The Ultimate Steganalysis Benchmark?

Andrew Ker

adk@comlab.ox.ac.uk

Royal Society University Research Fellow

Oxford University Computing Laboratory

ACM Multimedia & Security Workshop

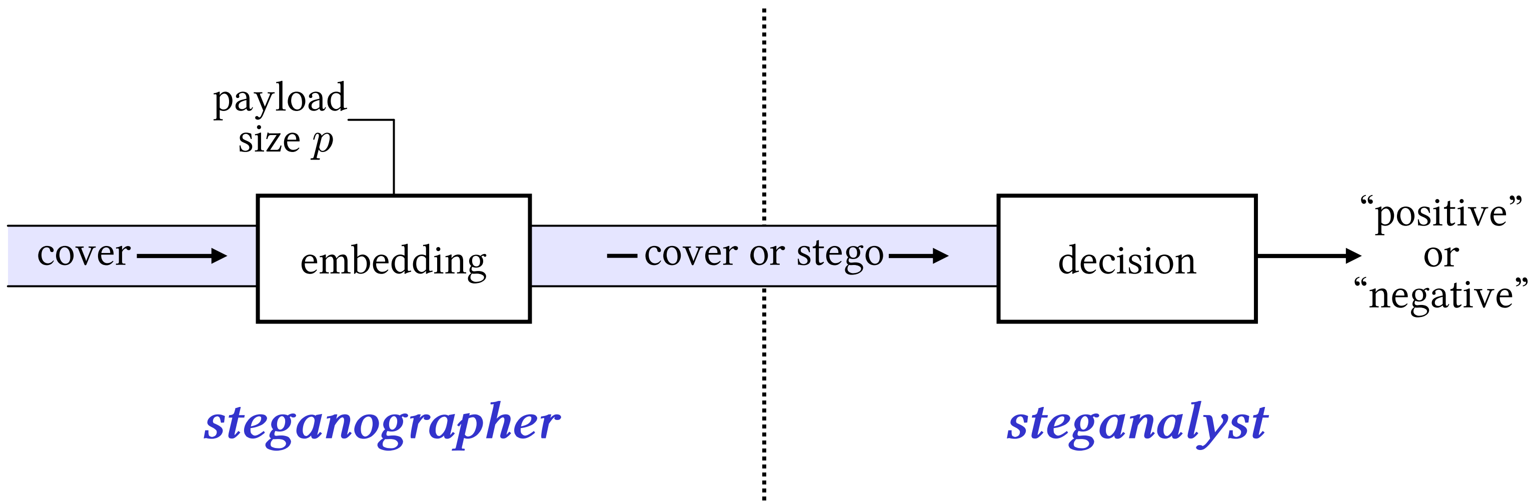
21 September 2007

The Ultimate Steganalysis Benchmark?

Outline

- *Currently-used benchmarks not ideal*
- *New benchmark based on KL divergence*
- *Difficulties estimating the benchmark value*
- *Examples*

Binary Steganalysis



Common Benchmarks

- ROC curve

  difficult to rank; too much information

- Area under ROC
- Minimize sum of false positive & negative

  assumes false positive and false negatives are equivalent

- False negative rate at fixed false positive
- False positive rate at fixed false negative

  impossible to justify numbers objectively

Common Benchmarks

- ROC curve

  difficult to rank; too much information

- Area under ROC
- Minimize sum of false positive & negative

  assumes false positive and false negatives are equivalent

- False negative rate at fixed false positive
- False positive rate at fixed false negative

