Secure Payload Scaling for Source Adaptive Payload Allocation

Eli Dworetzky, Edgar Kaziakhmedov, and Jessica Fridrich, Department of ECE, SUNY Binghamton, NY, USA, {edworet1, ekaziak1, fridrich}@binghamton.edu

Abstract

Assuming that Alice commits to an embedding method and the Warden to a detector, we study how much information Alice can communicate at a constant level of statistical detectability over potentially infinitely many uses of the stego channel. When Alice is allowed to allocate her payload across multiple cover objects, we find that certain payload allocation strategies that are informed by a steganography detector exhibit super-square root secure payload (scaling exponent 0.85) for at least tens of thousands of uses of the stego channel. We analyze our experiments with a source model of soft outputs of the detector across images and show how the model determines the scaling of the secure payload.

Introduction

In steganography with digital media objects, no matter how hard the steganographer tries, there exists a model within which the embedding will be detectable [4].[1] A consequence of this postulate is the so-called square root law (SRL) [17, 15] that states the existence of a critical rate that determines whether the covert communication will be asymptotically perfectly secure or whether the steganographer will be eventually caught with certainty.

In this paper, we take a look at a complementary setup. Starting with a Warden's detector built for a known steganographic technique, we ask how much information the steganographer can communicate over multiple uses of the stego channel at a fixed detectability w.r.t. Warden's detector by cleverly allocating payloads to individual images but without making any adjustments to the embedding algorithm. We study this problem within the context of batch steganography and pooled steganalysis [18, 23, 21, 26, 25, 1, 14] with an increasing size of the bag of images.

First, we consider this problem for the hypothetical scenario when the steganographer has access to Warden's detector (worst case for the Warden) as well as the more realistic setup when she does not have access to Warden's detector. Surprisingly, under both scenarios our experiments indicate that there exist payload allocation strategies whose secure payload exhibits super-SRL scalings. To understand why, we adopt a source model on the soft output of a detector and link its numerical characteristics to the scaling exponent of the secure payload.

The paper is structured as follows. After introducing the necessary concepts and definitions in the next section, in Section "Experiments", we study the scaling of secure payload experimentally. Inspired by the results, in Section "Analysis" we impose a statistical model of the soft output of Warden's detector and derive the relationship between this source model and the secure payload scaling. This allows us to explain the experimentally observed trends and how they are affected by the source model and the payload allocation strategy. In Section "Conclusions," we summarize our work and point out possible future directions.

Basic concepts

This section introduces some basic concepts and constructs needed in this paper. Throughout the paper, we denote cover images using the symbol X and images (either cover or stego) intercepted by the Warden using the symbol Y. We use boldface symbols to denote n-tuples of objects. In particular, $\mathbf{X} = (X_1, \ldots, X_n)$ denotes a bag of n cover images, and $\mathbf{Y} = (Y_1, \ldots, Y_n)$ denotes a bag (either cover or stego) of n images intercepted by the Warden.

Warden's detector

The Warden detects steganography in a bag in two phases (known as pooled steganalysis). She applies a single-image detector (SID) to each of the communicated images and then pools the soft scalar outputs of the SID.

Formally, the SID is a mapping $d : \mathcal{X} \to \mathbb{R}$, where $\mathcal{X}$ is the space of all images. Having intercepted n images, the Warden's pooler is of the form $\pi : \mathbb{R}^n \to \mathbb{R}$. The Warden infers whether the sender uses steganography by computing $d(Y_i)$ for all n intercepted images Y_i, $i = 1, \ldots, n$, and comparing $\pi(d(Y_1), \ldots, d(Y_n))$ against a threshold determined by some application-dependent requirements, such as controlling the false alarm.

Response curve

We use C to denote the maximum embedding capacity of a cover image $X \in \mathcal{X}$. For a ternary embedding scheme in the spatial domain, $C \leq \log_2 3$ bits per pixel (bpp). Since most steganographic schemes avoid making changes to saturated pixels, the capacity can be strictly smaller than $\log_2 3$.

A response curve (RC) for a cover image X and de-

[1] More precisely, the steganographic Fisher information within this model will be positive.

tector d is the function $\varrho : [0,C] \to \mathbb{R}$ defined by[2]

$$\varrho(\alpha) = \mathbb{E}[d(X(\alpha;k))], \quad 0 \le \alpha \le C, \tag{1}$$

where $X(\alpha;k)$ is X embedded with a secret message of relative length α bpp and stego key k. The expectation is taken over random messages and stego keys. Furthermore, we define the expected shift of the detector response

$$s(\alpha) = \varrho(\alpha) - \varrho(0). \tag{2}$$

Payload allocation strategies

In this paper, we use three different payload allocation strategies (senders).

Let us assume that the sender has a bag of n cover images $X_1, \ldots, X_n$ with embedding capacities $0 \le C_i \le \log_2 3$ bpp. Let $P(n) \in \left[0, \sum_{i=1}^{n} C_i\right]$ bpp be the total payload the sender wants to communicate in the bag. Her payload allocation strategy is an algorithm that assigns relative payloads α_i to each image X_i subject to the constraints $\sum_{i=1}^{n} \alpha_i = P(n)$ and $\alpha_i \in [0, C_i]$ for each i.

The *uniform sender* spreads the payload uniformly across all images in the bag. Since embedding capacities can vary, the uniform sender employs a water filling algorithm. It first finds $\overline{\alpha}$ such that $\sum_{i=1}^{n} \min\{\overline{\alpha}, C_i\} = P(n)$ and then embeds $\min\{\overline{\alpha}, C_i\}$ bpp to each image X_i in the bag. If all $C_i = C$, each image receives the same payload $\alpha_i = \overline{\alpha} = P(n)/n$.

Assuming the steganographer has access to a steganography detector d, the next two payload allocation strategies consider feedback from the detector.

The *greedy sender* fully concentrates the payload in images that induce the smallest shift in detector response. The sender first permutes the bag of images so that $s_i(C_i) = \varrho_i(C_i) - \varrho_i(0)$ are sorted in non-decreasing order $s_{(1)}(C_{(1)}) \le \cdots \le s_{(n)}(C_{(n)})$. The permutation is expressed as $X_{(1)}, \ldots, X_{(n)}$. Let k be the largest integer for which $\sum_{i=1}^{k} C_i < P(n)$. The sender fully embeds images $X_{(1)}, \ldots, X_{(k)}$ with $\alpha_i = C_i$ while leaving $X_{(k+2)}, \ldots, X_{(n)}$ empty. The $k+1$st image $X_{(k+1)}$ is partially embedded with the remaining relative payload $\alpha_{k+1} = P(n) - \sum_{i=1}^{k} C_i$ bpp. We note that when $C_i = C$ for all i, $k = \lfloor P(n)/C \rfloor$.

