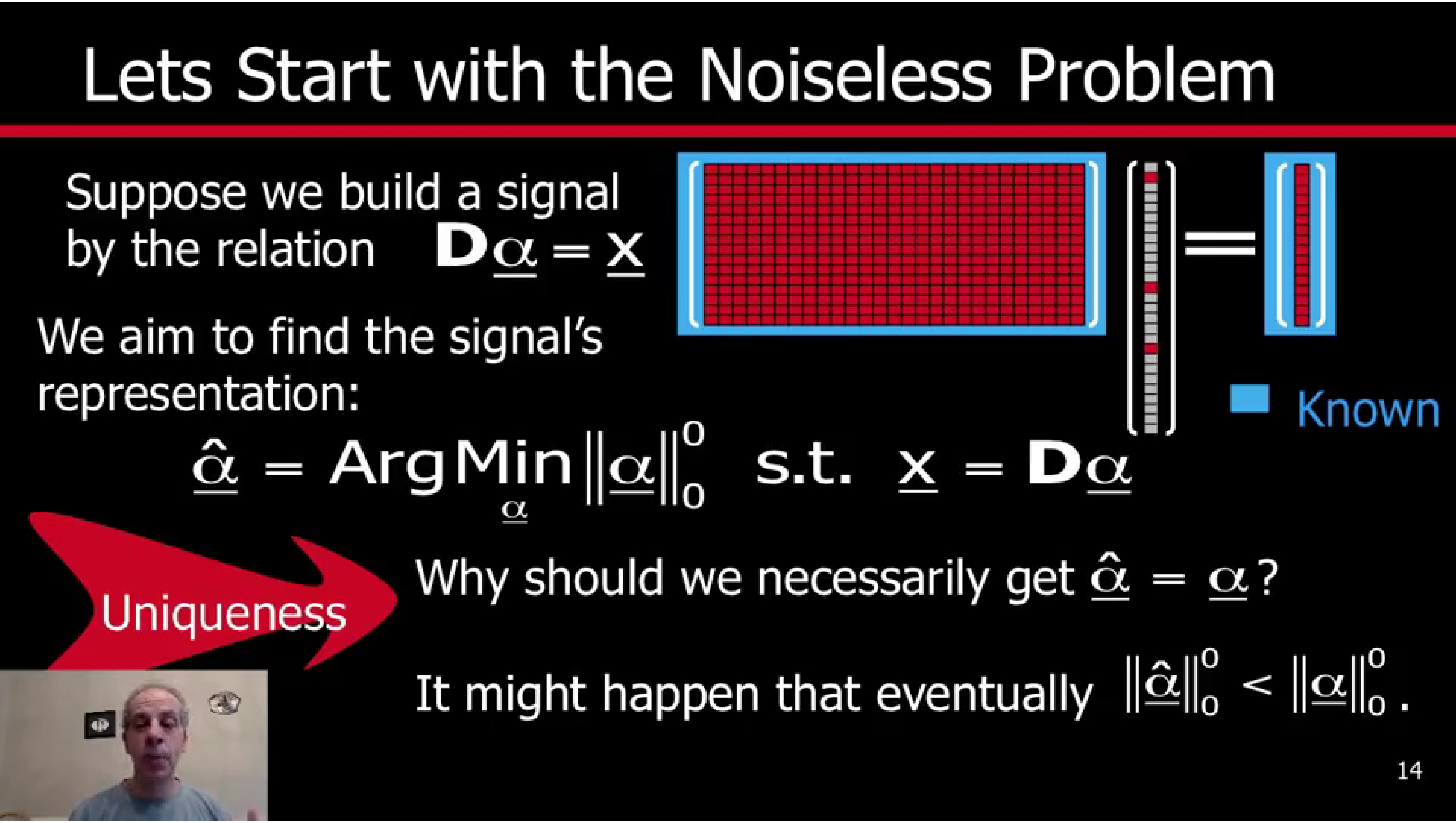
Sparse Modeling: Some Theory and Implementation

Image and Video Processing: From Mars to Hollywood with a Stop at the Hospital

Guillermo Sapiro

Duke
UNIVERSITY



Lets Start with the Noiseless Problem

Suppose we build a signal by the relation $\mathbf{D}\underline{\alpha} = \underline{\mathbf{x}}$

