Feature Restoration & Distortion Metrics

Ventsislav Chonev & Andrew Ker

ventsislav.chonev@oriel.ox.ac.uk adk@comlab.ox.ac.uk

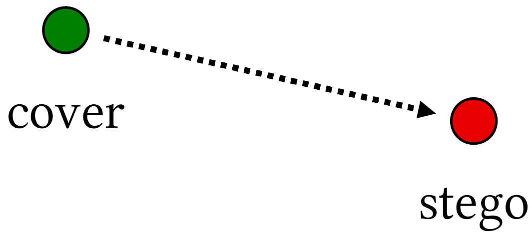
Oxford University Computing Laboratory

SPIE/IS&T Electronic Imaging, San Francisco

24 January 2011

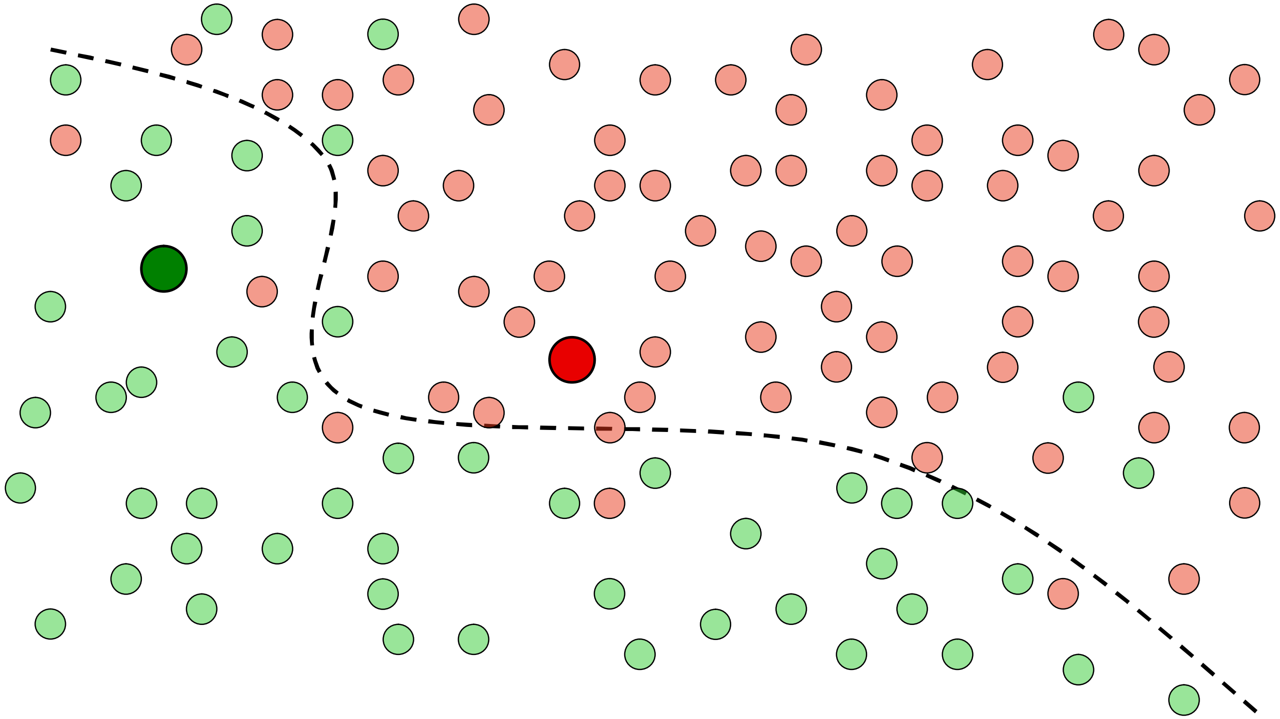
Feature restoration

- Consider some fixed set of statistical ‘features’.