The *shift-limited sender* (SLS) [25] enforces the shift hypothesis [14] by considering the impact of the embedding on the statistical distribution of detector outputs across cover images. The SLS finds the smallest $\delta > 0$ that leads to the same expected detector output shift when embedding payload α_i in X_i, satisfying $\sum_{i=1}^{n} \alpha_i = P(n)$, and $\delta = s_i(\alpha_i)$ for all i for which $s_i(C_i) \ge \delta$. For images that do not satisfy this condition (images with "flat" response curves), the SLS sets $\alpha_i = C_i$.

We note that in the rare case when the RC of an image is not monotonically increasing, the greedy sender uses the absolute value $|s_i(\alpha)|$ and SLS uses the cumulative max[3] of $|s_i(\alpha)|$ in their implementations.

These senders were selected as a diverse set of payload allocation strategies that are computationally inexpensive. This requirement is important since we plan to experiment with very large bags with tens of thousands of images. In particular, excessive implementation complexity prevented us from working with the minimum deflection sender [25], which minimizes statistical detectability within the model analyzed in Section "Analysis." The same applies to the image merging sender [23], which treats the bag as one large image and lets the steganographic method distribute the payload.

Experiments

This section reports on a series of experiments aimed at establishing the scaling of secure payload when Alice uses modern content-adaptive steganography and the Warden uses state-of-the-art empirical single-image detectors constructed using machine learning. Loosely speaking, the secure payload is the maximal absolute payload that can be communicated by the sender while guaranteeing a fixed level of statistical undetectability. We are interested how the size of the secure payload scales with the number of images sent by such a detectability limited sender (DLS). We first discuss a subtle yet important distinction between implementing a DLS in practice and implementing a DLS to observe scaling laws.

In practice, Alice must implement a DLS based on some knowledge of a detector. For example, she could adopt a model within which a relationship between payload and error rates of an optimal detector can be established [22, 16]. Another method involves fixing the empirical detectability of *Alice's estimate* of the Warden's pooled detector (SID and pooler). In particular, Alice determines payload based on empirical error rates of her own pooled detector (SID and pooler) when evaluated on a dataset of bags. We refer the reader to Sec II of [11] for additional discussion and examples.

The disadvantage of these approaches is that the detectability is bounded only within the model or bounded only with respect to Alice's estimated pooled detector, respectively. There is no guarantee that detectability will be bounded in practice with respect to the *Warden's* pooled detector. In this paper, we wish to study the scaling of secure payload with respect to the detector the Warden committed to. Therefore, to properly observe such a scaling in a non-asymptotic regime we must, for every bag size n, fix the detectability of the Warden's pooled detector by empirical evaluation (see Section "Simulating ..."). We emphasize that Alice can only achieve this DLS in real life conditions if she has full knowledge of the Warden's pooled detector.

Setup of experiments

We assume that there is a source of cover images available to both the sender and the Warden. The sender commits to an embedding scheme for secret communication and the Warden commits to a SID d^{W} and pooler π. The Warden is always fully aware of the actions potentially taken by the sender. This means that the Warden knows the

[2] The RC also depends on the steganographic embedding scheme.

[3] The unidirectional search used to implement SLS requires that RCs are monotonically increasing [25].

sender's embedding scheme and payload allocation strategy. We experimentally determine the scaling of the secure payload for versions of the above payload allocation strategies with two different poolers. In particular, we wish to determine the largest payload the sender can communicate under any circumstances, such as when allocating payloads with Warden's SID d^{W} as well as the case when the sender does not have access to Warden's SID d^{W}.

All experiments were executed on the image dataset ALASKA II [7] developed as in [7] without the final JPEG compression step. This dataset contains 75,000 images, which we randomly split into three disjoint parts of the same size for our experiments (Splits 1–3). Split 1 and Split 2 are used for training detectors while Split 3 was used for assessing the secure payload scaling. In all experiments, the sender uses the embedding algorithm HILL [19], which is simulated to perform on the rate–distortion bound.

A single-image detector d is trained as an SRNet [5] and the Efficient Net B4 [20, 24] pre-trained on ImageNet with the binary task of steganalyzing J-UNIWARD [12] (the so-called JIN pre-training exactly as described in [6]). The refinement of both detectors to detect HILL was done with stego images embedded with relative payloads randomly drawn from the uniform distribution on the set of relative payloads (in bpp)

$$\mathcal{P} = \{0.05, 0.1, 0.2, \ldots, 1.4, 1.5\}. \tag{3}$$

SRNet was trained on Split 1 while B4 was trained on Split 2. Each split was randomly partitioned into disjoint subsets of 22k, 1k, and 2k images for training, validation, and testing, respectively. The CNNs logit is used as the detector's response.

The response curves were estimated on the same grid of payloads $\mathcal{P}$. We computed the average detector response $\hat{\varrho}(\alpha)$ and the standard deviation of detector outputs $\hat{\sigma}(\alpha)$ using 100 stego images (with different PRNG seeds in the embedding simulator) for each payload $\alpha \in \mathcal{P}$.

Poolers

Our experiments include the simple average pooler, which is agnostic w.r.t. the sender's payload allocation strategy

$$\pi_{\mathrm{avg}}(d^{\mathrm{W}}(\mathbf{Y})) = \frac{1}{n}\sum_{i=1}^{n} d^{\mathrm{W}}(Y_i) \tag{4}$$

and the correlator pooler introduced in [25]

$$\pi_{\mathrm{corr1}}(d^{\mathrm{W}}(\mathbf{Y})) = \sum_{i=1}^{n} d^{\mathrm{W}}(Y_i)\overline{s}(\alpha_i). \tag{5}$$

This pooler makes use of a weighting function $\overline{s}(\alpha)$ which is a logistic fit over all embedding shifts $\hat{s}_i(\alpha) = \hat{\varrho}_i(\alpha) - \hat{\varrho}_i(0)$ of the Warden's detector across the 2k image test set of Split 1 or Split 2 (depending on if SRNet or B4 is used, repectively). Note that this pooler is given the true payloads α_i that might reside in the images. In practice, α_i would need to be estimated by the Warden from the images at hand. As shown in [25], the effect of estimating the payloads has a negligible impact on Warden's detection performance for the studied in this prior art.