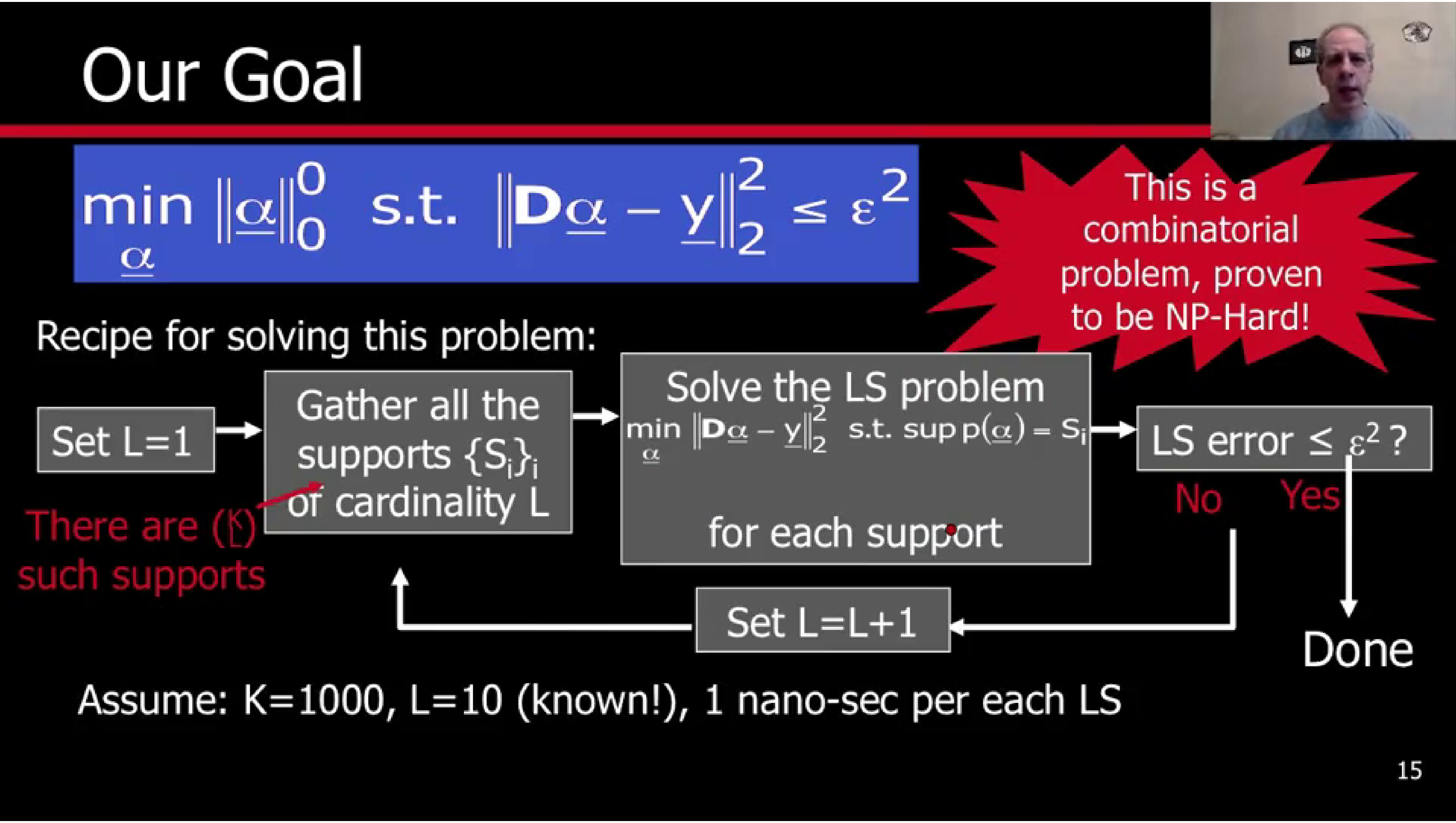
We aim to find the signal's representation:

$$\hat{\underline{\alpha}} = \underset{\underline{\alpha}}{\text{ArgMin}} \|\underline{\alpha}\|_0^0 \quad \text{s.t.} \quad \underline{\mathbf{x}} = \mathbf{D}\underline{\alpha}$$

Uniqueness

Why should we necessarily get $\hat{\underline{\alpha}} = \underline{\alpha}$?

It might happen that eventually $\|\hat{\underline{\alpha}}\|_0^0 < \|\underline{\alpha}\|_0^0$.



Our Goal

$$\min_{\underline{\alpha}} \|\underline{\alpha}\|_0^0 \quad \text{s.t.} \quad \|\mathbf{D}\underline{\alpha} - \underline{y}\|_2^2 \le \varepsilon^2$$

This is a combinatorial problem, proven to be NP-Hard!