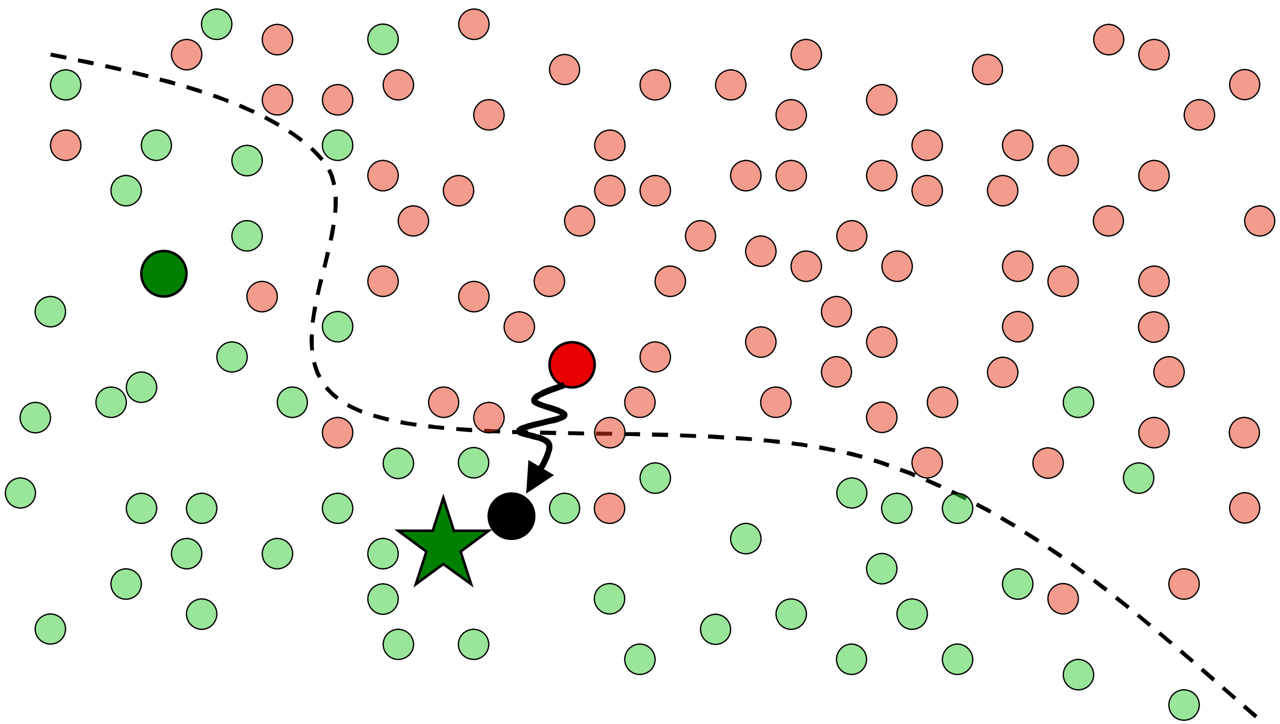
Feature restoration

- Consider some fixed set of statistical ‘features’.



Feature restoration

- Consider some fixed set of statistical 'features'.
- Use spare pixels/locations to try to restore the features to a 'target'.



1. **Which target?**
2. **How to measure distance?**
3. **How to achieve the best outcome?**

Feature restoration

- Consider some fixed set of statistical 'features'.
- Use spare pixels/locations to try to restore the features to a 'target'.

Idea first proposed in [Kodovský, 2008].

Not the same as:

Source coding/matrix embedding [Crandall, ...]
– reduce the number of embedding changes.

Distortion minimization [Kim, Filler, ...]
– during embedding, choose favourable changes.

Formalization

$$\text{minimize} \quad d\Big(\boldsymbol{t}, \phi\big(S + \textstyle\sum_{c \in \mathcal{C}} c\big)\Big) \quad \text{subject to} \quad \mathcal{C} \in \mathcal{A}$$

$d(-,-)$	is a distance metric
$\boldsymbol{t}$	is a target feature vector
ϕ	is the feature map
S	is the stego object
$S + c$	is the application of a ‘change’
$\mathcal{A}$	is a set of allowable combinations of changes

Formalization

$$\text{minimize} \quad d\Big(\boldsymbol{t}, \phi(S + \textstyle\sum_{c \in \mathcal{C}} c)\Big) \quad \text{subject to} \quad \mathcal{C} \in \mathcal{A}$$

Possible targets

- *Original cover image*
- *Estimated mean cover feature ('least suspicious')*

Quadratic form distance

$$d(\boldsymbol{u}, \boldsymbol{v}) = (\boldsymbol{u} - \boldsymbol{v})^T \Delta (\boldsymbol{u} - \boldsymbol{v})$$

- $\boldsymbol{\Delta} = \mathbf{I}$ *Euclidean distance*
- $\boldsymbol{\Delta} = \hat{\boldsymbol{\Sigma}}^{-1}$ $\boldsymbol{\Sigma}$ *the covariance of cover features: Mahalanobis distance*
- $\boldsymbol{\Delta} = \boldsymbol{\Lambda}^{-1}$ $\boldsymbol{\Lambda}$ *the diagonal of* $\boldsymbol{\Sigma}$*: standardized Euclidean distance*

Additivity

$$\phi(S + \textstyle\sum_{c \in \mathcal{C}} c) \;\approx\; \phi(S) + \sum_{c \in \mathcal{C}} \phi(S + c) - \phi(S)$$

Simplest formalization

Take an easy version of the feature restoration problem:

- no disallowed combinations of changes,
- all changes exactly additive.

This reduces to

$$\text{minimize} \quad \boldsymbol{x}^T \Theta \boldsymbol{x} + \boldsymbol{k}^T \boldsymbol{x} \quad \text{subject to} \quad \boldsymbol{x} \in \{0,1\}^n,$$

<u>Theorem</u> The above problem is NP-complete.

We will have to try iterative heuristics to approximate a solution.

Heuristics

Greedy

Test every pixel change and immediately apply all that are beneficial (reduce distance to the target).

Biased greedy

Try pixel changes in noisy regions first.

Random

Test random **batches** of pixel changes, and apply whenever beneficial.

*Genetic**

Maintain a population of pixel changes and 'evolve' the best.