Algorithm 1 Detectability Limited Sender's binary search for secure payload. This algorithm is performed independently for each bag size n.

```
// Set detectability $\delta$, bag size $n$, number of bags sampled $N$
for $m = 1$ to $N$:
  $\mathbf{X}_m \leftarrow$ sample $n$ images from Split 3 without replacement
$P_{\mathrm{upper}} \leftarrow \min_m\{$total capacity of $\mathbf{X}_m\}$
$P_{\mathrm{lower}} \leftarrow 0$
loop:
  $P \leftarrow (P_{\mathrm{lower}} + P_{\mathrm{upper}})/2$
  for $m = 1$ to $N$:
    $\boldsymbol{\alpha}_m \leftarrow$ compute payload allocation for bag $\mathbf{X}_m$
    $\mathbf{Y}_m \leftarrow$ simulate embedding payloads $\boldsymbol{\alpha}_m$ in bag $\mathbf{X}_m$
    Compute $\pi(d^{\mathrm{W}}(\mathbf{X}_m)), \pi(d^{\mathrm{W}}(\mathbf{Y}_m))$
  $P_{\mathrm{E}}(P,n,N) \leftarrow$ compute $P_{\mathrm{E}}$ from $\{\pi(d^{\mathrm{W}}(\mathbf{X}_m)), \pi(d^{\mathrm{W}}(\mathbf{Y}_m))\}_{m=1}^{N}$
  if $|P_{\mathrm{E}}(P,n,N) - \delta| < 10^{-3}$:
    break
  else if $P_{\mathrm{E}}(P,n,N) > \delta$:
    $P_{\mathrm{lower}} \leftarrow P$
  else:
    $P_{\mathrm{upper}} \leftarrow P$
}
$P_\delta(n) \leftarrow P$
return $P_\delta(n)$
```

We note that our recent work which conceived of detector-informed batch steganography [25] did not include the greedy sender. As far as we know, there is no published benchmark comparing the greedy sender and SLS nor is there a benchmark testing the greedy sender against different poolers. Thus, we initially experimented with additional poolers and found that a modified version of the correlator

$$\pi_{\mathrm{corr2}}(d^{\mathrm{W}}(\mathbf{Y})) = \sum_{i=1}^{n} d^{\mathrm{W}}(Y_i)s_i(\alpha_i), \tag{6}$$

where $s_i(\alpha_i)$ is the shift in the RC of Warden's detector, performs better than π_{corr1} when Alice uses the greedy sender. Such a pooler is justified as the most powerful pooler from a statistical model of detector response in Section "Analysis." Note that π_{corr2} is clairvoyant as it requires knowledge of the exact shift in response $s_i(\alpha_i)$ as opposed to a "global average" $\overline{s}(\alpha_i)$. However, π_{corr1} still universally performs better than π_{corr2} when Alice uses SLS. Therefore, in all experiments, we use π_{corr2} when the setup includes the greedy sender, and we use π_{corr1} when the setup includes SLS.

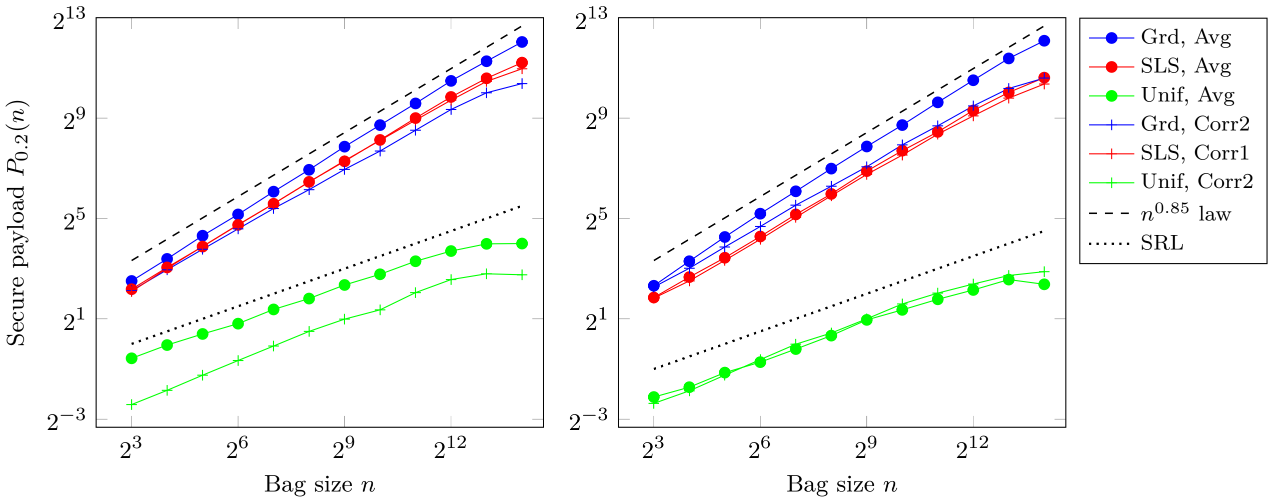

Figure 1. Log-log plot of secure payload vs. bag size n for various payload allocation strategies and poolers. For every n, the Warden's pooler π achieves constant detectability $P_\mathrm{E} = 0.2$. Left: SID SRNet. Right: SID B4.

In summary, we wish to point out that giving the exact payloads α_i and shifts s_i to the Warden is done intentionally to consider the worst case scenario for the steganographer. If a super SRL secure payload is observed for such unrealistically empowered Warden, then in practice the secure payload will be at least as large as what we determine under our assumptions. These clairvoyant pooled detectors are also significantly cheaper to experiment with than detectors that need to estimate these quantities. This aspect is especially important for our study with large bags and an empirical detectability criterion at the Warden's side.

Simulating a detectability-limited sender

For a fixed statistical detectability $\delta \geq 0$ and bag size n, we wish to determine the secure payload size $P_\delta(n)$ that gives the Warden's pooler a prescribed detectability δ. In this paper, we use the detector's minimum total average error probability $P_\mathrm{E} = \min \frac{1}{2}(P_\mathrm{FA} + P_\mathrm{MD})$ as our detectability measure. Of course, alternative measures could certainly be used, such as wAUC, true positive rate for a fixed false alarm, etc.

This problem formulation is known as a detectability-limited sender (DLS). We reiterate that our DLS fixes detectability across a collection of N bags rather than fixing detectability within a statistical model of a specific bag. We therefore must use a payload-limited sender (PLS) to solve for the total payload that achieves the desired detectability δ. In particular, we implement the DLS as a binary search for $P_\delta(n)$. Note that this DLS results in the same payload $P_\delta(n)$ being embedded in every bag while a traditional DLS using a model can have variable payloads depending on the bag.