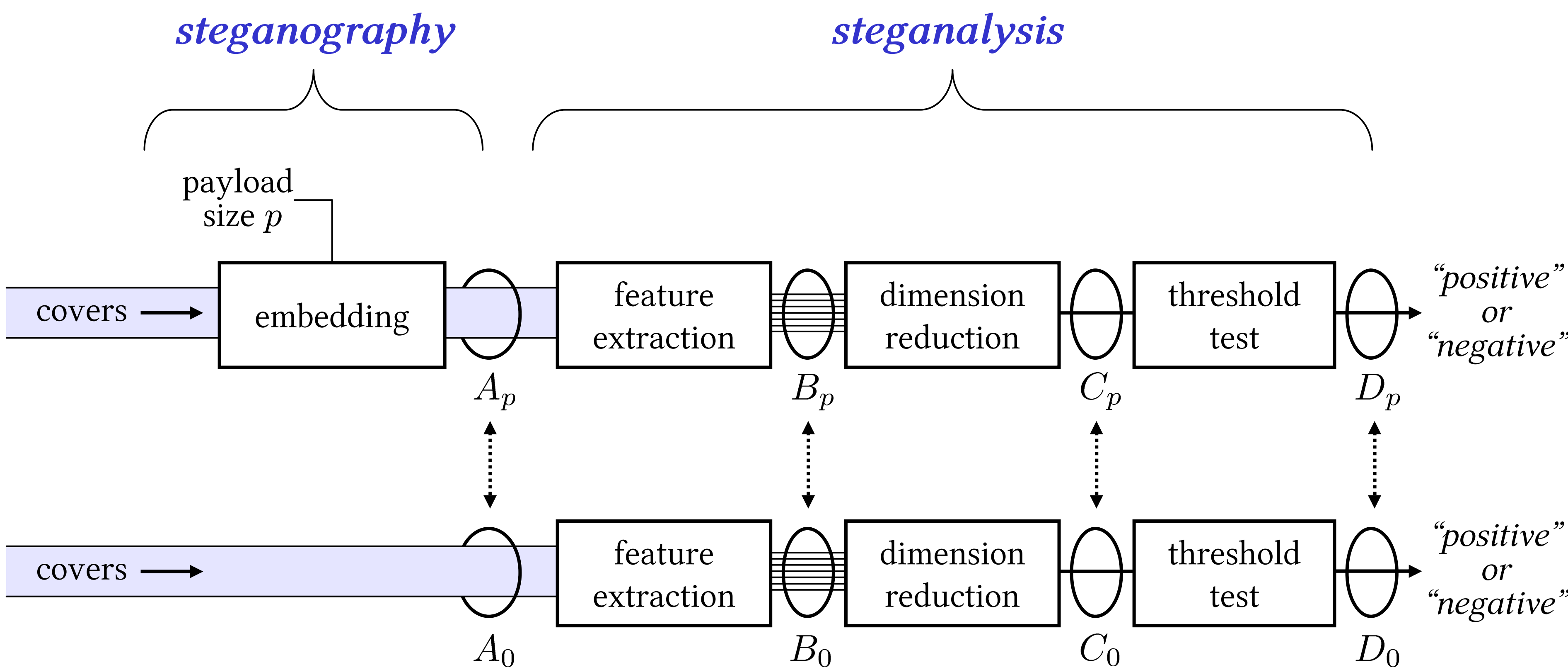
  impossible to justify numbers objectively

also depend on payload size

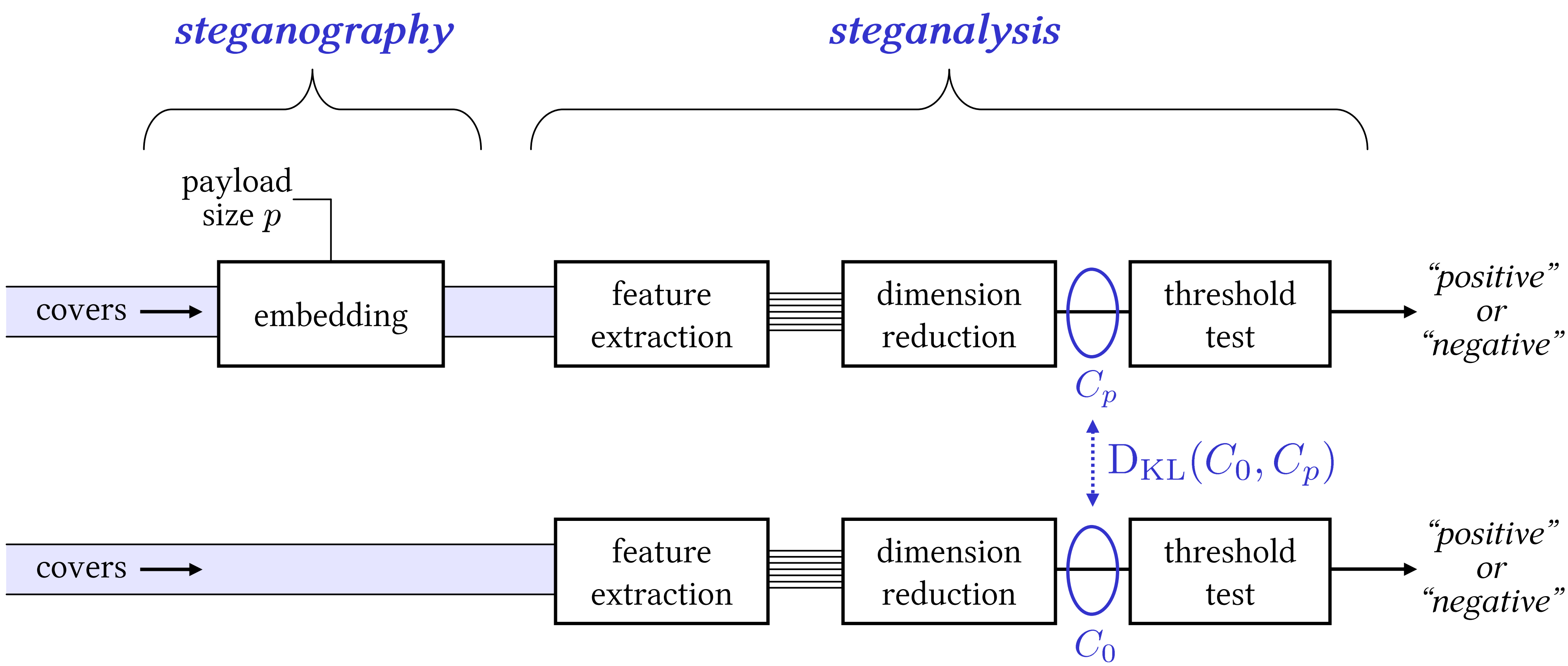
- Minimum payload detectable at fixed false positive & false negative rate

