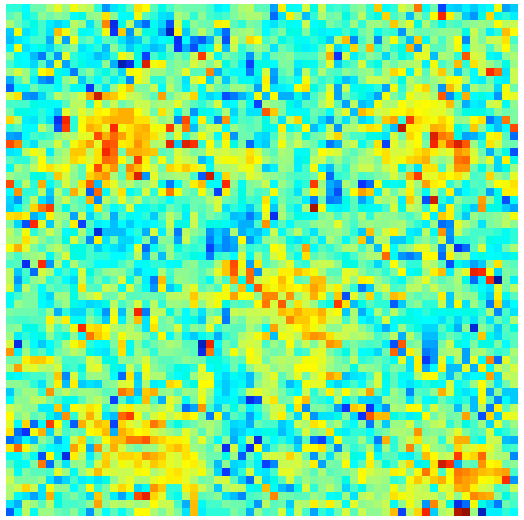
Original Image
4096 pixels
256 levels

## Reconstruction on unit sphere



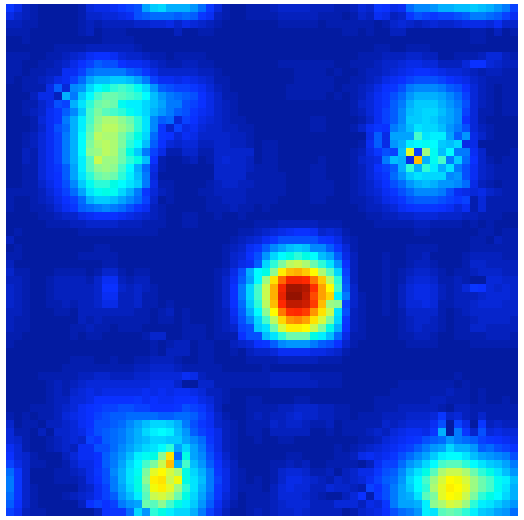
4096 measurements
1 bit per measurement
4096 bits (1 bit per pixel)

## Classical Compressive Sensing



512 measurements
8 bits per measurement

## Reconstruction on unit sphere



512 measurements
1 bit per measurement
512 bits (0.125 bits per pixel)

# Concluding Remarks

- Practical systems may eliminate amplitude information
- Reconstruction on the unit sphere is necessary
- The sphere is a well-behaved manifold
- Unit sphere constraint reduces the search space

- Several still open questions