Each iteration of the binary search does the following (also follow Algorithm 1). We randomly select n cover images $\mathbf{X}$ from Split 3 without replacement[4] and use a payload allocation strategy to compute $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)$ given a desired total payload of P bpp. We generate each stego image $Y_i = X_i(\alpha_i)$ using a random key, producing the stego bag $\mathbf{Y} = \mathbf{X}(\boldsymbol{\alpha})$. The pooled detector is then applied to the cover and stego bags, $\left(\pi(d^\mathrm{W}(\mathbf{X})), \pi(d^\mathrm{W}(\mathbf{Y}))\right)$. This is repeated with a newly sampled cover bag a total of N times. Denoting the outputs of Warden's pooled detector with $\left\{\pi(d^\mathrm{W}(\mathbf{X}_m)), \pi(d^\mathrm{W}(\mathbf{Y}_m))\right\}_{m=1}^{N}$, we compute from this data the empirical detectability $P_\mathrm{E}(P, n, N)$. The search ultimately solves for the payload P such that $P_\mathrm{E}(P, n, N) = \delta$. The payload found this way will be denoted $P_\delta(n)$, omitting the dependence on N since $P_\mathrm{E}(P, n, N)$ saturates for large enough N.

To determine the scaling of the secure payload across a range of bag sizes, the entire binary search is repeated for $n \in \{2^1, 2^2, \ldots, 2^{14}\}$. The maximum bag size that we can study, 2^{14}, is determined by the size of the split3 (25,000). The number of bags N was adjusted with n to control the computational complexity. In particular, $N = 1000$ for $n = 2^1, \ldots, 2^8$, $N = 500$ for $n = 2^9, \ldots, 2^{12}$ and $N = 300$ for $n = 2^{13}, 2^{14}$.

We note that the entire procedure described above includes expensive operations, such as running the embedding simulator and computing forward passes of the CNNs for $O(N \times n)$ images per iteration of the binary search. To speed up our experiments, we sampled logits of stego images in a Monte Carlo fashion by drawing a sample from the Gaussian distribution[5] $\mathcal{N}(\hat{\varrho}(\alpha), \hat{\sigma}^2(\alpha))$ by linearly interpolating $\hat{\varrho}(\alpha), \hat{\sigma}^2(\alpha)$ from the two closest grid points from $\mathcal{P}$. We can verify that the scaling we observe in our simulations is roughly the scaling for real images by running the PLS with real images using the secure payload $P_\delta(n)$ found from the simulation. The verification comes from observing that empirical detectability for real images is approximately our desired detectability δ.

[4] Given the size of the splits, if bags are formed without replacement, we are limited to bag size on the order of $\approx 20,000$.

[5] The Gaussian model is justified in Section "Analysis."

Results

Figure 1 contains log-log plots of $P_\delta(n)$ as a function of n at detectability $\delta = P_\mathrm{E} = 0.2$ for different combinations of senders and Warden's pooled detectors. These results assume that the SIDs used by Alice and the Warden are matched. The left plot uses SRNet and the right plot uses B4. The black reference lines correspond to power laws $n^{0.85}$ and $n^{1/2}$. The most notable result is the scaling for SLS and the greedy senders whose secure payload scales as n^γ with $\gamma \approx 0.85$ with both poolers and both SIDs. The secure payload for the uniform sender scales according to the square root law and serves as a sanity check. In Figure 2, we verify that the super SRL scaling for the greedy sender holds across a range of fixed detectability levels $P_\mathrm{E} \in \{0.1, 0.2, 0.3, 0.4\}$.

We note that the slight departure from power law scaling for the largest bag size is due to the finite dataset (split). The statistical spread of the pooler output on the largest bags whose size is comparable to the dataset size ($n = 16{,}384$ on a split with 25,000 images) is smaller than what it would be for a much larger (infinite) dataset. In the extreme case when the bag size is the entire split, the statistical spread would be zero as there is only one bag.

The authors wish to emphasize that Alice is not changing the embedding algorithm itself but she is strategizing on how to distribute her payload across images based on feedback from Warden's SID. Thus, the observed scaling is perhaps more fittingly interpreted as a property of the source and the Warden's pooled detector rather than a result for an adversarial setting. For a fixed cover source, embedding method, and a pooled detector, the scaling tells us how much information can be communicated through the channel by allocating the payload without triggering the pooled detector. Thus, one can think of the result as a scaling law of source-adaptive payload allocation.

The next batch of experiments aims at establishing the secure payload scaling in a more realistic setting when the payload allocation is carried out with feedback from a *different* SID than the one used by the Warden for detection. To this end, we used the same setup as above but with Alice distributing her payload based on feedback from SRNet while the Warden uses B4 for detection and vice versa. We note that SRNet was trained on Split 1 and B4 on Split 2 to include a mismatch in training data as well. Split 3 is still used for assessing the secure payload scaling.

Figure 3 (left) shows the payload scaling when the sender spreads using feedback from B4 and the Warden uses SRNet for detection. In the right figure, the sender uses SRNet and the Warden uses B4. When the Warden uses the average pooler, the secure payload follows the super-SRL scaling. The effect of using the clairvoyant correlator poolers depends on who has which SID. When the sender uses SRNet and Warden B4, the greedy sender offers a larger secure payload that follows the super SRL scaling while the SLS falls under the SRL. This situation reverses when the sender is given B4 for spreading and Warden SRNet for detection. This asymmetric behavior could be caused by one of the SIDs being superior to the other. Since the greedy sender is more aggressive than SLS in how it uses the SID feedback, when the sender uses the more powerful SID she can communicate larger secure payload with this more aggressive spreading. However, when using the inferior SID, the greedy sender "overdoes" the spreading and becomes more vulnerable to the better SID used by the Warden. In either case, nevertheless, the sender can still communicate super SRL payload despite the mismatched SIDs.

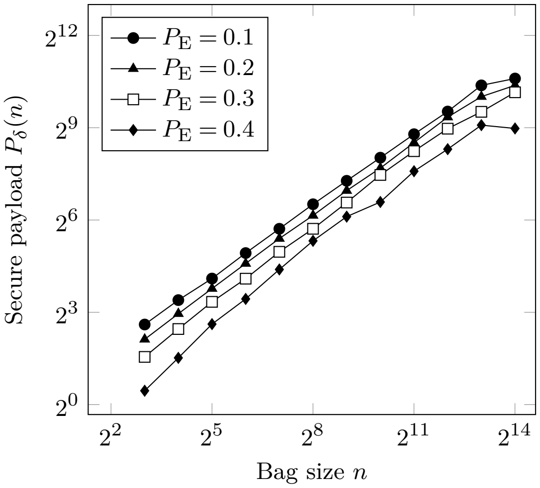

Figure 2. Log-log plot of secure payload vs. bag size n for a range of fixed detectability levels P_E. Alice uses greedy sender and embedding algorithm HILL. Warden uses π_corr2. Both use the same SRNet as their SID.

