

On the Relationship Between Embedding Costs and Steganographic Capacity

Andrew D. Ker

Department of Computer Science

University of Oxford

Oxford, UK

adk@cs.ox.ac.uk

ABSTRACT

Contemporary steganography in digital media is dominated by the framework of additive distortion minimization: every possible change is given a cost, and the embedder minimizes total cost using some variant of the Syndrome-Trellis Code algorithm. One can derive the relationship between the cost of each change c_i and the probability that it should be made π_i , but the literature has not examined the relationship between the costs and the total capacity (secure payload size) of the cover. In this paper we attempt to uncover such a relationship, asymptotically, for a simple independent pixel model of covers. We consider a ‘knowing’ detector who is aware of the embedding costs, in which case $\sum \pi_i^2 c_i$ should be optimized. It is shown that the total of the inverse costs, $\sum c_i^{-1}$, along with the embedder’s desired security against an optimal opponent, determines the asymptotic capacity. This result also recovers a Square Root Law. Some simple simulations confirm the relationship between costs and capacity in this ideal model.

KEYWORDS

Adaptive steganography; optimal embedding; square root law

1 INTRODUCTION

In steganography, *additive distortion minimization* is a framework for selecting steganographic changes, aiming to choose the least detectable combination. Each possible change is enumerated – in images, this might include incrementing or decrementing each pixel or transform coefficient – and each change is assigned a *cost*. Then the combination of changes is selected which both encodes the desired payload and minimizes total cost, most famously using the Syndrome-Trellis Code (STC) algorithm [3]. With different heuristics for cost, this is the dominant method in digital media steganography at this time: the UNIWARD methods [4] and HILL [7] are the standard benchmarks for still image steganography, and STCs are now also used in audio [8] and video [10] steganography.

When the attacker (Warden, steganalyst) is aware of the costs, minimizing total cost is not optimal. Such a situation was examined in [6], which derives optimal behaviour for the the embedder against this so-called *knowing attacker*.

Against either an ignorant or a knowing attacker, there are simple calculations showing the relationship between the *cost* of a change and the *probability* of making such a change, in a single steganographic embedding. What the results do not show is the relationship between a cover’s entire *set* of costs and its *capacity* (by which we mean secure payload size), yet one would expect such a relationship to exist: if most costs are large, most locations

are difficult to hide in (and easy to detect changes of), so the capacity should be low; conversely if plenty of costs are low, the capacity should be high. This paper shows that the relationship is, asymptotically, rather simple.

We restrict our attention to a simple independent pixel model of covers, and the case of a knowing attacker: our result is an asymptotic analysis of the optimization problem derived in [6]. Note that this is *not* the most common optimization problem encountered in steganography, where the coding is typically solved by STCs, but it is the correct behaviour against a detector that knows the embedding costs. Indeed, this work is not strictly a result about steganography, rather about the asymptotic solutions to a certain optimization problem that has applications in steganography.

In Sect. 2 we will summarise the theoretical result of [6], describing the model at hand and the optimal behaviour of payload- and distortion-limited senders, as well as identifying some pathological cases to exclude. We state the result in Sect. 3, and prove it in Sect. 4. Some simple simulations are performed in Sect. 5 to confirm the result, and we discuss further directions in Sect. 6.

1.1 Asymptotic Notation

Throughout the paper we write asymptotic relationships

$$f(x) \sim g(x) \text{ as } x \rightarrow a$$

($a \in \mathbb{R}$ or $a = \pm\infty$; the limit may be one-sided) to mean

$$\frac{f(x)}{g(x)} \rightarrow 1$$

in the same limit. Recall that $f(x) \sim g(x)$ as $x \rightarrow a$ implies $(f \circ h)(x) \sim (g \circ h)(x)$ as $x \rightarrow h^{-1}(a)$, at least if h^{-1} exists and is continuous at a (or continuous approaching a from the relevant side, in the case of one-sided limits).

2 OPTIMAL EMBEDDING AGAINST A KNOWING ATTACKER

We adopt the model from [6], and we briefly summarise the definition and pertinent results here.

Consider a cover generated as n random elements (which we call pixels, but they need not be) taking values x_1, \dots, x_n in some finite alphabet Σ . We suppose that the pixels are independent, but not identically distributed: each x_i has a potentially-different mass function p_i .

We imagine that embedding consists of two stages: embedding locations are selected, then a fixed operation is applied at the selected locations. (This model only covers binary embedding, but see Sect. 6 for more on this.) This process is determined by the hidden

payload, but if the payload is unknown it can be modelled probabilistically: location i will be selected with probability π_i . When selected, the pixel distribution is changed to some other distribution \mathbf{q}_i . The embedder's strategy is to choose the probabilities π_i . The distributions \mathbf{p}_i and \mathbf{q}_i are assumed to be public knowledge.