Recipe for solving this problem:

Set L=1

Gather all the supports $\{S_i\}_i$ of cardinality L

There are $\binom{K}{L}$ such supports

Solve the LS problem $\min_{\underline{\alpha}} \|\mathbf{D}\underline{\alpha} - \underline{y}\|_2^2$ s.t. $\text{supp}(\underline{\alpha}) = S_i$ for each support

LS error $\le \varepsilon^2$?

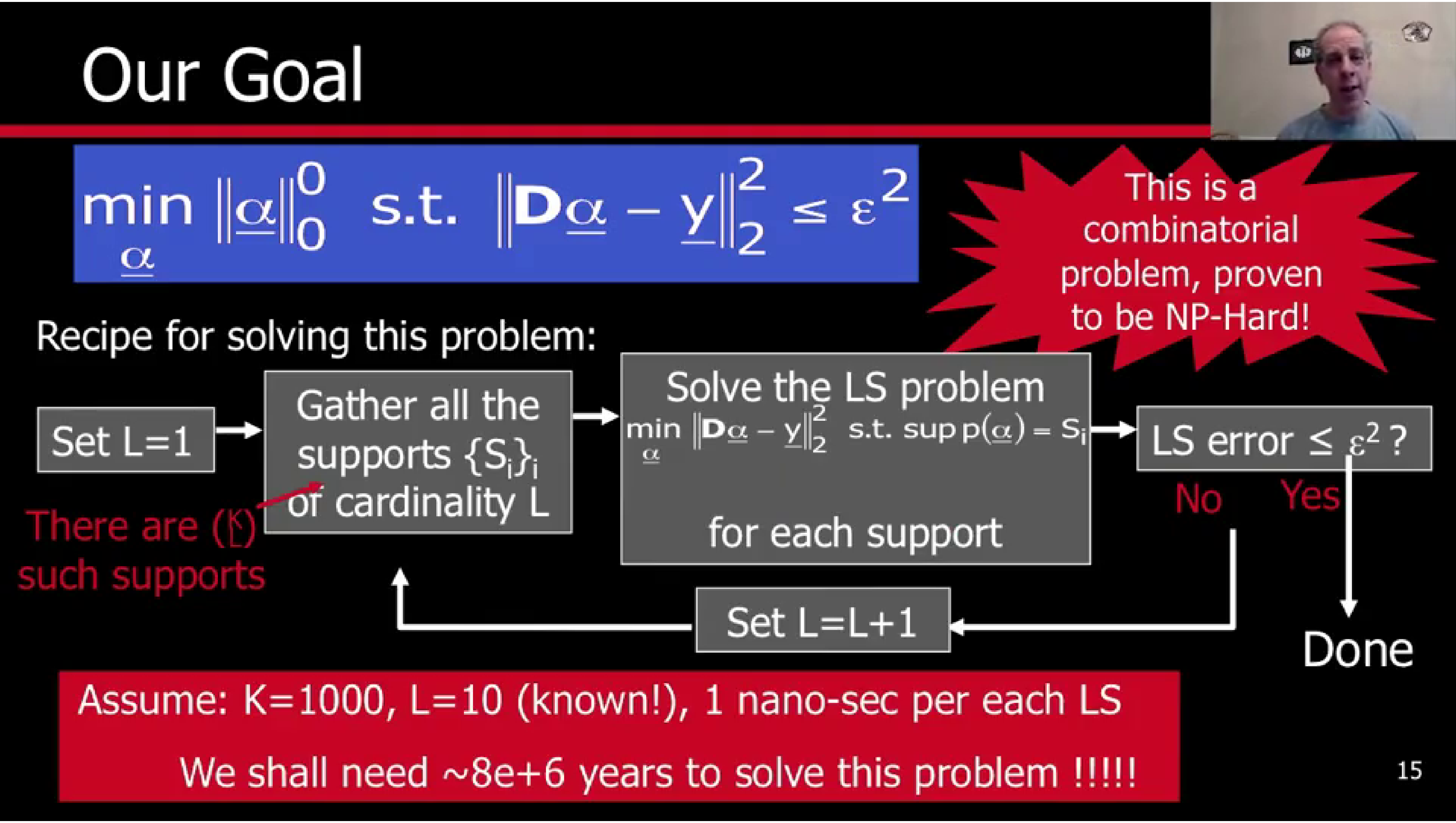
No

Yes

Set L=L+1

Done

Assume: K=1000, L=10 (known!), 1 nano-sec per each LS



Our Goal

$$\min_{\underline{\alpha}} \left\|\underline{\alpha}\right\|_0^0 \text{ s.t. } \left\|\mathbf{D}\underline{\alpha} - \underline{y}\right\|_2^2 \le \varepsilon^2$$

This is a combinatorial problem, proven to be NP-Hard!