We remind the reader that all experiments above were executed with a simulated response of the SIDs. To confirm that the payload scaling we observe holds when the Warden feeds actual images to her SID, we perform an iteration of the PLS using real images and the secure payloads found from the simulations. Figure 4 includes 8 cases of SIDs, senders, and poolers. Observe that there is very minor deviation from the target detectability $P_\mathrm{E} = 0.2$ such payloads would achieve in the simulations. We note that the largest bag size 2^{14} is omitted and re-emphasize that it was infeasible for us to run such an expensive binary search using real images given our time constraints.

The experiments presented in the figures so far are for the case when Alice adjusts payload for constant detectability of the Warden's pooled detector. As discussed previously, in practice Alice will likely not be able to determine such secure payloads and will have to either choose secure payloads w.r.t. her own pooled detector or use a PLS determined heuristically. In Figure 5, we show the results for the most realistic case in our study; we have Alice and the Warden use feedback from mismatched SIDs and have Alice determine her absolute payloads from a simple formula $P(n) = 0.5n^\gamma$ which does not give any guarantees on security. Colors are used to represent cases of different SIDs, spreaders, and pooling strategies. Marker symbols are used to represent the value of γ used. Note that $\gamma = 0.9$ tends towards perfect detection while $\gamma = 0.7$ and 0.8 remain roughly in the center of the dynamic range of P_E.

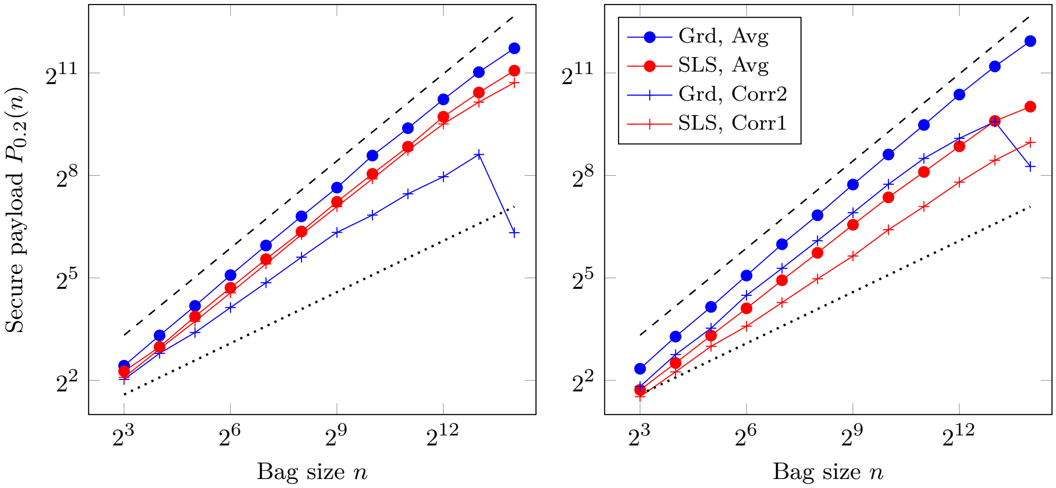

Figure 3. *Log-log plot of secure payload vs. bag size n for the greedy and SLS payload allocation strategies with two poolers. Embedding algorithm HILL, $P_E = 0.2$, SRNet trained on Split 1, B4 on Split 2, evaluated on Split 3. Left: Alice uses B4 and Warden SRNet. Right: Alice SRNet and Warden B4.*

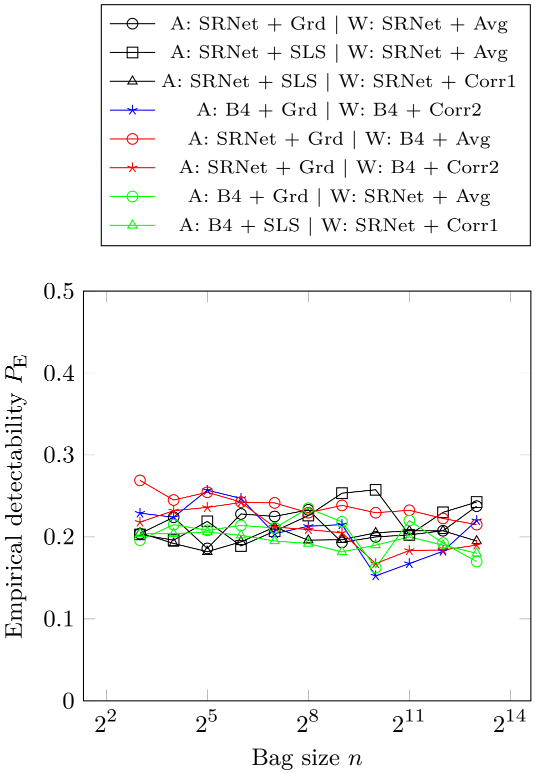

Figure 4. *P_E vs. bag size n when detecting real stego bags. The payloads embedded are the secure payloads $P_\delta(n)$ determined from the simulated experiments. Each curve corresponds to a particular case of Alice's SID & spreader and Warden's SID & pooler. We hypothesize that this figure's legend is one of the world's largest.*

Analysis

In this section, we analyze the experimental results in an attempt to explain the secure payload scaling observed in practice. To this end, we adopt a statistical model of detector soft outputs and study the secure payload of the greedy sender for Warden's most powerful pooler. The greedy sender was selected because the analysis allows us to obtain the main results in an easily interpretable analytic form, which is not the case for the SLS and other senders, such as the MDS or the IMS. Finally, we note that our analysis is restricted to the matched detector case when both the steganographer and the Warden make use of the same single-image detector for their respective goals. This is done intentionally to avoid dealing with the difficult problem of modeling the mismatch between detectors, which would significantly complicate the analysis.

Sampling from the cover source

We envision the process of sampling from the cover source as a two-stage stochastic process. First, the sender selects a (noise-free) scene and then acquires it using a digital imaging sensor. The acquisition is affected by numerous noise sources, such as the shot (photonic) noise, the readout noise, and thermal noise [13]. Thus, taking multiple images of the exact same scene with the same camera would produce slightly different images that follow a statistical distribution that is conditioned on the scene and, technically, also on the camera. We call the random variable following this distribution an *acquisition oracle*, a concept commonly found in steganography in the past [2, 9, 10, 8, 25].