The attacker wants to distinguish cover and stego objects. A *knowing attacker* is granted knowledge of the embedder's strategy π_1, \dots, π_n , and will adjust their detector accordingly. In [6] the zero-sum payoff is modelled by the *deflection* of a linear detector, between the cases of cover and stego object, which determines (amongst other things) the large-sample false positive rate when the true detection rate is 50%. It is shown that a) linear detectors are an optimal subclass under Neyman-Pearson criteria, and b) at this game's equilibrium between embedder and attacker, the squared deflection is

$$\sum \pi_i^2 c_i$$

for some constants c_i called *costs* determined from \mathbf{p}_i and \mathbf{q}_i , and hence known to both players.

The paper contrasts this with the result when the attacker is not granted knowledge of the embedder's strategy π_1, \dots, π_n , the so-called *ignorant* attacker: there the deflection is proportional to $\sum \pi_i c_i$ for the same costs c_i , which justifies the minimization of total average distortion in such a case. But in this work we only consider the knowing attacker.

Our starting point, then, is the following optimal embedding paradigm. Given n pixels in which to hide, and corresponding costs c_1, \dots, c_n , a *distortion limited sender*, who wants to hide the maximum payload subject to a deflection constraint¹, solves the constrained optimization problem

$$\arg \max_{\pi_1, \dots, \pi_n} m = \sum_{i=1}^n H(\pi_i) \text{ subject to } \sum_{i=1}^n \pi_i^2 c_i \leq D. \quad (\text{DLS})$$

Here D is the maximum permissible squared deflection, and then m is the resulting length of payload: the relative entropy of the stego object given the cover which, thanks to the Gel'fand-Pinsker theorem, is the capacity of the cover to convey hidden payload under perfect coding. Although it is optimistic to assume that perfect coding exists, linear codes approach such capacity asymptotically [3], justifying this assumption.

Alternatively, a *payload limited sender* has a fixed payload length to communicate, and solves the constrained optimization problem

$$\arg \min_{\pi_1, \dots, \pi_n} \delta^2 = \sum_{i=1}^n \pi_i^2 c_i \text{ subject to } \sum_{i=1}^n H(\pi_i) \geq M, \quad (\text{PLS})$$

where M is the desired payload size, and then δ^2 is the resulting squared deflection.

We highlight a few properties of these two convex optimization problems, adapting arguments used for the ignorant detector [3]. We may exclude any $\pi_i > \frac{1}{2}$, since such a choice can never be optimal: $1 - \pi_i$ would reduce distortion for the same entropy. Within this range, both $\sum \pi_i^2 c_i$ and $\sum H(\pi_i)$ are increasing in each π_i , so at the optimum the constraints of (DLS) and (PLS) hold with equality. Furthermore, both (DLS) and (PLS) have (up to reparameterizing

of the multiplier) the same Lagrangian

$$\mathcal{L}(\pi_1, \dots, \pi_n, \lambda) = \sum H(\pi_i) + \lambda \sum \pi_i^2 c_i.$$

Therefore, setting its gradient to zero, they both have the same solution, occurring when, for each i ,

$$\frac{\pi_i}{H'(\pi_i)} = \frac{1}{\lambda c_i}. \quad (1)$$

2.1 Exclusion of Pathological Cases

In the theory of steganographic capacity, two types of pathological case are typically excluded; they were identified in [5] as conditions termed *no free bits* and *no determinism*. We have similar conditions for the result here.

We cannot have too many cases of $c_i = 0$: each represents a 'free bit' for the embedder, who can set $\pi_i = 0.5$ to carry a payload bit while contributing nothing to the total cost. A regular supply of free payload bits will break the square root law, and contradicts the result of this paper. We also need to ban weird asymptotic behaviour where some non-negligible subset of c_i are not zero, but tend to it (which might happen if, for example, the pixels' dynamic range depends on n). Our *no free bits* condition can be simply

$$\exists \underline{c} > 0. \forall i. c_i > \underline{c}, \quad (\text{A})$$

but in fact we need not always ban free bits completely. It suffices for their number to be asymptotically negligible, compared with the steganographic capacity. Such a weaker *no free bits* condition is

$$\exists \underline{c} > 0. \frac{\#\{c_i < \underline{c}\}}{\sqrt{n} \log n} \rightarrow 0 \text{ as } n \rightarrow \infty. \quad (\text{A}')$$

Clearly (A) implies (A'), and the reverse is also true if c_i are generated from some stationary ergodic² process, but otherwise the second condition is weaker.