Recipe for solving this problem:

Set L=1 → Gather all the supports $\{S_i\}_i$ of cardinality L → Solve the LS problem $\min_{\underline{\alpha}} \left\|\mathbf{D}\underline{\alpha} - \underline{y}\right\|_2^2$ s.t. $\sup p(\underline{\alpha}) = S_i$ for each support → LS error $\le \varepsilon^2$?

Yes → Done

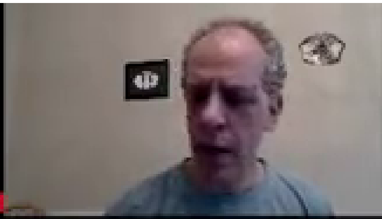
No → Set L=L+1 → Gather all the supports

There are $\binom{K}{L}$ such supports

Assume: K=1000, L=10 (known!), 1 nano-sec per each LS

We shall need ~8e+6 years to solve this problem !!!!!

Lets Approximate



$$\min_{\underline{\alpha}} \left\|\underline{\alpha}\right\|_0^0 \text{ s.t. } \left\|\mathbf{D}\underline{\alpha} - \underline{y}\right\|_2^2 \le \varepsilon^2$$

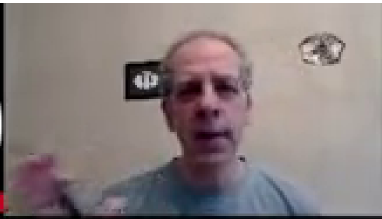
Relaxation methods

Smooth the L_0 and use continuous optimization techniques

Greedy methods

Build the solution one non-zero element at a time

Relaxation – The Basis Pursuit (BP)

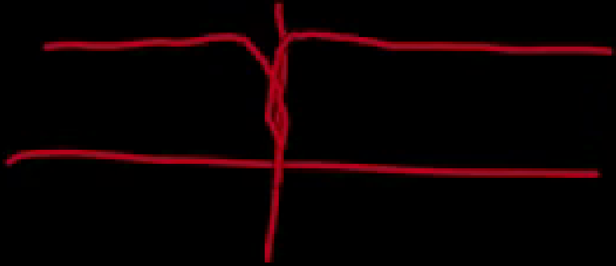


Instead of solving

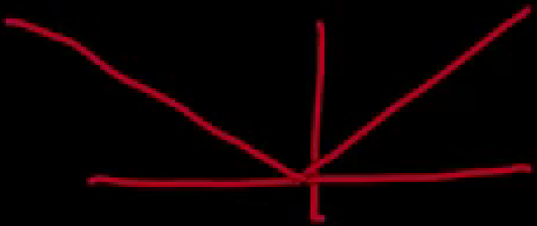
$$\min_{\underline{\alpha}} \|\underline{\alpha}\|_0^0 \text{ s.t. } \|\mathbf{D}\underline{\alpha} - \underline{y}\|_2 \le \varepsilon$$