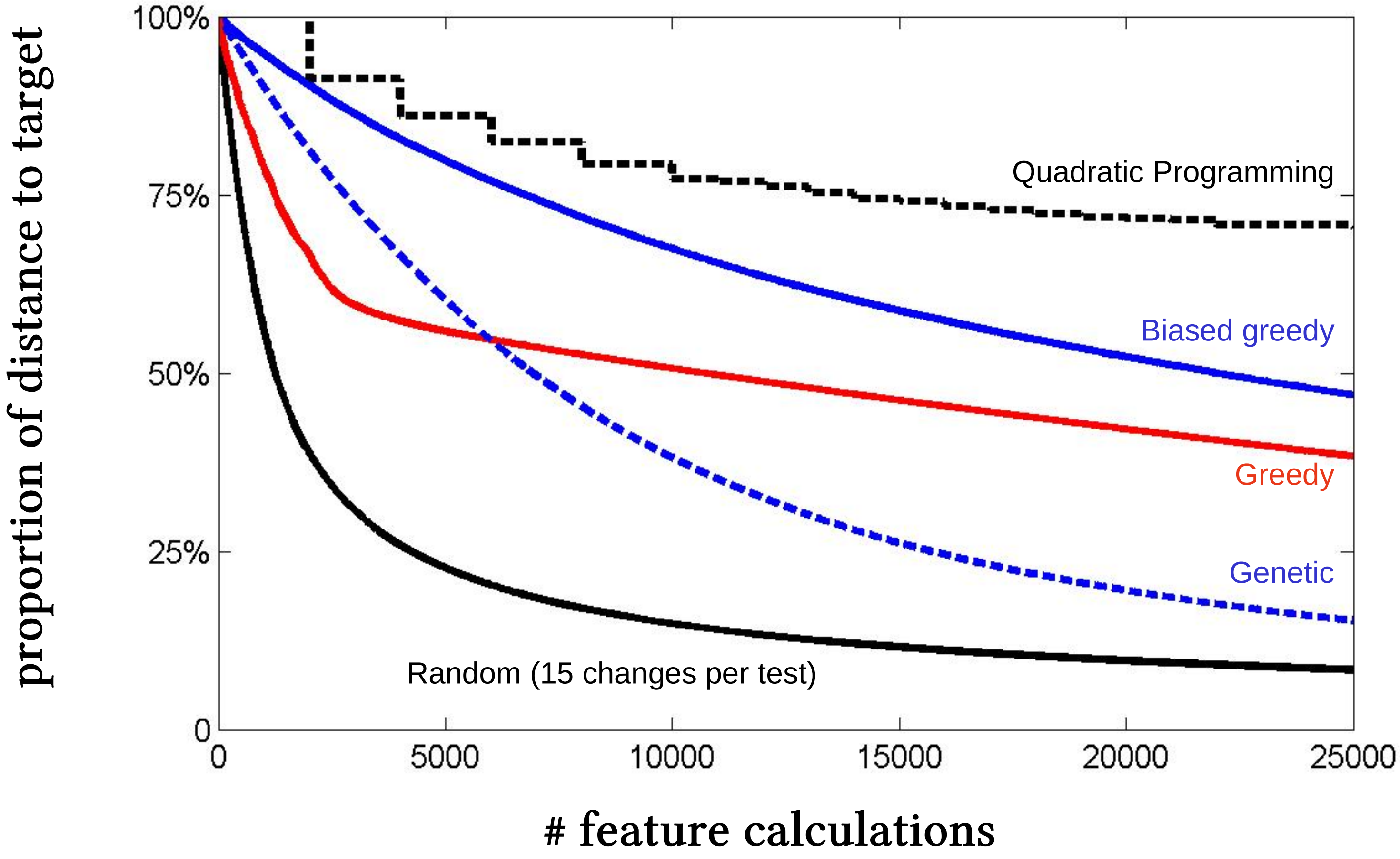
*Quadratic Programming**

Approximately solve the NP-complete problem by relaxing the integrality constraint.