To avoid the potentially infeasibly complex task of modeling the oracle itself, we make assumptions on the distribution of detector outputs on the realizations of the oracle. We take advantage of the fact that the distribution of the acquisitions X for a fixed scene is concentrated on a small subset of $\mathcal{X}$ (multiple images of the same scene taken with the same camera differ only slightly). Since

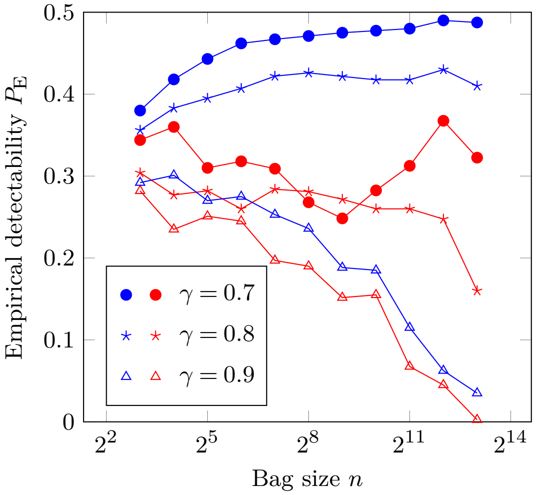

Figure 5. *P_E vs. bag size n when detecting real stego bags for a PLS of the form $P(n) = 0.5n^\gamma$. Each curve corresponds to a different case of scaling exponent $\gamma = 0.7, 0.8, 0.9$ and mismatched SIDs. For blue curves, Alice uses B4 and SLS. Warden uses SRNet and π_{corr1}. For red curves, Alice uses SRNet and greedy sender. Warden uses B4 and π_{corr2}.*

differentiable non-linear functions are approximately linear on sufficiently small neighborhoods, the central limit theorem (CLT) implies that[6]

$$d(X) \sim \mathcal{N}(\mu, \sigma^2), \tag{7}$$

where μ and σ^2 are the expected value and variance of d on cover images generated by the acquisition oracle for a fixed scene. Since stego schemes try to preserve statistical properties of X, the embedding process will also preserve the concentration. Therefore, by the same argument we assume that the detector output on X embedded with relative message length α (denoted as $X(\alpha)$) is also Gaussian

$$d(X(\alpha)) \sim \mathcal{N}(\mu + s(\alpha), \sigma^2). \tag{8}$$

Note that we implicitly assume that only the mean is affected by embedding but not the variance. At the end of this section, we comment on the case when $\sigma^2(\alpha)$ increases with α and its effect on secure payload scaling. This *local* shift hypothesis is a much weaker assumption than the shift hypothesis (as adopted in, e.g., [23]) about the *global* distribution of detector response which is not satisfied for modern steganalyzers in the form of rich models and convolutional neural networks (CNNs) (see Sec. 3.2 in [25]).

Optimal pooler

Let $\mathbf{X} = (X_1, \ldots, X_n)$ be a bag of independently sampled cover images. Equipped with a SID d that adheres to the assumptions above, given a bag of images $\mathbf{Y} = (Y_1, \ldots, Y_n)$ the Warden's hypothesis test becomes:

$$\begin{aligned} \mathcal{H}_0: &\quad d(Y_i) \sim \mathcal{N}(\mu_i, \sigma^2) \quad \text{for all } i \\ \mathcal{H}_1: &\quad d(Y_i) \sim \mathcal{N}(\mu_i + s_i(\alpha_i), \sigma^2) \quad \text{for all } i, \end{aligned} \tag{9}$$

where α_i is the relative payload residing in the ith cover image X_i. The outputs $d(Y_i)$ are pooled by the Warden using pooler π to detect the use of steganography by the sender.

Assuming the parameters of the distributions in the hypothesis test (9) are known to the Warden, the test becomes simple and, due to the independence of cover images, the Warden's most powerful pooled detector is the likelihood ratio test. The detectability of steganography based on evidence from n images $\mathbf{Y} = \mathbf{X}(\boldsymbol{\alpha})$ is thus determined by the deflection coefficient

$$\Delta^2(\mathbf{Y}) = \frac{1}{\sigma^2} \sum_{i=1}^{n} s_i^2(\alpha_i). \tag{10}$$

Cover source model

In this paper, a cover source is modeled by adopting a statistical model for detector response curves (1). As already mentioned in the introduction, modeling the source through the response of a detector is a compromise because the conclusions reached are with respect to a specific detector and embedding scheme. On the other hand, this arrangement gives us a substantial advantage in the form of tractable analysis and estimability of all modeling parameters in practice.

A particularly simple model of response curves is the linear model. For cover image $X_i \in \mathbf{X}$

$$s_i(\alpha) = \varrho_i(\alpha) - \varrho(0) = b_i \alpha, \tag{11}$$

where b_i is the RC's slope. Since the greedy sender either embeds an image fully or not at all,[7] we can equivalently consider a model for the expected increase in detector response[8] at their embedding capacity, $s_i(C) = \varrho_i(C_i) - \varrho(0)$. For this model, the deflection of Warden's optimal pooler is given by

$$\Delta^2(\mathbf{Y}) = \frac{1}{\sigma^2} \sum_{i=1}^{k} s_i^2(C_i) + \frac{1}{\sigma^2} s_{k+1}^2(\alpha_{k+1}), \tag{12}$$

where k is the largest integer such that $\sum_{i=1}^{k} C_i < P$, where P is the payload to be embedded in the bag.

A cover source model is a probabilistic distribution imposed on $s_i^2(C_i)$ for images from a given source:

$$s_i^2(C_i) \sim F, \tag{13}$$

where F is a cumulative distribution function (CDF) supported on $[0, \infty)$. We wish to emphasize that F depends on the cover source, steganographic method, and Warden's detector.

[6] Modern single-image detectors d are often neural networks with differentiable structure.

[7] With the exception of the last image, which may be embedded only partially.

[8] Expectation taken over embeddings with different messages / stego keys.

Secure payload

Given a bag of n cover images $\mathbf{X} = (X_1, \ldots, X_n)$ sampled independently from the cover source, we define the *secure payload of a bag of size* n at detectability $\delta \geq 0$, $P_\delta(n)$, as the largest total expected payload $\sum_{i=1}^{n} \alpha_i$, $\alpha_i \leq C_i$, that can be communicated in these n images that satisfies

$$\mathbb{E}[\Delta^2(\mathbf{X}(\boldsymbol{\alpha}))] \leq \delta, \qquad (14)$$

with the expectation taken over bags of size n.

We will study the scaling of the secure payload only for the greedy sender because the results are available in a closed form. Our goal is to explain the scaling observed in experiments in Section "Results." For simplicity and without loss of generality, we will assume that $\sigma^2 = 1$ and that the embedding capacity of all images is the same, $C_i = C = \log_2 3$ for all i. We further simplify by ignoring the contribution of the last partially embedded image to the deflection (12) as this will asymptotically for large n become negligible.

This main result is proved in the appendix:

Theorem: Let $s^2(C) \sim F$ with $F(x) > 0$ for $x > 0$ continuous and invertible on some right neighborhood of zero. The secure payload of the greedy sender for bag size n and detectability $\delta > 0$ is

$$P_\delta(n) = \log_2 3 \times k(n), \qquad (15)$$

where $k(n)$ is the largest integer satisfying $n \int_0^{k(n)/n} F^{-1}(x)\mathrm{d}x \leq \delta$. In particular, when $F(x) \propto x^\beta$, $\beta > 0$, on some right neighborhood of 0, $P_\delta(n) \propto n^{\frac{1}{\beta+1}}$.