  impossible to justify numbers objectively

Distribution Differences



Distribution Differences



New Benchmark

- **Based on $\mathrm{D}_{\mathrm{KL}}(C_0, C_p)$, where C_p is the univariate distribution produced just before threshold test.**

<u>From steganalysis/info theory literature</u>

If steganography is repeated at a fixed embedding rate, the probability of detection tends to 1. [Cachin; Moulin; Ker; ...]

- **For long-run performance we should concentrate on payload sizes tending to zero.**

<u>A theorem by S. Kullback</u>

Let F_p be a family of distributions satisfying certain regularity conditions.

Then $\lim_{p\to 0} \frac{\mathrm{D}_{\mathrm{KL}}(F_0, F_p)}{p^2}$ *exists and is nonzero.* [adapted from Kullback, 1968]

- **If we believe that the regularity conditions are satisfied, then $\mathrm{D}_{\mathrm{KL}}(C_0, C_p)$ is, locally to zero, a multiple of p^2.**

New Benchmark

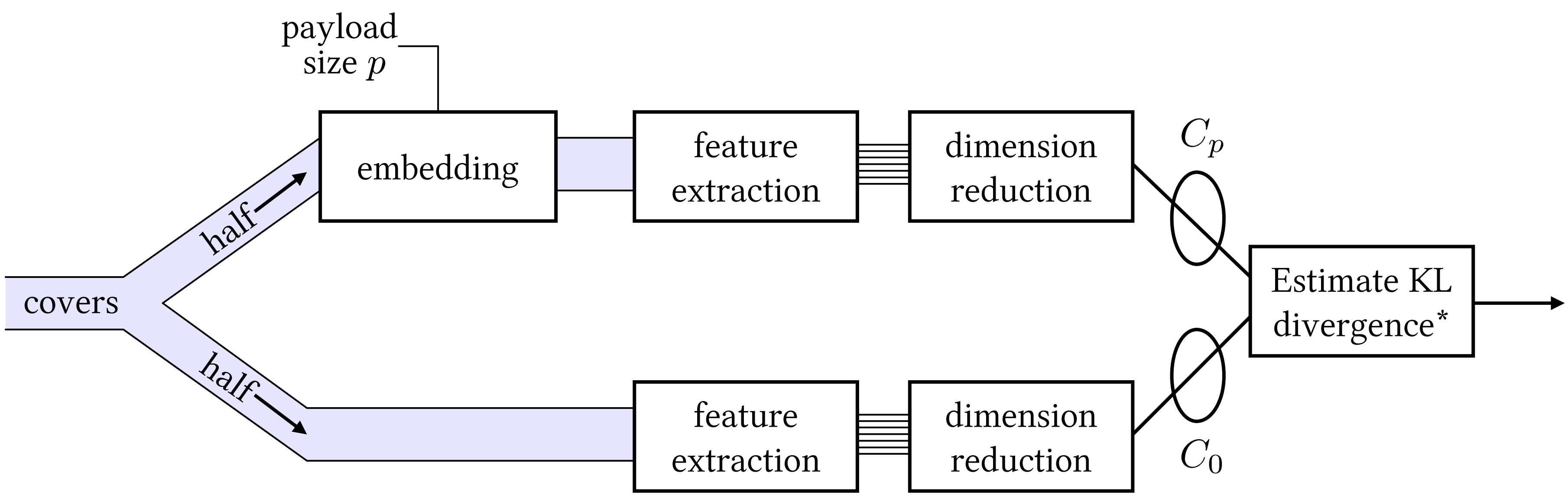
The quantity $Q = \lim_{p \to 0} \frac{\mathrm{D_{KL}}(C_0, C_p)}{p^2}$

tells us how quickly "evidence" accumulates. This is the proposed benchmark.