We also cannot have too many $c_i = \infty$: these are unusable locations for the embedder, either because the corresponding cover location is deterministic, or because altering it would violate some constraint of the cover. We also need to ban weird asymptotic behaviour where most of the c_i are not infinite, but tend to infinity. We need not ban unusable locations completely, nor even have them happen negligibly-often: we only need at least linearly many usable locations. So for this paper the *no determinism* condition is

$$\exists \bar{c}. \frac{\#\{c_i > \bar{c}\}}{n} \rightarrow \kappa < 1 \text{ as } n \rightarrow \infty. \quad (\text{B})$$

3 RESULT

Our result concerns the asymptotic behaviour of solutions to (DLS) and (PLS), as $n \rightarrow \infty$. In the first case, we have a fixed acceptable deflection and determine how the costs influence the rate at which m grows; in the second we have a fixed payload size and determine how the costs influence the rate at which δ^2 diminishes.

THEOREM 1. *Suppose that c_1, c_2, \dots is a fixed infinite sequence of costs (perhaps drawn from some stationary distribution, but this need not be the case) satisfying conditions (A) and (B). Write $C = \sum_{i=1}^n 1/c_i$ for the total inverse cost in a cover of size n , which we call the **capacity coefficient**.*

¹This is referred to as the *detectability limited sender* in [2].

²Ergodic in the sense that the ensemble mean equals the time-average mean: a cost with strictly positive probability happens linearly often, with probability 1.

(a) Fix D . As $n \rightarrow \infty$, the solution to **(DLS)** satisfies

$$m \sim \frac{\sqrt{DC}}{2} \log_2(C/D). \quad (2)$$

(b) Fix M . As $n \rightarrow \infty$, the solution to **(PLS)** satisfies

$$\delta^2 \sim \frac{M^2}{C(\log_2(C/M))^2}. \quad (3)$$

(c) We may weaken the assumption that D is fixed, to $D/n \rightarrow 0$, and in (b) to $M/n \rightarrow 0$. Furthermore, (a) also holds with **(A)** weakened to **(A')**, but (b) does not.

So in both cases the behaviour of payload/distortion is determined by the capacity coefficient C , which is the total of the inverse costs in the cover. This uncovers the asymptotic relationship between capacity, security, and the set of costs.

The result makes intuitive sense. Parts of the cover with higher costs contribute relatively little to the overall capacity, because they must be seldom used; as the cost tends to infinity, its contribution to capacity tends to zero; infinite cost locations contribute nothing.

The generalization of (a) to the case of negligibly-many zero costs is helpful since, in the real world, we might encounter a few such locations. Allowing D or M to diminish with n is practically useful since the embedder may wish to increase M with n (but sublinearly!). Indeed, probing this relationship further, we recover the famous Square Root Law by simple manipulation of (3):

COROLLARY. If

- (i) $m \sim r\sqrt{n} \log \sqrt{n}$, and
- (ii) the costs are ergodic, so that $C \sim an$ where a is the **average inverse cost** $\lim \frac{1}{n} \sum 1/c_i$, then

$$\delta^2 \rightarrow \frac{r^2}{a}. \quad (4)$$

This result shows that, even in the adversarial model where the detector knows the embedder's costs, the critical rate for embedding is $r\sqrt{n} \log \sqrt{n} = \frac{r}{2} \sqrt{n} \log n$, and the asymptotic deflection can be determined from the 'root rate' (here 'root-times-log rate') r and the average inverse cost. If m grows asymptotically strictly faster than $\sqrt{n} \log n$ then (3) ensures that the deflection tends to infinity, and if m grows strictly slower than the critical rate then the deflection tends to zero.

4 PROOF

Define the following four functions:

$$H(x) = -x \log_2 x - (1-x) \log_2(1-x), \quad J(x) = \frac{x}{\log_2 x},$$

$$G(x) = \frac{x}{H'(x)} = \frac{x}{\log_2(1-x) - \log_2 x}, \quad K(x) = \frac{x}{(\log_2 x)^2},$$

extended to $H(0) = G(0) = J(0) = K(0) = 0$. These functions are continuous approaching zero from above.

LEMMA.

- (i) $G^{-1}(x) \sim -x \log_2 x$ as $x \rightarrow 0^+$,
- (ii) $(H \circ G^{-1})(x) \sim x(\log_2 x)^2$ as $x \rightarrow 0^+$,
- (iii) $J^{-1}(x) \sim x \log_2 x$ as $x \rightarrow \infty$,
- (iv) $K^{-1}(x) \sim x(\log_2 x)^2$ as $x \rightarrow \infty$.

PROOF. For (i), compute

$$\begin{aligned} \frac{-G(x) \log_2 G(x)}{x} &= \frac{-1}{H'(x)} \log \left(\frac{x}{H'(x)} \right) \\ &= \frac{\log_2 x}{\log_2 x - \log_2(1-x)} + \frac{\log H'(x)}{H'(x)}. \end{aligned}$$

Since $H'(x) \rightarrow \infty$ as $x \rightarrow 0^+$, the first term tends to one and the second to zero, establishing $-G(x) \log_2 G(x) \sim x$. Composing both sides with the continuous (at zero from above) function G^{-1} gives the first result. Then (ii) follows by composing $H(x) \sim -x \log_2 x$ with the previous result.