Discussion

The key to understanding the super SRL scaling observed in experiments is the source model and the CDF F. In Figure 6 we show the PDF (left) and the log-log plot of the CDF (right) of $s^2(C)$ for the SID in the form of an SRNet trained for HILL. The PDF shows that images with increasingly small $s(C)$ are very common in our dataset, which is the reason for the super SRL scaling observed in experiments. These images are typically very noisy (taken at high ISO setting) and / or contain complex texture. The CDF includes a line of best fit whose slope is $\beta \approx 0.22$. Using the rule $P_\delta(n) \propto n^{\frac{1}{\beta+1}}$, the line of best fit predicts a power scaling of $\gamma \approx 0.82$ which is roughly what was observed in the experiments with matched detectors in Section "Results."

The secure payload scaling for the greedy sender has been derived under some simplifying assumptions, which will inevitably cause deviations between the theory and experiments. In particular, the hypothesis test (9) that gives rise to the theoretical result is simple as the Warden is given the means μ_i, which is the response of the detector response on the cover $\varrho_i(0)$. In practice, the Warden will not have access to this data and her detector will thus perform suboptimally, allowing the sender to communicate more at the same level of statistical detectability. Moreover, our analysis is limited to all capacities being equal, $C_i = C$, across all images. To properly extend our study to this case, we would need to adopt a source model on the capacities and also most likely consider the joint distribution of $(C, s(C))$, which might be difficult to model and estimate in practice.

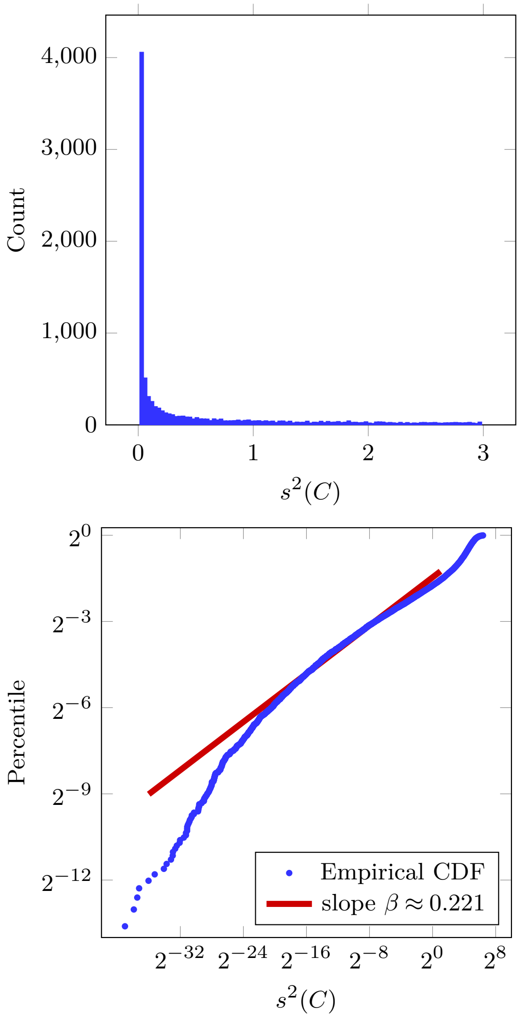

Figure 6. Left: Plot of the PDF F of $s^2(C)$ across Split 3. Right: The corresponding Log-log plot of the CDF with a best fit line of slope $\beta \approx 0.22$.

Our modeling assumption that the variance σ^2 of the detector soft output is constant (9) is also an approximation. Our experiments suggest that the combined standard deviation due to acquisition and embedding, $\sigma_i(\alpha)$, increases approximately linearly with $s_i(\alpha)$; $\sigma_i(\alpha) \approx \sigma + c s_i(\alpha)$, where $c > 0$ does not depend on the image i. The optimal pooler in this case is a sum of a correlator and an energy detector as the Warden can also detect embedding by increased variance of the SID output. For a linear model of response curves $s_i(\alpha) = b_i \alpha$, it can be shown that the leading term of the deflection (in terms of c being small) changes only by a constant multiplicative factor and thus has no effect on secure payload scaling.

At this point, we also wish to contrast our findings with Ker's scaling law of secure payload for content-adaptive steganography [16]. The obvious parallel here is the similarity between payload allocation across images and content-adaptive steganography (payload allocation across pixels). In [16], the author derived a relationship between secure payload and the sum of reciprocals of pixels' embedding costs (Fisher information). One of the key assumptions in this paper that guaranteed the SRL of secure payload was banning a non-negligible source of free bits in the form of pixels' diminishing costs. While this assumption is indeed reasonable on the level of (integer-valued) pixels, outputs of a detector trained on entire images exhibit a much larger diversity. In particular and unlike for pixels, we can accurately estimate the properties of the detector output as a function of payload on individual images. The fact that the distribution of $s^2(C)$ exhibits a spike at zero gives rise to a CDF that leads to the observed super SRL secure payload scaling.

In fact, we have observed this spike for other datasets (BOSSbase [3]) and other embedding algorithms than HILL (not reported in this paper). We plan to investigate the universality of this spiky distribution of response curves' slopes across a wider spectrum of detectors, datasets, and embedding schemes. Should this characteristic of steganography detectors be universally observed, it would indicate a new type of law that the Warden should be aware of in practice:

> *No matter how hard the Warden works to build a detector, there exists a payload allocation strategy that would allow the steganographer to communicate super SRL secure payload.*

Finally, in experiments on finite datasets we will be unable to distinguish between "true" payload scaling n^γ as $n \to \infty$ and a transient scaling that eventually becomes $n^{1/2}$. For example, we might see a different power scaling of the CDF F very close to zero if we were to substantially increase the size of the dataset. A linear scaling $F(x) \propto x$ or $F(x) = 0$ on some small right neighborhood of 0 would imply a secure payload that eventually must follow the SRL for sufficiently large n. Thus, our conclusion concerning the secure payload scaling needs to be understood within the context (and limitation) of our experiments. Technically, we cannot claim that the scaling will be observed for sufficiently large n.