*** assumes some form of additivity.**

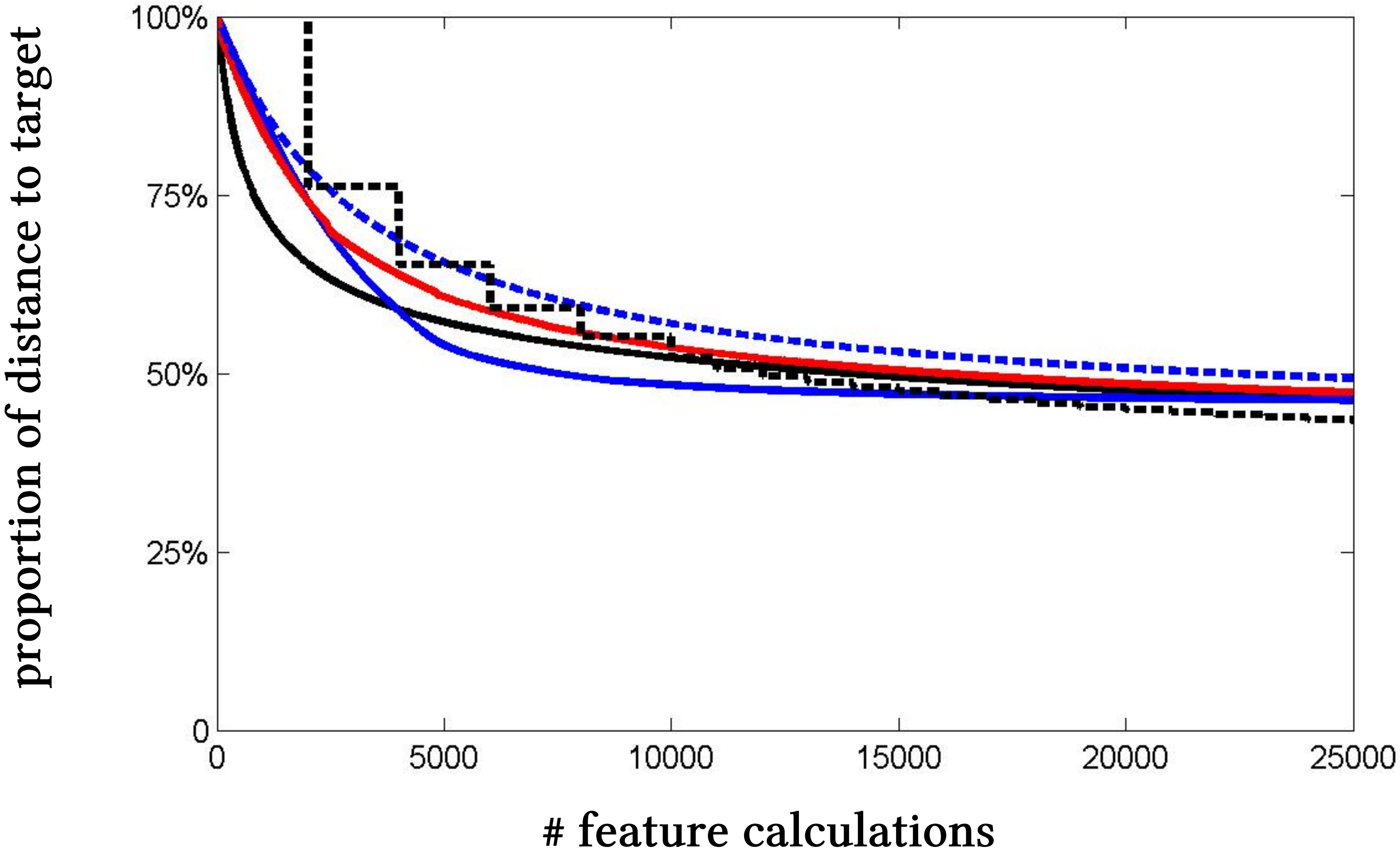
Performance

- WAM features (27), 2000 grayscale images, LSB matching 0.5bpp
- Distortion metric: **Euclidean distance** to cover



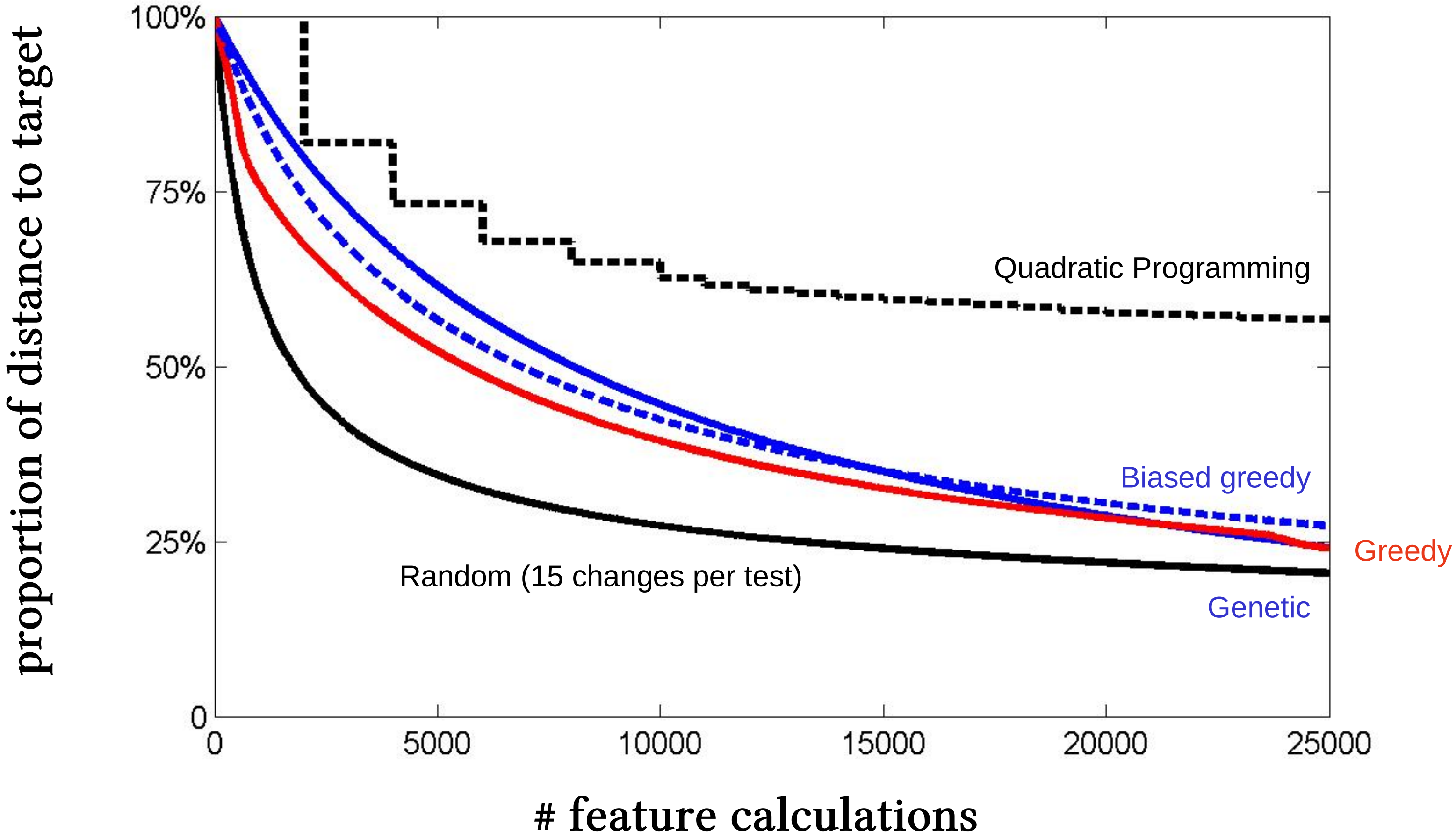
Performance

- WAM features (27), 2000 grayscale images, LSB matching 0.99bpp
- Distortion metric: **Mahalanobis distance** to mean