(iii) follows from

$$\frac{J(x) \log_2 J(x)}{x} = \frac{\log_2 x - \log_2 \log_2 x}{\log_2 x} \rightarrow 1, \quad (5)$$

as $x \rightarrow \infty$, so $J(x) \log_2 J(x) \sim x$. Compose with J^{-1} for the result. (iv) is similar. \square

PROOF OF THEOREM 1(a). Write (1) as $\pi_i = G^{-1}(\frac{1}{\lambda c_i})$, but also rewrite it as $\pi_i c_i = \frac{H'(\pi_i)}{\lambda}$. Note that m, C, λ and the optimal solution π_1, \dots, π_n all depend on n , but we leave this dependence implicit for tidiness of notation. Now observe: G is increasing, $G(1/3) = 1/3$, H' is decreasing, and $H'(1/3) = 1$; therefore

$$H'(\pi_i) \geq 1 \iff \pi_i \leq \frac{1}{3} \iff \frac{1}{\lambda c_i} \leq \frac{1}{3} \iff c_i \geq \frac{3}{\lambda}.$$

Now consider the total distortion, bounding below:

$$\begin{aligned} \sum \pi_i^2 c_i &\geq \sum_{i: c_i \leq \frac{3}{\lambda}} \pi_i^2 c_i + \sum_{i: \frac{3}{\lambda} \leq c_i \leq \bar{c}} \frac{H'(\pi_i)^2}{\lambda^2 c_i} \\ &\geq \frac{n_1 \underline{c}}{9} + \frac{n_2}{\bar{c} \lambda^2}, \end{aligned}$$

where n_1 is the number of $c_i \leq \frac{3}{\lambda}$ and n_2 the number of $\frac{3}{\lambda} \leq c_i \leq \bar{c}$. By **(B)**, $n_1 + n_2$ is linear in n . It follows that, if

$$\sum \pi_i^2 c_i \leq D, \text{ for all } n, \quad (6)$$

with D fixed, then $\lambda \rightarrow \infty$ as $n \rightarrow \infty$.

Now we know that $1/\lambda c_i \rightarrow 0$ for each i , $\pi_i \rightarrow 0$ so that any particular embedding change has diminishing probability, we can use the asymptotics of G^{-1} to obtain

$$\pi_i = G^{-1}\left(\frac{1}{\lambda c_i}\right) \sim \frac{1}{\lambda c_i} \log_2 \lambda c_i.$$

We wish to draw conclusions about $\sum \pi_i^2 c_i$ and $\sum H(\pi_i)$, but it is not in general true that $a_i \sim b_i$ for all i implies $\sum a_i \sim \sum b_i$ ³. In this case, however, because c_i is bounded below by \underline{c} , the arguments x_i to each G^{-1} are bounded above by $1/\underline{c}\lambda$, which tends to zero. Therefore the convergence of π_i to $-x_i \log_2 x_i$ is uniform in i , so it follows that

$$D = \sum \pi_i^2 c_i \sim \sum \frac{(\log_2 \lambda c_i)^2}{\lambda^2 c_i}. \quad (7)$$

On the other hand, use the asymptotics of $H \circ G^{-1}$ and consider the payload:

$$m = \sum H(\pi_i) = \sum (H \circ G^{-1})\left(\frac{1}{\lambda c_i}\right) \sim \sum \frac{(\log_2 \lambda c_i)^2}{\lambda c_i}. \quad (8)$$

³Take for example $a_i = (1 + i/n)$ and $b_i = 1$ for all i .

Comparing (7) and (8), we see that the payload is asymptotically equal to λD . It remains to determine the asymptotics of λ . We get this by expanding (7),

$$D \sim \frac{1}{\lambda^2} \left[(\log_2 \lambda)^2 \sum \frac{1}{c_i} + (\log_2 \lambda) \sum \frac{\log_2 c_i}{c_i} + \sum \frac{(\log_2 c_i)^2}{c_i} \right].$$