Note:

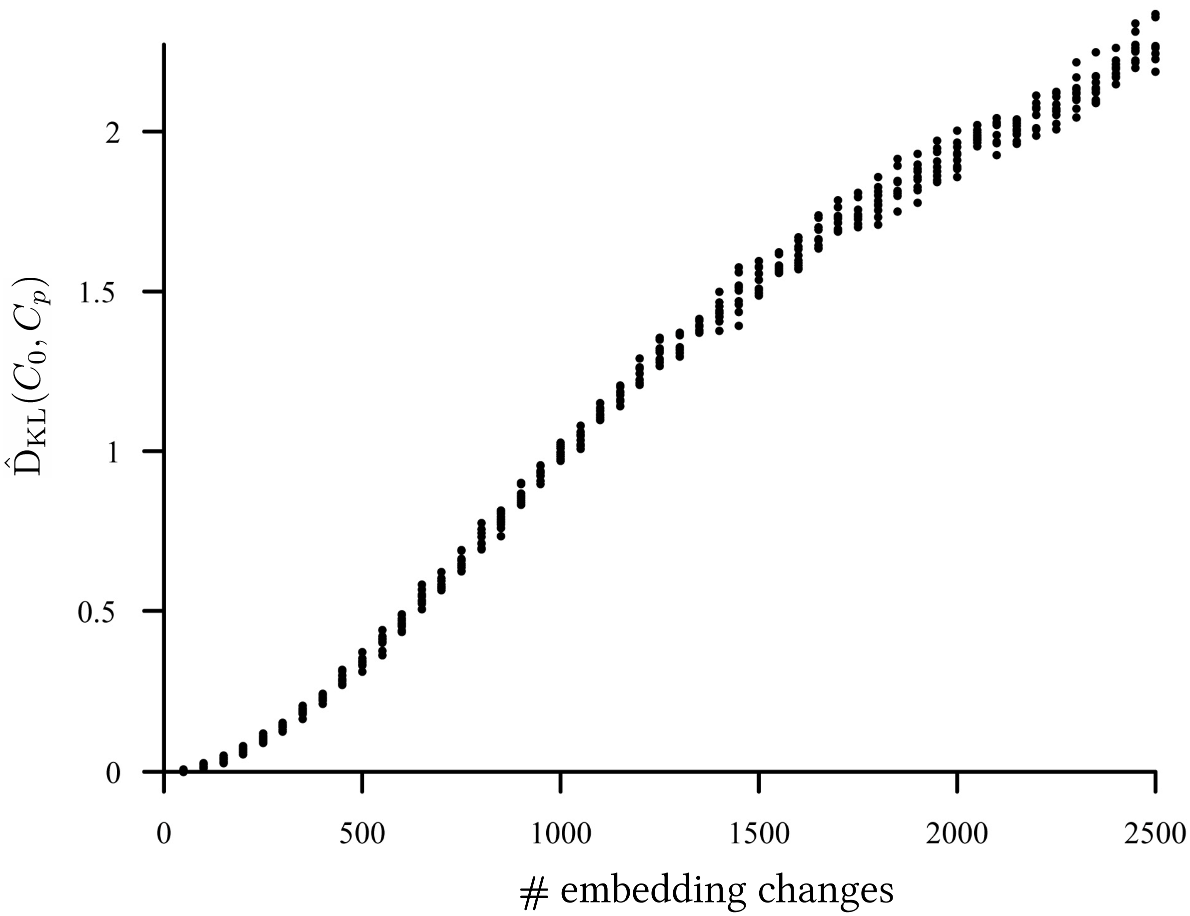
- ***"Payload size" should be measured by number of embedding changes***
- ***Then Q is measured in "nats per embedding change squared"***