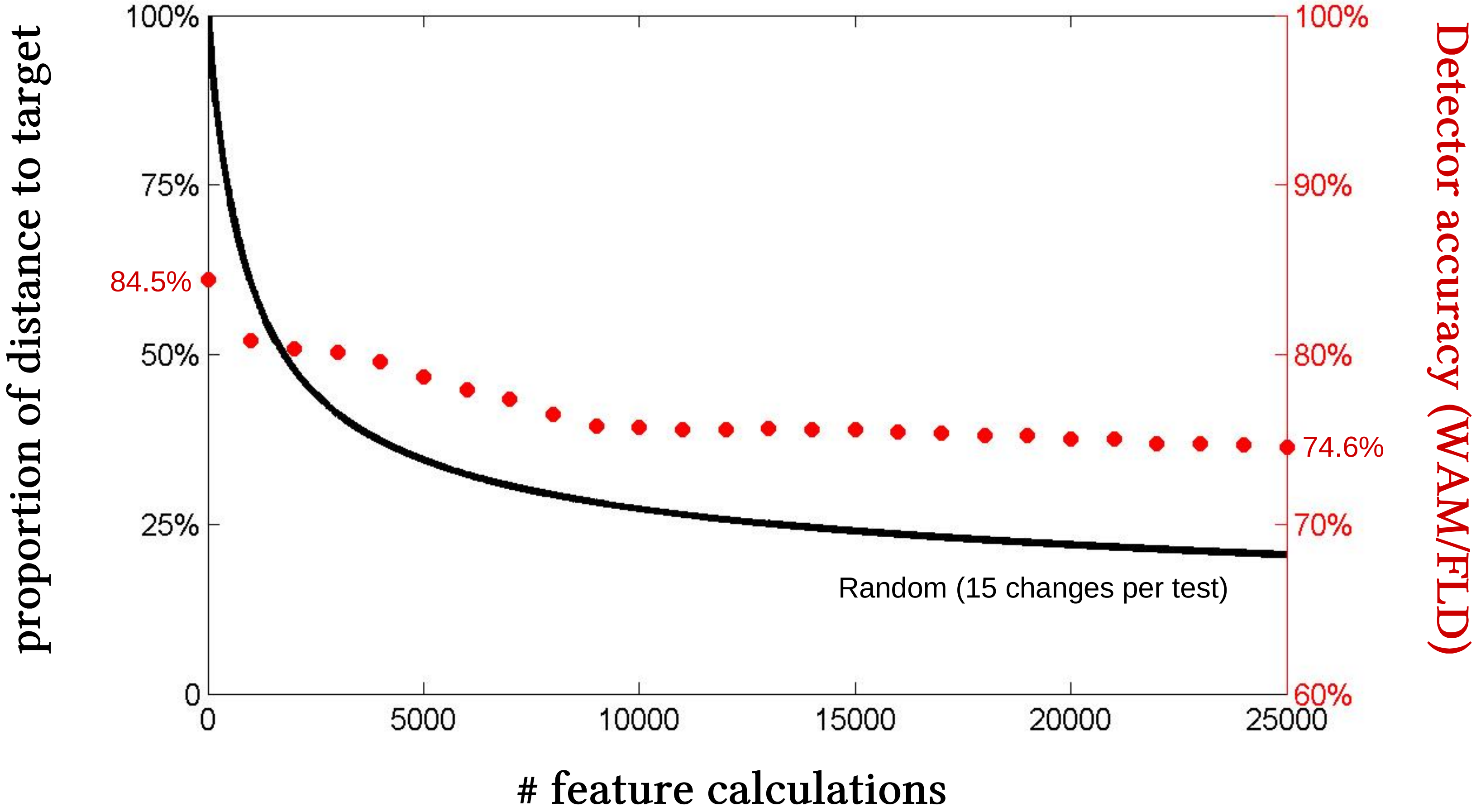
Performance

- WAM features (27), 2000 grayscale images, LSB matching 0.9bpp
- Distortion metric: **Mahalanobis distance** to cover