Note that $\sum \frac{1}{c_i}$ is at least linear in n , thanks to (B), whereas both $\sum \frac{\log_2 c_i}{c_i}$ and $\sum \frac{(\log_2 c_i)^2}{c_i}$ are at most linear in n , because $|\log x|/x$ and $(\log x)^2/x$ are bounded for x bounded away from zero, and (A) ensures that c_i is bounded away from zero. Therefore the first term dominates and

$$D \sim \frac{(\log_2 \lambda)^2}{\lambda^2} \sum \frac{1}{c_i} = \frac{C}{J(\lambda)^2},$$

giving $\lambda \sim J^{-1}(\sqrt{C/D})$ and hence

$$m \sim \lambda D \sim DJ^{-1}(\sqrt{C/D}). \quad (9)$$

Plugging in the asymptotics of J^{-1} gives the result. \square

PROOF OF THEOREM 1(b). Follows a similar argument. To get started we must show $\lambda \rightarrow \infty$, this time using the fact that the optimum of (PLS) occurs when the constraint is equality. Bounding the payload size below:

$$\sum H(\pi_i) \geq \sum_{c_i < \bar{c}} (H \circ G^{-1})\left(\frac{1}{\lambda c_i}\right) \geq n_1(H \circ G^{-1})\left(\frac{1}{\lambda \bar{c}}\right).$$

By (B), n_1 is linear in n , and by monotonicity of $H \circ G^{-1}$ it follows that $\lambda \rightarrow \infty$.

By exactly the same calculations as before, $\delta^2 \sim M/\lambda$, and the asymptotics of $H \circ G^{-1}$ also give

$$M = \sum H(\pi_i) \sim \sum \frac{(\log_2 \lambda c_i)^2}{\lambda c_i} \sim \sum \frac{(\log_2 \lambda)^2}{\lambda} \frac{1}{c_i}.$$

Therefore $\lambda \sim K^{-1}(C/M)$ and hence

$$\delta^2 \sim M/\lambda \sim M/K^{-1}(C/M). \quad (10)$$

Plugging in the asymptotics of K^{-1} gives the result. \square

PROOF OF THEOREM 1(c). First note that C is linear in n , since $\frac{n(1-\kappa)}{\bar{c}} \leq C \leq \frac{n}{\bar{c}}$. This follows from (A) and (B).

To see that the assumption in (a) may be weakened to $D/n \rightarrow 0$, observe that we only needed to control D at two places: after (6) to ensure $\lambda \rightarrow \infty$, which is still true as long as D is sublinear in n , and after (9) to ensure $C/D \rightarrow \infty$, true since C is linear in n . The same applies for PLS.

To weaken (A) to (A') in the DLS case, note that for fixed D , $m = \Theta(\sqrt{n} \log n)$. If we have a sequences of costs where (A') holds, but not (A): set all costs below \underline{c} to zero, embed one bit for free in each such location, and then remove those locations from consideration. (A) is now true for the rest of the locations, and we can apply the previous result. By (A'), giving the embedder this many free bits increased the payload m by asymptotically strictly less than $\sqrt{n} \log n$, so m was unchanged asymptotically compared with (2). The same does not hold for the PLS case when M is fixed, because the free (or asymptotically free) bits could be enough to cover the entire payload. It would be true for cases where M grows strictly between $\Omega(\sqrt{n} \log n)$ and $O(n)$. \square

5 SIMULATIONS

We now perform simulations to confirm the theoretical results. Ideally we would wish to demonstrate their significance in steganalysis, for example in the case of PLS comparing the performance of detectors against the predictions for δ^2 as a function of M and C . However, the theory applies to knowing detectors against embedders using statistically accurate costs (see discussion in [6, §6]); there are some empirically-determined knowing detectors, but there is no reason to believe that they are optimal, and there are no embedders based on true statistical costs: the closest would be MiPOD [9], which is optimal only for an independent nonstationary Gaussian model of cover pixels. We will have to wait for the development of truly optimal cost-based steganography and steganalysis.

But recall that our result is not only about steganography: it gives asymptotic behaviour of solutions to a class of optimization problems. We do not need to perform steganalysis to confirm the asymptotic behaviour of the solutions to (DLS) and (PLS), for values of the parameters informed by realistic steganographic examples.

Most of our tests relate to (DLS), where we anticipate that the relationship with inverse cost would be most applicable (informing a steganographer how much they can safely embed). We will generate a sequence of costs, fix a maximum tolerable square deflection D , and compare the payload achieved at the optimum of (DLS) with (2). We will also compare with the more precise result (9). The latter still connects m with C and D , but it will turn out to be more accurate because the asymptotics of J^{-1} , used to derive (2), only converge very slowly⁴. The experiments will be performed for cover sizes $n \in \{100, 200, 400, 1000, 2000, \dots, 10^8\}$.

Our first experiment simulates discrete costs. We drew a sequence c_i independently from a Poisson distribution with parameter 5, adding 1 to avoid zero costs. We set $D = 2$. We plot the true and estimated maximum payloads in Fig. 1. The more precise equation (9) converges rapidly to the true solution and is more accurate than (2), but for large n even the latter is no more than a small constant multiple from the true value.