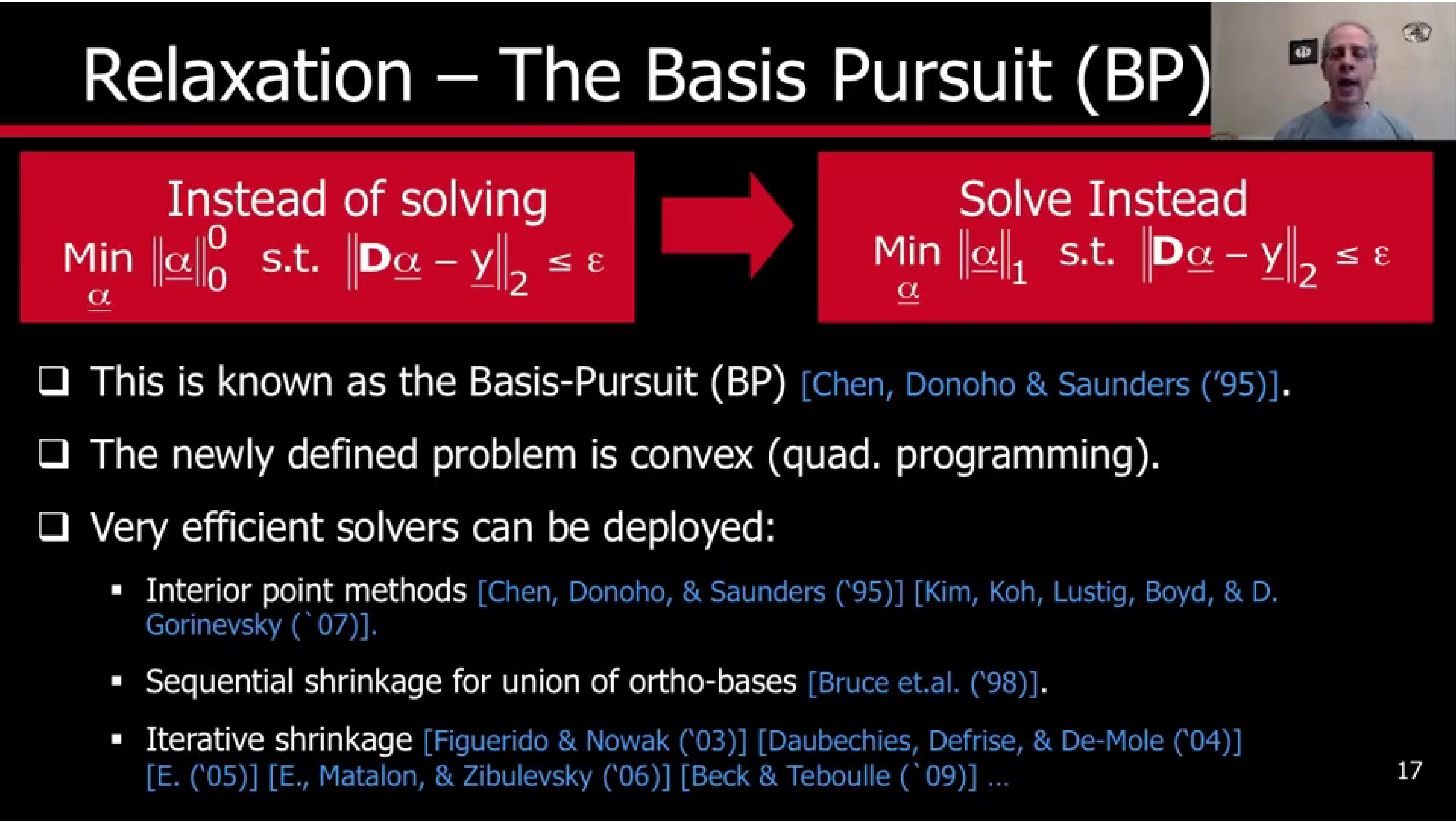
Solve Instead

$$\min_{\underline{\alpha}} \|\underline{\alpha}\|_1 \text{ s.t. } \|\mathbf{D}\underline{\alpha} - \underline{y}\|_2 \le \varepsilon$$



D, L





Relaxation – The Basis Pursuit (BP)

Instead of solving

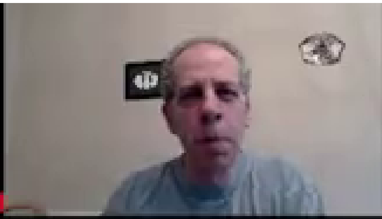
$$\min_{\underline{\alpha}} \left\|\underline{\alpha}\right\|_0^0 \ \text{ s.t. } \ \left\|\mathbf{D}\underline{\alpha}-\underline{y}\right\|_2 \le \varepsilon$$

Solve Instead

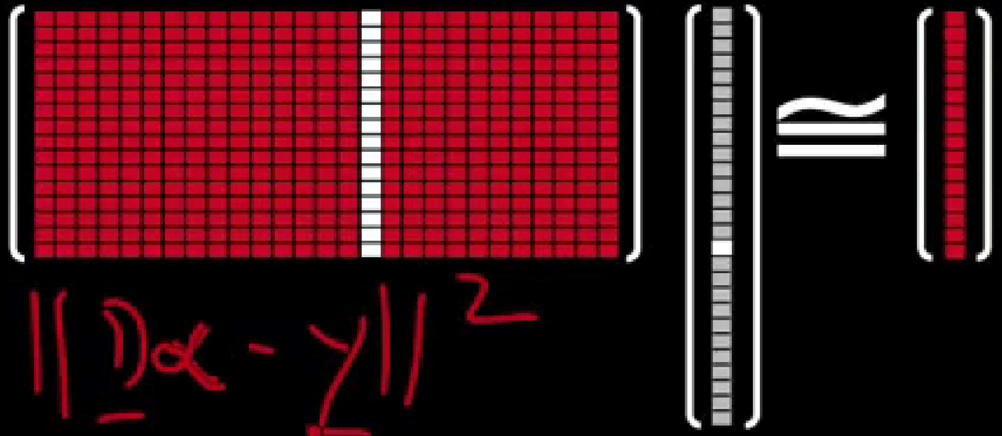
$$\min_{\underline{\alpha}} \left\|\underline{\alpha}\right\|_1 \ \text{ s.t. } \ \left\|\mathbf{D}\underline{\alpha}-\underline{y}\right\|_2 \le \varepsilon$$