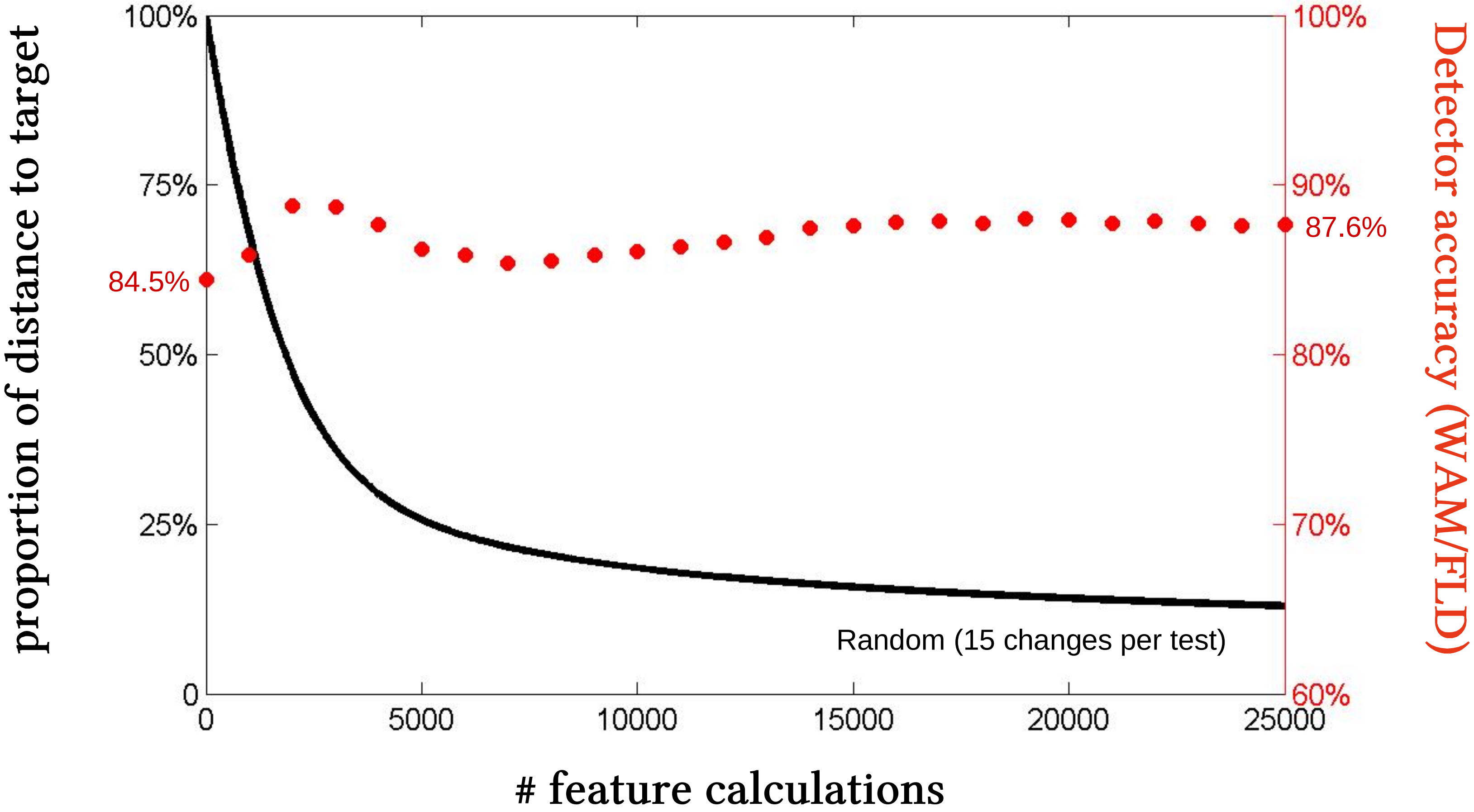
Detectability

- WAM features (27), 2000 grayscale images, LSB matching 0.9bpp
- Distortion metric: **Mahalanobis distance** to cover