Experimental Results



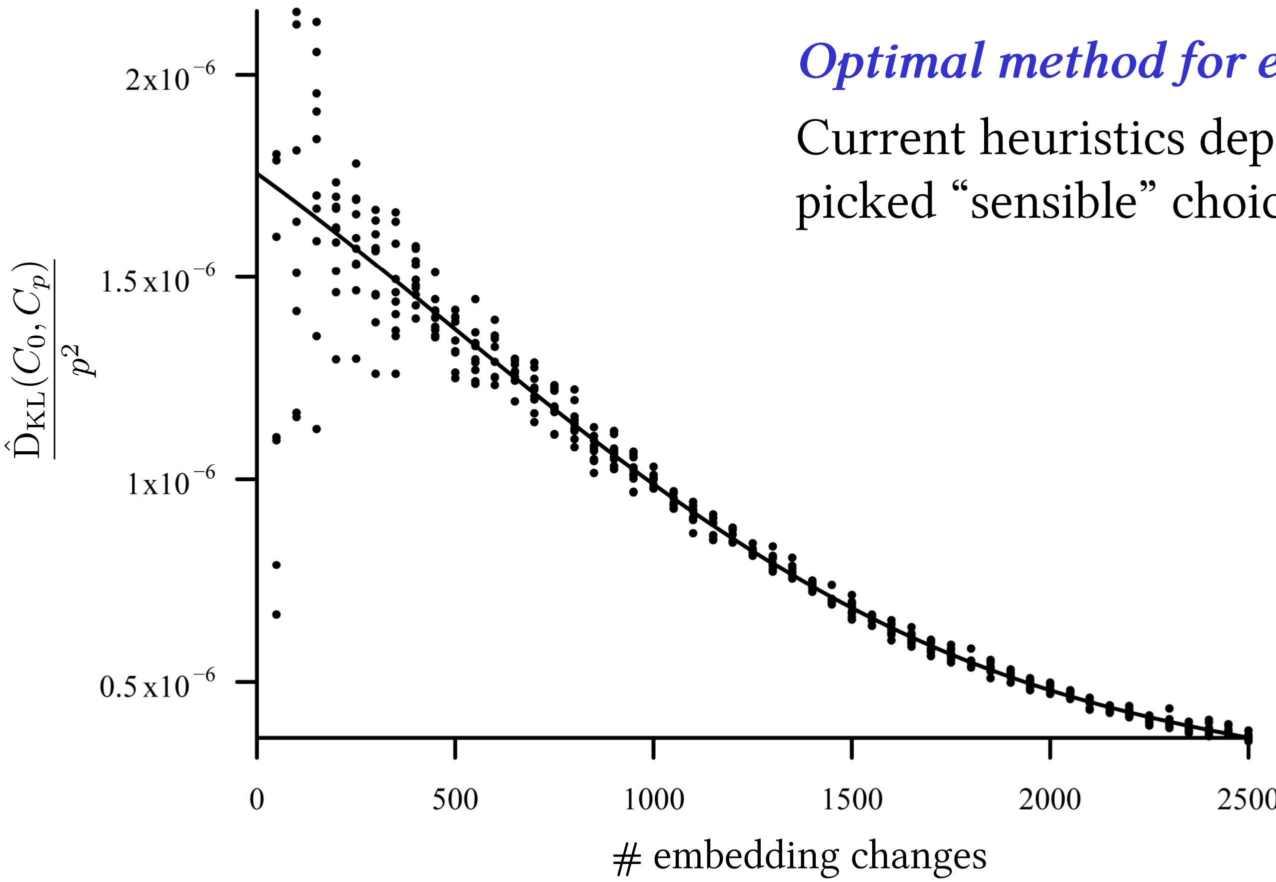
*KL divergence estimation by [Wang, Kulkarni, & Verdu, 2005]

Experimental Results



- 10000 cover images
- LSB replacement embedding, 50 payload sizes, repeated 10 times each
- “Triples” steganalysis