- This is known as the Basis-Pursuit (BP) [Chen, Donoho & Saunders ('95)].
- The newly defined problem is convex (quad. programming).
- Very efficient solvers can be deployed:
    - Interior point methods [Chen, Donoho, & Saunders ('95)] [Kim, Koh, Lustig, Boyd, & D. Gorinevsky (`07)].
    - Sequential shrinkage for union of ortho-bases [Bruce et.al. ('98)].
    - Iterative shrinkage [Figuerido & Nowak ('03)] [Daubechies, Defrise, & De-Mole ('04)] [E. ('05)] [E., Matalon, & Zibulevsky ('06)] [Beck & Teboulle (`09)] …

Go Greedy: Matching Pursuit (MP)

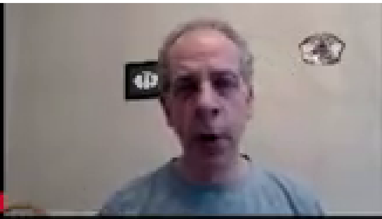


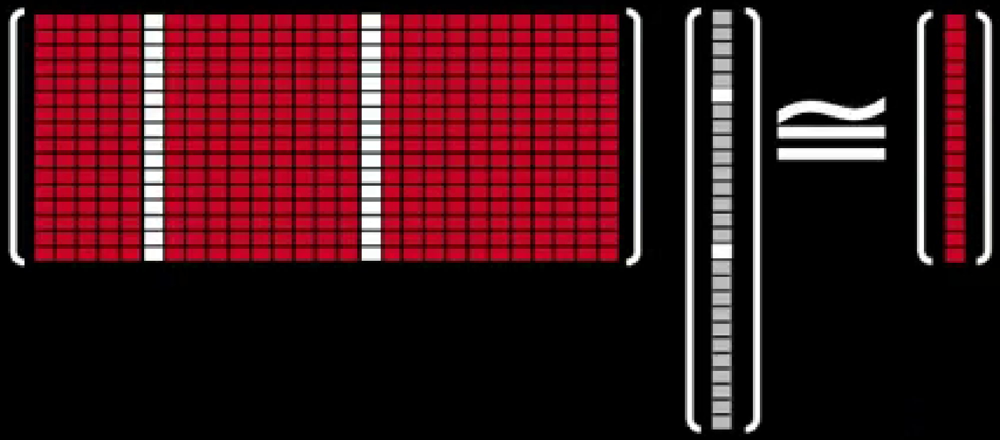
- The MP is one of the greedy algorithms that finds one atom at a time [Mallat & Zhang ('93)].
- Step 1: find the one atom that best matches the signal.



$\|\underline{D}\alpha - \underline{y}\|^2$

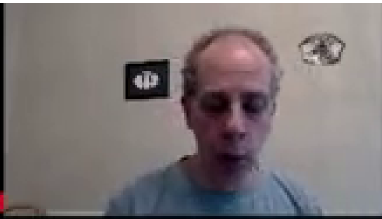
Go Greedy: Matching Pursuit (MP)





- The MP is one of the greedy algorithms that finds one atom at a time [Mallat & Zhang ('93)].
- Step 1: find the one atom that best matches the signal.
- Next steps: given the previously found atoms, find the next **one** to best fit the rsidual.
- The algorithm stops when the error $\|\mathbf{D}\underline{\alpha} - \underline{y}\|_2$ is below the destination threshold.
- The Orthogonal MP (OMP) is an improved version that re-evaluates the coefficients by Least-Squares after each round.