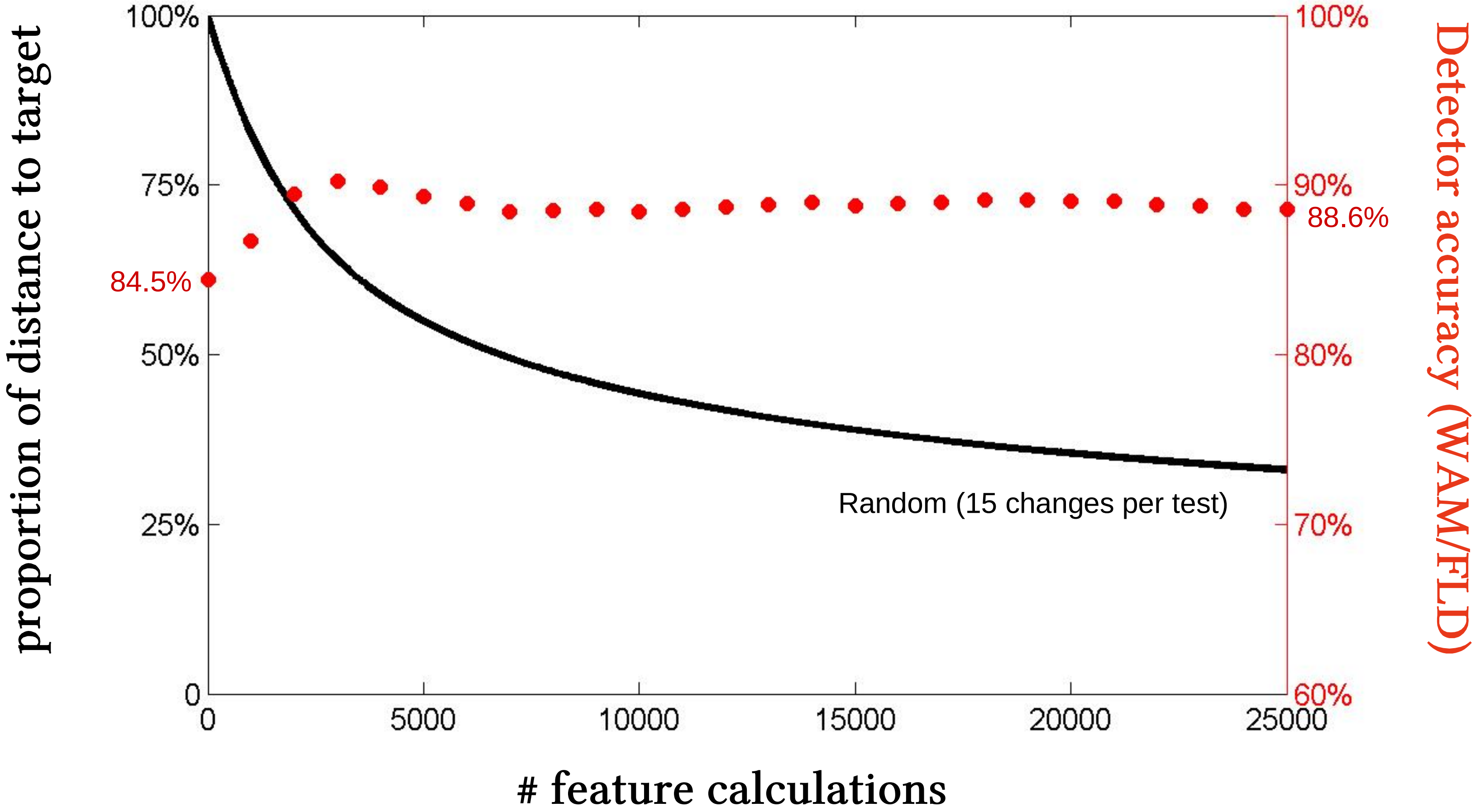
Detectability

- WAM features (27), 2000 grayscale images, LSB matching 0.9bpp
- Distortion metric: **Euclidean distance** to cover



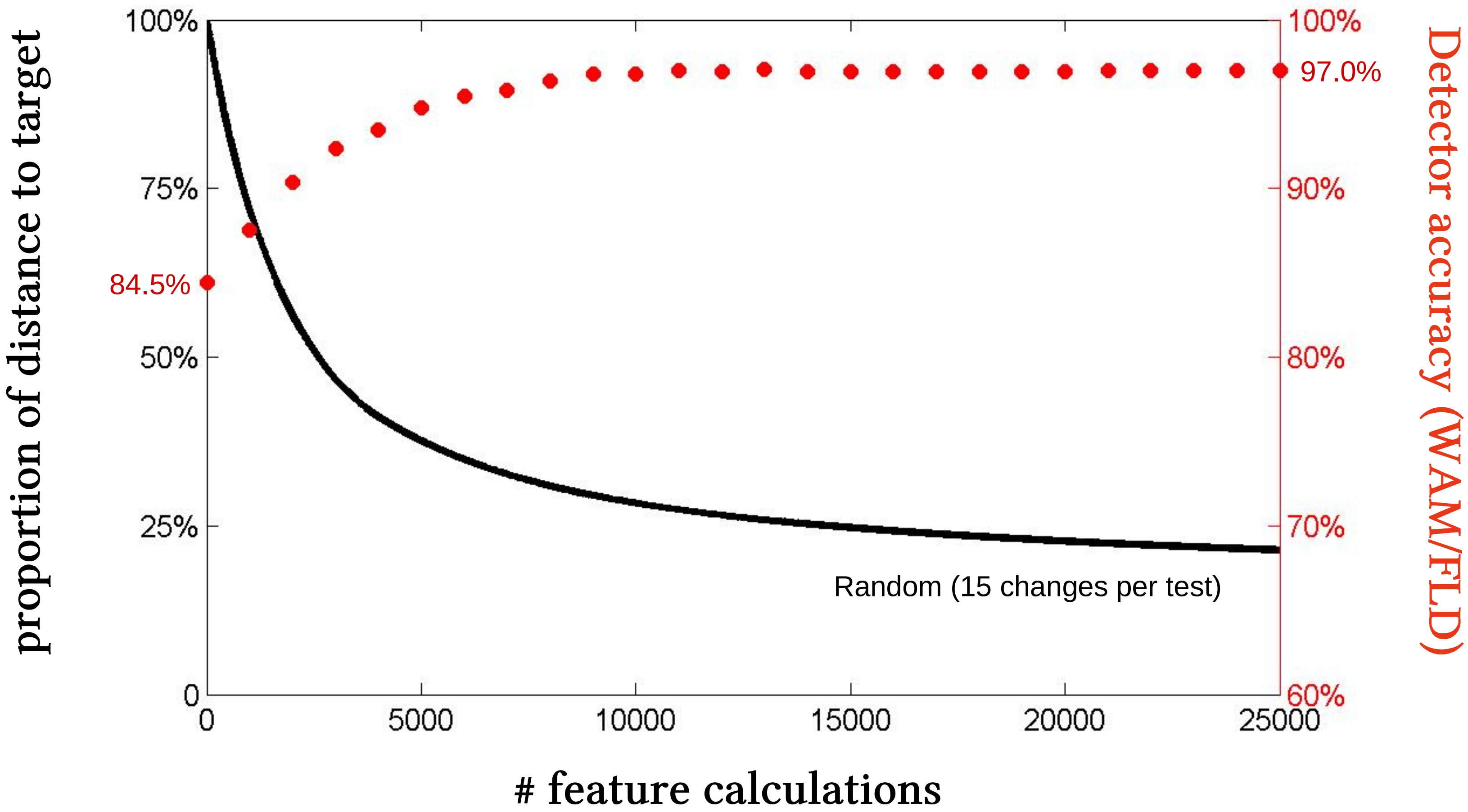
Detectability

- WAM features (27), 2000 grayscale images, LSB matching 0.9bpp
- Distortion metric: **standardized Euclidean distance** to cover