Experimental Results

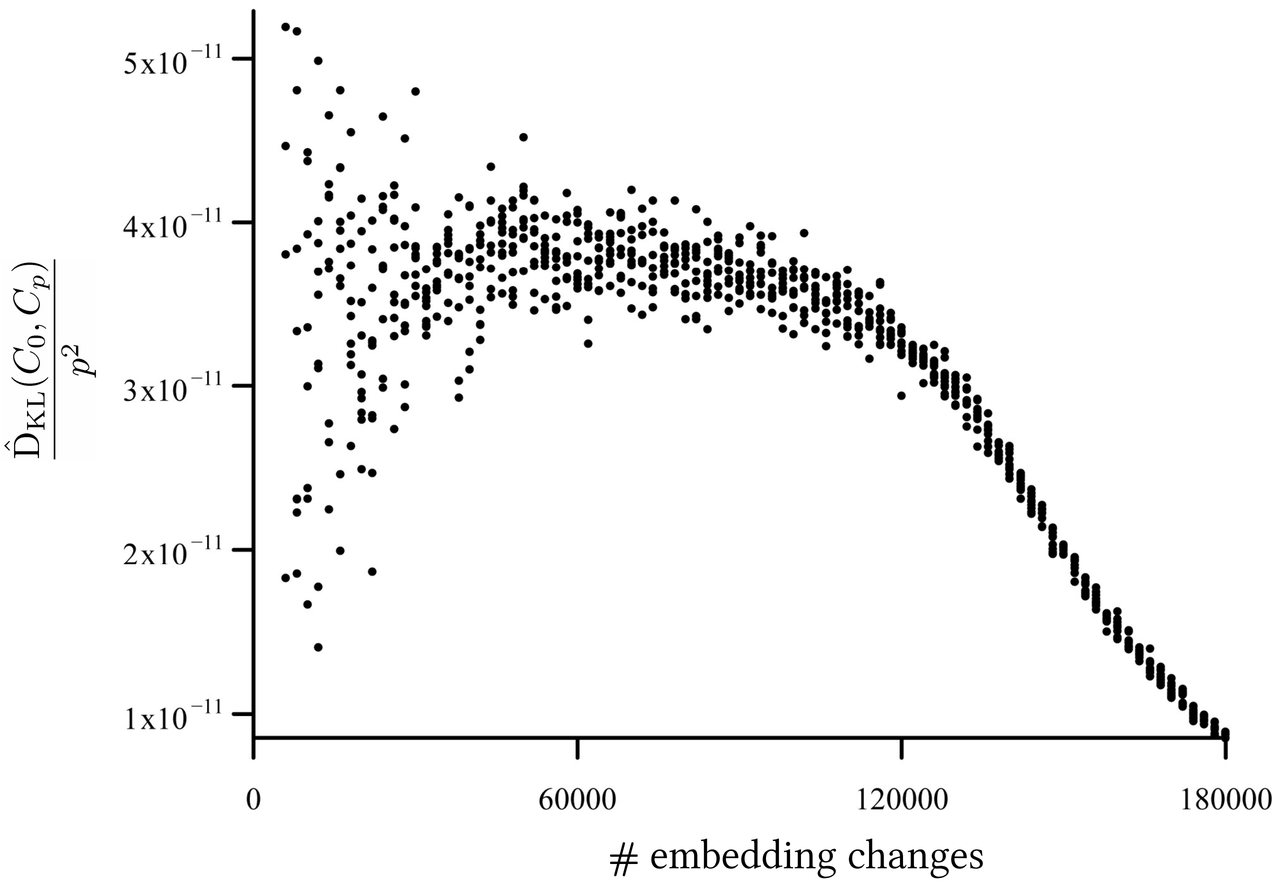


Optimal method for estimating Q?

Current heuristics depend on hand-picked "sensible" choice of p.

- 10000 cover images
- LSB replacement embedding, 50 payload sizes, repeated 10 times each
- "Triples" steganalysis

Experimental Results



- 20000 cover images
- LSB matching (±1) embedding, 90 payload sizes, repeated 10 times each
- “Calibrated HCF COM” steganalysis

Conclusions

- There is a need for an application-independent benchmark.
- The new "Q-factor" benchmark measures how quickly **information** is accumulated as payload increases.
- More work needed for good empirical estimation of "Q":
  - *Currently seems to need a very large experimental base*
  - *Test objects should be the same size*
  - *Optimal estimation?*

Conclusions

- There is a need for an application-independent benchmark.
- The new "Q-factor" benchmark measures how quickly **information** is accumulated as payload increases.
- More work needed for good empirical estimation of "Q":
  - *Currently seems to need a very large experimental base*
  - *Test objects should be the same size*
  - *Optimal estimation?*

Steganalysis	3000 grayscale bitmap covers	10000 colour JPEG covers
SPA [Dumitrescu et al, IHW 2002]	16.1	28.3
SPA/LSM [Lu et al, IHW 2004]	12.1	161
Triples [Ker, IHW 2005]	20.7	1500
Triples/WLSM [Ker, SPIE EI 2007]	16.1	1500

nanonats per embedding change squared

End

adk@comlab.ox.ac.uk