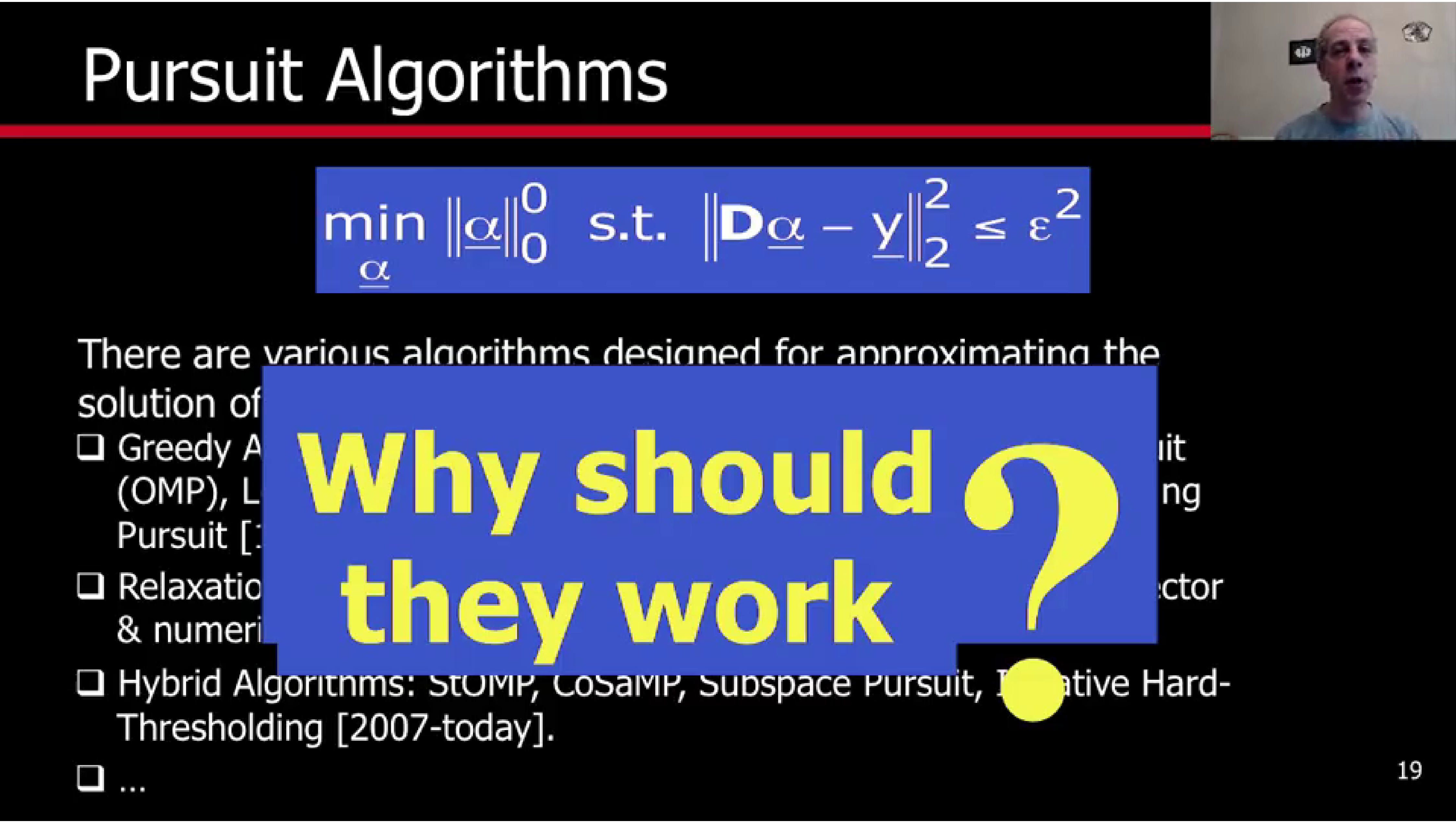
Pursuit Algorithms



$$\min_{\underline{\alpha}} \|\underline{\alpha}\|_0^0 \text{ s.t. } \|\mathbf{D}\underline{\alpha} - \underline{y}\|_2^2 \leq \varepsilon^2$$

There are various algorithms designed for approximating the solution of this problem:

- Greedy Algorithms: Matching Pursuit, Orthogonal Matching Pursuit (OMP), Least-Squares-OMP, Weak Matching Pursuit, Block Matching Pursuit [1993-today].
- Relaxation Algorithms: Basis Pursuit (a.k.a. LASSO), Dnatzig Selector & numerical ways to handle them [1995-today].
- Hybrid Algorithms: StOMP, CoSaMP, Subspace Pursuit, Iterative Hard-Thresholding [2007-today].
- …