For a second experiment we simulated continuous, *statistically accurate* costs. Following [6, §4] we generated covers of independent binary pixels, where p_i the probability for pixel i is itself drawn from a Beta(5, 50) distribution; the true costs are then $c_i = (1 - 2p_i)^2 / p_i(1 - p_i)$ (see Eq. (7) of [6], but note that it contains a typo). This cost distribution has a very long tail, the extent of which cannot be seen in the histogram in Fig. 2: 0.35% of the costs are over 50 and the largest is approximately 600. (A tail of high costs is often observed in image and video steganography.) This time we chose $D = 1$. Since the costs are statistically accurate for the generated covers, we can interpret D as the square deflection of an optimal detector, who therefore in this case must make approximately 15.9% false positives if they make 50% true positive detections. Subject to this constraint, payloads are plotted in Fig. 2; as in the previous experiment, adherence to the theory is observed.

We also tested against a set of costs used in typical image steganography, the S-UNIWARD costs (with parameter $\sigma = 1$ [4]) from a single BOSSBase image [1]. About 5.7% of these costs are infinite. Drawing with replacement, we sampled them to get the sequence c_i . We know that these are not true statistical costs, so we set a

⁴Looking at (5), only as fast as $(\log_2 \log_2 x)/\log_2 x \rightarrow 0$.

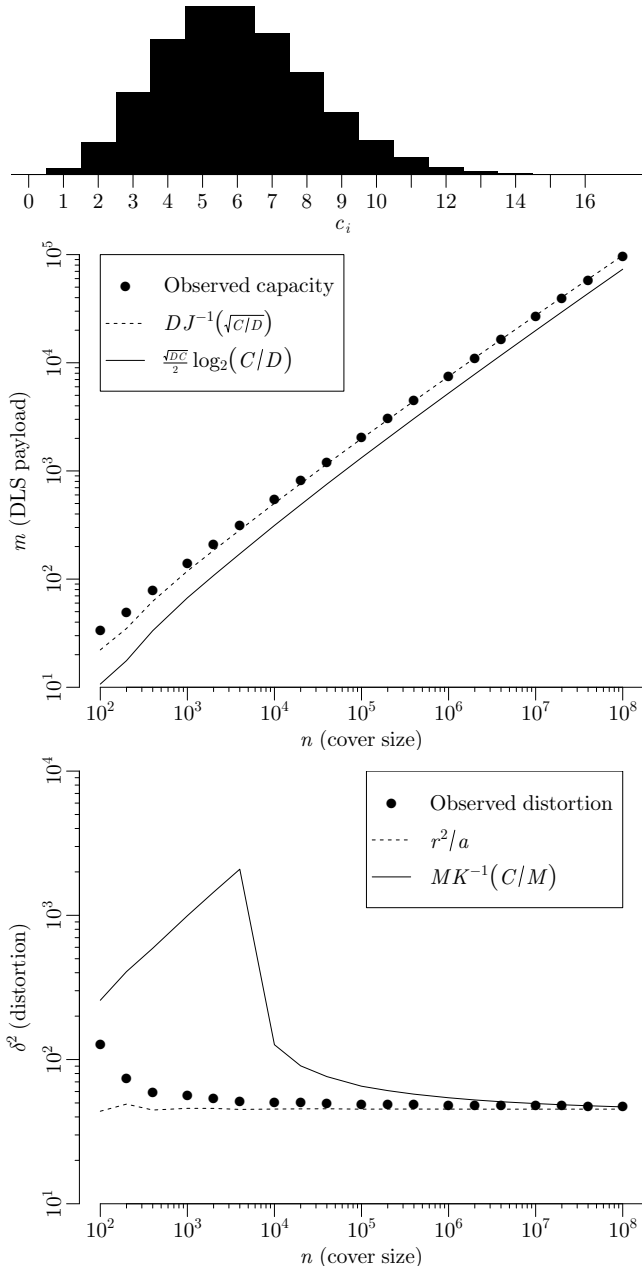


Figure 1: Top: costs drawn from $\text{Poi}(5) + 1$. Middle: true maximum payloads under DLS, and those estimated using (2) and (9), for different cover sizes and $D = 2$. Bottom: squared deflection for the PLS case, when $m \sim 3\sqrt{n} \log \sqrt{n}$, compared to the theoretical predictions (10) and (4). Log-log axes.

deflection that makes sense in practice: $D = 1000$ gives a secure payload of around 0.1 bits per pixel for covers sized like the BOSS-Base images, which is the order of magnitude typically tested in steganalysis literature (e.g. [4, 7]), so this seems appropriate. True and estimated payloads are plotted in Fig. 3. The formulae predict