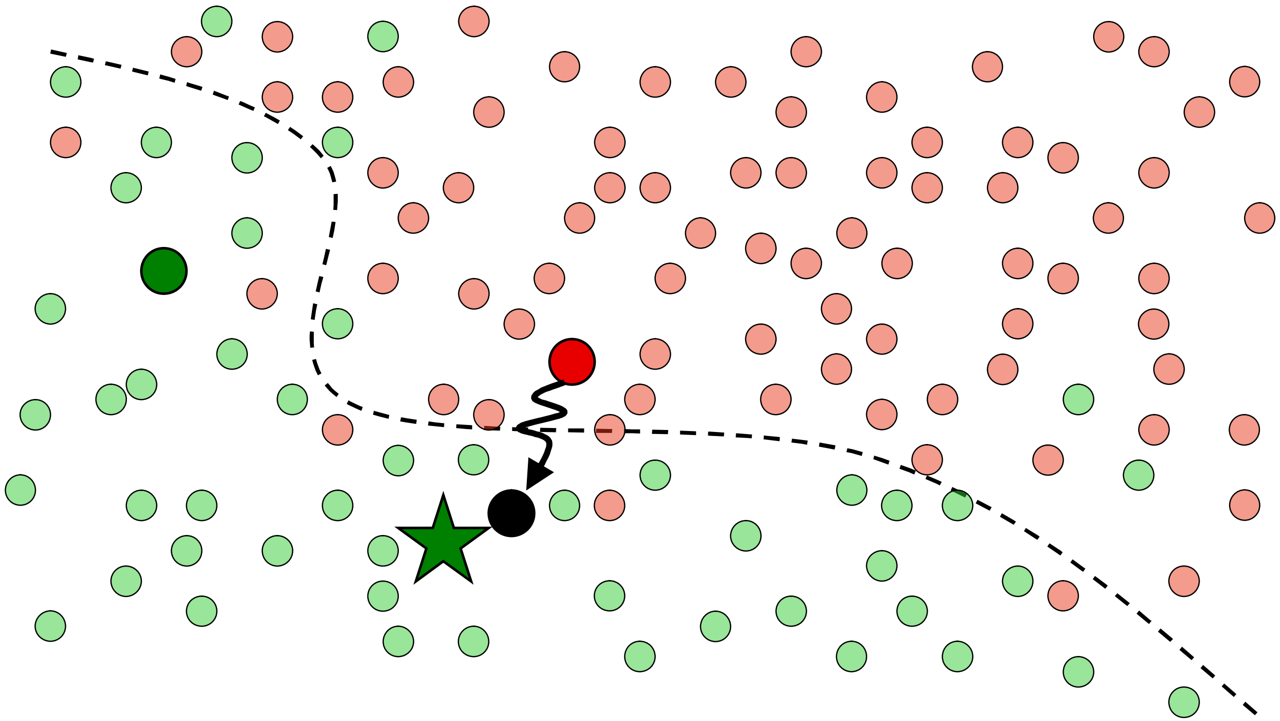


Detectability

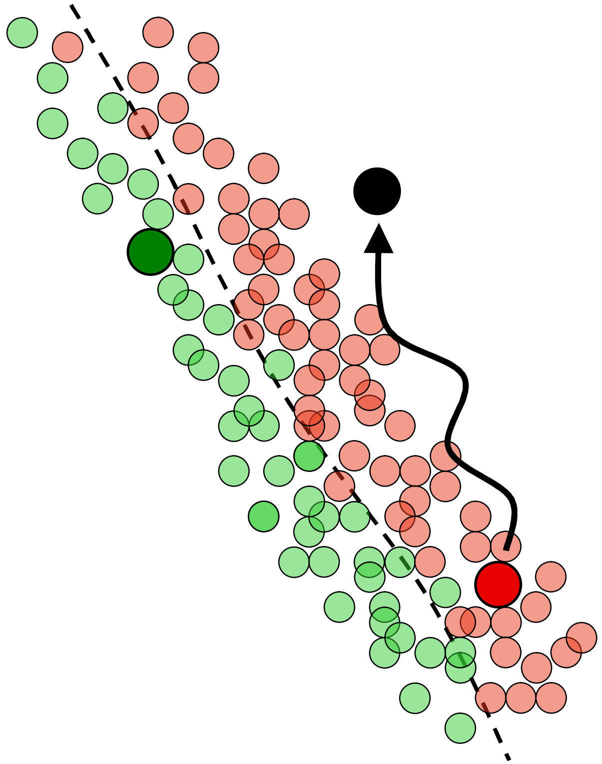
- WAM features (27), 2000 grayscale images, LSB matching 0.9bpp
- Distortion metric: **Euclidean distance** to mean



The real picture



The real picture



With WAM features, correlation so strong that even computing Σ^{-1} can be numerically unstable!

Lesson: check the condition number of your features' covariance matrix.

Conclusions

- Feature restoration should be used with caution.
  - *If targetting the wrong features, it might be disastrous.*
  - *But can be bolted on to other embedding methods.*
  - *Seems to work well with payloads as high as 90%-99%.*

- Randomized algorithms provide a way to approximate solutions to this NP-complete problem.

- It is **critical** to use the right distortion metric.
  - *Euclidean distance, standardized or not, seems to be poor.*
  - *Mahalanobis distance whitens the features, but could be unstable.*
  - *This lesson is important for other areas of steganography.*