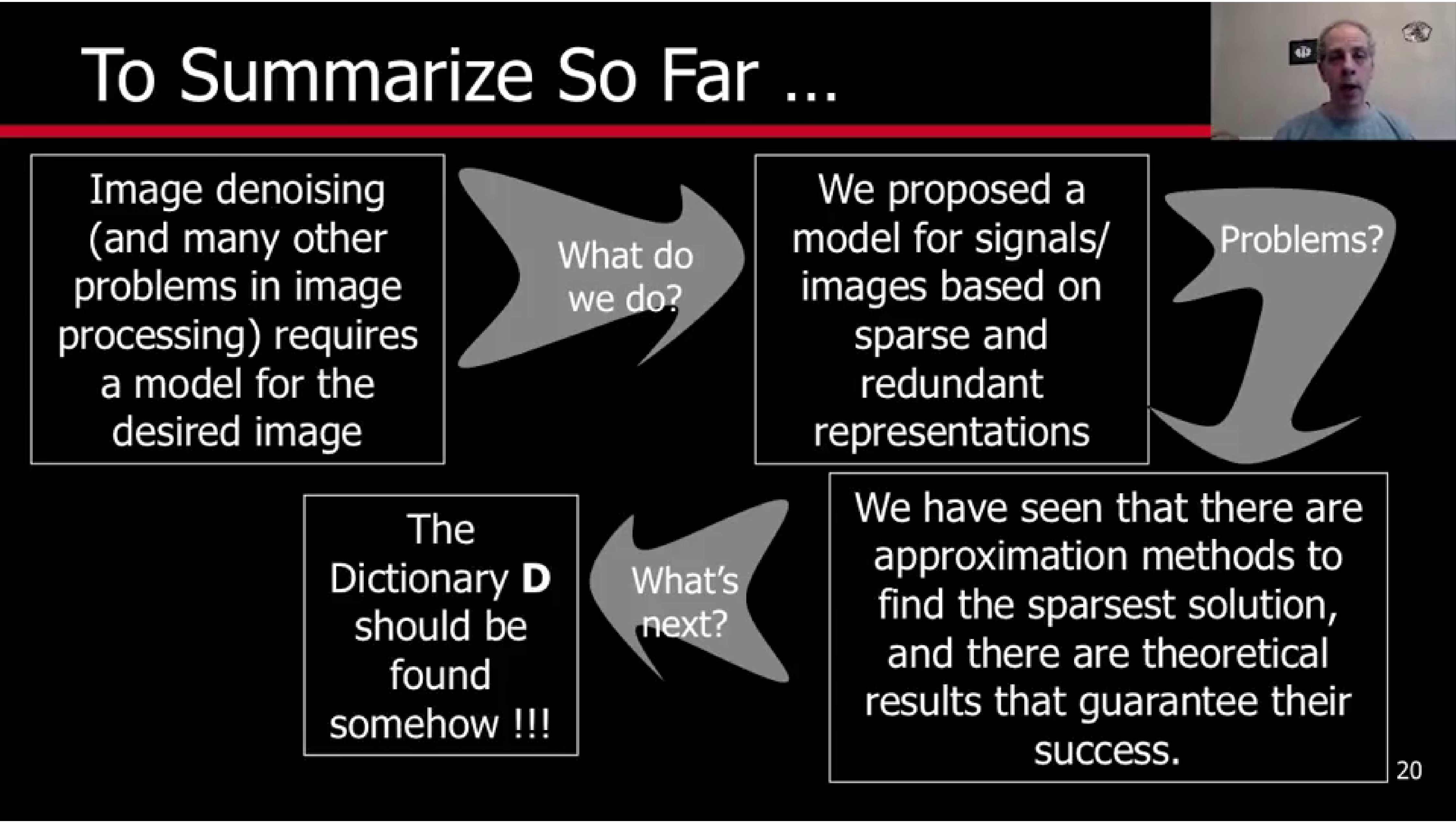
Pursuit Algorithms

$$\min_{\underline{\alpha}} \left\|\underline{\alpha}\right\|_0^0 \text{ s.t. } \left\|\mathbf{D}\underline{\alpha} - \underline{y}\right\|_2^2 \le \varepsilon^2$$

There are various algorithms designed for approximating the solution of

- ☐ Greedy A … iit (OMP), L … ng Pursuit [
- ☐ Relaxatio … ector & numeri
- ☐ Hybrid Algorithms: StOMP, CoSaMP, Subspace Pursuit, I ative Hard-Thresholding [2007-today].
- ☐ …

Why should they work ?



To Summarize So Far …

Image denoising (and many other problems in image processing) requires a model for the desired image

What do we do?

We proposed a model for signals/ images based on sparse and redundant representations

Problems?

We have seen that there are approximation methods to find the sparsest solution, and there are theoretical results that guarantee their success.

What's next?

The Dictionary **D** should be found somehow !!!