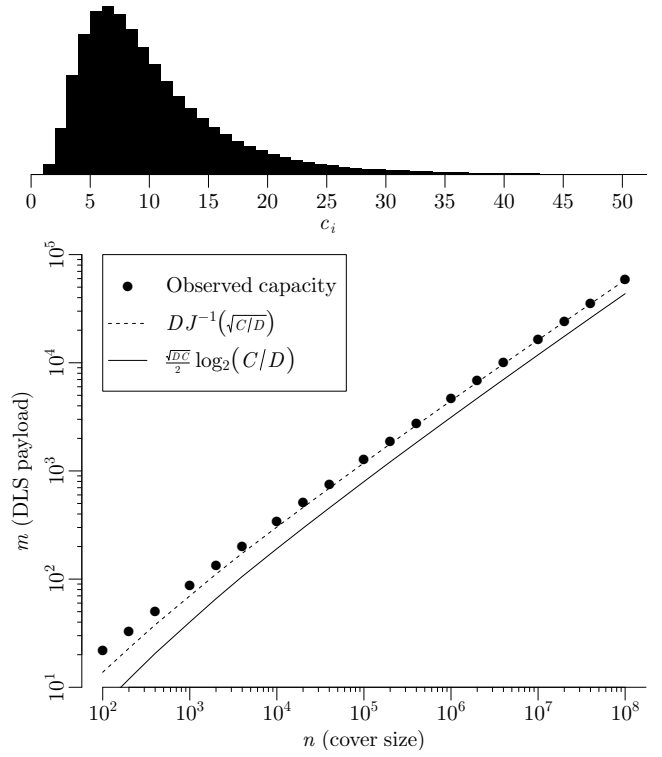


Figure 2: Above: statistically accurate costs for artificial binary images. Below: true maximum payloads, and those estimated using (2) and (9), for different cover sizes and $D = 1$. Log-log axes.

nonsense answers for small n , but converge rapidly to the true answer for reasonable n .

Finally, we also tested the PLS version of our result, in this case focussing on the root rate calculation (4). For the same set of cover sizes, and using the Poisson cost sequence, we set each payload size at $M = r\sqrt{n} \log_2 \sqrt{n}$ for $r = 3$, plotting the observed square deflection obtained in Fig. 1 (bottom). These distortion values, found at the optimum of (PLS), converge rapidly to the theoretically-predicted⁵ limit $r^2/a = 45/(1 - e^{-5})$. It is curious that they do so even faster than the prediction of (10), perhaps the result of fortuitous cancellation of log log factors.

Similar behaviour occurs for the other cost distributions, as long as r is not too large, but we omit the graphs for lack of space.

6 FURTHER DIRECTIONS

This short paper uncovers the relationship between a cover's costs and its capacity, highlighting the significance of the *capacity coefficient* $C = \sum 1/c_i$. We presented an asymptotic result (skimming some of the mathematical details) and some simple simulations. There are several further directions.

⁵For this cost distribution, a can be determined analytically: the expectation of $1/X$, where $X \sim 1 + \text{Poi}(\lambda)$, is $\sum_{i=0}^{\infty} e^{-\lambda} \frac{\lambda^i}{i(i+1)} = \frac{1}{\lambda} \sum_{i=1}^{\infty} e^{-\lambda} \frac{\lambda^i}{i!} = \frac{1}{\lambda} (1 - e^{-\lambda})$.

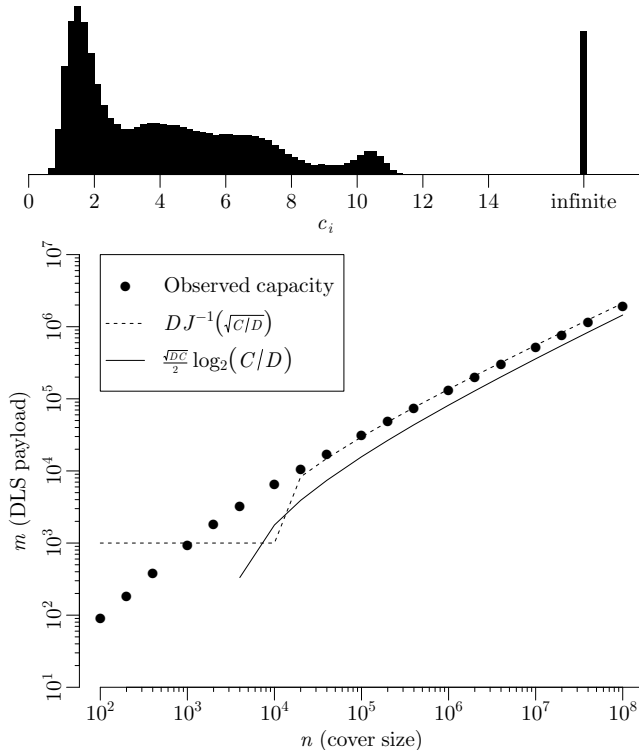


Figure 3: Above: genuine S-UNIWARD costs from one of the BOSSBase images. Below: true maximum payloads, and those estimated using (2) and (9), for different cover sizes and $D = 1000$. Log-log axes.

We can conceive one more asymptotic result, for what we call the *distortion- and payload-limited sender*, who varies the size of the cover to meet both distortion and payload bounds. We postpone it to further work.

In this paper we only considered binary embedding. In fact, similar asymptotic results hold for arbitrary k -ary embedding, and the asymptotic capacity is unchanged. The k -ary equivalent of ‘inverse cost’ is complex, however, and it depends on cost interactions within each pixel; this requires more space for a proper exposition. Furthermore, the choices of *which* k -ary changes are made turn out to be insignificant to payload size: consider a distribution of possible changes in one pixel $((1 - \pi), \pi a_1, \dots, \pi a_k)$, where π represents the probability that a change is made, and a_1, \dots, a_k the relative weights on each possible change. The entropy of this distribution is, asymptotically as $\pi \rightarrow 0$, only $-\pi \log_2 \pi$: most of the information is encoded in whether a change happens, not which change it is. Such a limit applies for any sublinear payload. This is not to say that the choices of k -ary changes do not matter to detectability.

Given that embedding minimizing total cost dominates the literature, a natural question is whether the same results hold when distortion $\sum \pi_i^2 c_i$ is replaced by $\sum \pi_i c_i$ (modelling the ignorant detector). There is an upper bound in terms of C , but in the limit as $D/n \rightarrow 0$ the capacity is governed by $\min c_i$ and all changes cluster

in the cheapest locations. This is a slightly absurd conclusion that only emphasises how exploitable is steganography that optimizes $\sum \pi_i c_i$.

Finally, we intend to examine whether these results are observed in practice, even with non-optimal steganography and steganalysis: if they were as practically-robust as the Square Root Law seems to be, the knowledge that total inverse cost determines capacity would be key information for steganographers.

ACKNOWLEDGMENTS

The author thanks Patrick Bas and Jessica Fridrich, who independently asked him about the relationship between embedding capacity and costs, prompting this work.

REFERENCES

- [1] P. Bas, T. Pevný, and T. Filler. 2011. BOSSBase: Break Our Steganographic Scheme. <http://webdav.agents.fel.cvut.cz/data/projects/stegodata/BossBase-1.01-cover.tar.bz2>. (May 2011).
- [2] R. Cigran, V. Sedighi, and J. Fridrich. 2017. Practical Strategies for Content-Adaptive Batch Steganography and Pooled Steganalysis. In *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '17)*. IEEE, 2122–2126.
- [3] T. Filler, J. Judas, and J. Fridrich. 2011. Minimizing Additive Distortion in Steganography using Syndrome-Trellis Codes. *IEEE Transactions on Information Forensics and Security* 6, 3 (2011), 920–935.
- [4] V. Holub, J. Fridrich, and T. Denemark. 2014. Universal Distortion Function for Steganography in an Arbitrary Domain. *EURASIP Journal on Information Security* 2014, 1 (2014), 1–13.
- [5] A. D. Ker. 2017. The Square Root Law of Steganography: Bringing Theory Closer to Practice. In *Proc. 5th Workshop on Information Hiding and Multimedia Security (IH&MMSec '17)*. ACM, New York, NY, 33–44.
- [6] A. D. Ker, T. Pevný, and P. Bas. 2016. Rethinking Optimal Embedding. In *Proc. 4th Workshop on Information Hiding and Multimedia Security (IH&MMSec '16)*. ACM, New York, NY, 93–102.
- [7] B. Li, M. Wang, J. Huang, and X. Li. 2014. A New Cost Function for Spatial Image Steganography. In *Proc. IEEE International Conference on Image Processing (ICIP)*. IEEE, 4206–4210.
- [8] W. Luo, Y. Zhang, and H. Li. 2017. Adaptive Audio Steganography Based on Advanced Audio Coding and Syndrome-Trellis Coding. In *Proc. International Workshop on Digital Forensics and Watermarking (Lecture Notes on Computer Science)*, Vol. 10431. Springer, 177–186.
- [9] V. Sedighi, R. Cigran, and J. Fridrich. 2016. Content-Adaptive Steganography by Minimizing Statistical Detectability. *IEEE Transactions on Information Forensics and Security* 11, 2 (2016), 221–234.
- [10] P. Wang, H. Zhang, Y. Cao, and X. Zhao. 2016. A Novel Embedding Distortion for Motion Vector-Based Steganography Considering Motion Characteristic, Local Optimality and Statistical Distribution. In *Proc. 4th ACM Workshop on Information Hiding and Multimedia Security (IH&MMSec '16)*. ACM, New York, NY, 127–137.