Conclusions

The square root law studies the secure payload scaling based on a postulate that perfectly secure steganography is practically unachievable. In our paper, we take a complementary stance from the perspective of the Warden being unable to construct a perfect detector. We ask the sender to commit to a steganographic algorithm and subsequently the Warden to commit to a detector of that algorithm and study the size of payload that can be communicated at a fixed level of statistical detectability as measured by Warden's pooled detector. Our experimental findings are surprising as they indicate that there exists a strategy for the steganographer to cleverly allocate payload across images that leads to super square root scaling of secure payload. We traced this to the properties of Warden's single-image detector, namely the distribution of the detector response curves across the cover source. This distribution exhibits a spike at zero, effectively meaning that there is a significant source of images with diminishing response to embedding. To quantitatively explain the scaling observed in our experiments, we derive a closed-form expression for the scaling exponent of secure payload size based on the distribution of the detector response.

It should be emphasized that the fact that there exists a payload allocation strategy with super SRL scaling does not mean that the steganographer will always be able to determine this strategy in practice because it requires access to Warden's detector. Nevertheless, we show that, within our experimental setup, the steganographer can make use of her own detector and still enjoy a super SRL scaling. Our findings shed new light on the old question of who will win in the long term – the steganographer or the Warden? A super SRL secure payload scaling certainly allows the steganographer to make a practical use of the cover source.

Our paper brings numerous new questions to the table that we intend to study in the future. The most important one is whether the observed concentration of Warden's detector diminishing response to embedding is a universal phenomenon that will be observed in all datasets, with all kinds of detectors, and for all steganographic methods. Since in our study the super SRL scaling is observed when the Warden is given powerful clairvoyant detectors built using state-of-the-art deep convolutional neural networks, we conjecture that the spike in diminishing detector response is indeed ubiquitous. We believe that it is important for steganalysts to be aware of the possibility that a steganographer might go undetected even when embedding super SRL payloads.

Acknowledgments

The work on this paper was supported by NSF grant No. 2028119.

Appendix

In this appendix, we prove the main result for scaling of the secure payload for the greedy sender. Below, we use the symbol X for the random variable whose realizations are $(\varrho(C) - \varrho(0))^2$, the squared expected detector increase when embedding maximum payload.

We first state some known properties of order statistics of a uniformly distributed random variable on the interval $[0,1]$, $\mathcal{U}[0,1]$. Let $U_{(k)}$ denote the kth order statistic of n i.i.d. uniform random variables $U_i \sim \mathcal{U}[0,1]$. The CDF of $U_{(k)}$ is

$$G_k(x) = \sum_{j=k}^{n} \binom{n}{j} x^j (1-x)^{n-j}. \tag{16}$$

Further inspection of the PDF of $U_{(k)}$ would reveal that $U_{(k)} \sim \mathrm{Beta}(k, n+1-k)$ which means $\mathbb{E}[U_{(k)}] = \frac{k}{n+1}$ and $\mathrm{Var}[U_{(k)}] = \frac{k(n-k+1)}{(n+1)^2(n+2)}$.

Consider n i.i.d. random variables $X_i \sim F$ as well as the kth order statistic of the X_i, denoted by $X_{(k)}$. Suppose that $k = k(n)$ is some function of n satisfying $cn^{1/2} \leq k(n)$ (secure payload is at least $\propto n^{1/2}$) for some $c > 0$ and $k(n)/n \to 0$ (secure payload is sublinear) as $n \to \infty$.

Lemma. $F(X_{(k)})\frac{n+1}{k(n)} \to 1$ in probability as $n \to \infty$.

Proof: Since $F(X) \sim U[0,1]$ and since F is nondecreasing, $F(X_{(k)})$ is the kth order statistic $U_{(k)}$. Let $\epsilon > 0$. Applying Chebyshev's inequality gives us

$$\begin{aligned}
&\mathbb{P}\left\{\left|U_{(k)}\frac{n+1}{k(n)}-1\right| \geq \epsilon\right\}\\
&= \mathbb{P}\left\{\left|U_{(k)}-\frac{k(n)}{n+1}\right| \geq \epsilon\frac{k(n)}{n+1}\right\}\\
&\leq \frac{(n+1)^2}{k^2(n)\epsilon^2}\frac{k(n)(n+1-k)}{(n+1)^2(n+2)}\\
&= \frac{n+1}{n+2}\frac{1-k(n)/(n+1)}{k(n)\epsilon^2}. \qquad (17)
\end{aligned}$$

The upper bound (17) is $\mathcal{O}(n^{-1/2})$ as $n \to \infty$. Hence, we have that $F(X_{(k)})\frac{n+1}{k(n)} \to 1$ in probability, which completes the proof.■

Let $X[0,a]$ denote X conditioned on $[0,a]$. Given n i.i.d. samples X_i and $\epsilon > 0$, the lemma implies that $0 \leq F(X_{(i)}) \leq \frac{k(n)}{n}(1+\epsilon)$ and hence $0 \leq X_{(i)} \leq F^{-1}\left(\frac{k(n)}{n}(1+\epsilon)\right)$ for all $1 \leq i \leq k(n)$ with probability arbitrarily close to 1 for sufficiently large n. The sample statistics $X_{(i)}$, $1 \leq i \leq k(n)$ are realizations of $X\left[0, F^{-1}\left(\frac{k(n)}{n}(1+\epsilon)\right)\right]$. Denoting for compactness $a_{k,n} = F^{-1}\left(\frac{k(n)}{n}(1+\epsilon)\right)$ and realizing that $k\mathbb{E}[X[0,a_{k,n}]] = \delta$ from the detectability requirement, Chebyshev's inequality implies

$$\begin{aligned}
&\mathbb{P}\left(\left|\sum_{i=1}^{k} X_i[0,a_{k,n}] - k\mathbb{E}[X[0,a_{k,n}]]\right| > \epsilon_1\right)\\
&\leq \frac{k\mathrm{Var}[X[0,a_{k,n}]]}{\epsilon_1^2}\\
&= \frac{\delta\mathrm{Var}[X[0,a_{k,n}]]}{\epsilon_1^2\mathbb{E}[X[0,a_{k,n}]]}. \qquad (18)
\end{aligned}$$

This bound approaches 0 with $n \to \infty$ for any ϵ_1 because $a_{k,n} \to 0$ by our assumption on $k(n)$ and due to F being continuous and the fact that

$$\mathrm{Var}[X[0,a_{k,n}]] \leq a_{k,n}\mathbb{E}[X[0,a_{k,n}]]. \qquad (19)$$

Finally, the secure payload $k(n) \times \log_2 3$ is determined from the condition

$$\begin{aligned}
\delta &= k(n)\mathbb{E}[X[0,a_{k,n}]]\\
&= k(n)\frac{n}{k(n)}\int_0^{F^{-1}\left(\frac{k(n)}{n}\right)} x\mathrm{d}F(x)\\
&= n\int_0^{k(n)/n} F^{-1}(x)\mathrm{d}x. \qquad (20)
\end{aligned}$$

When $F(x) \propto x^{\beta}$, $F^{-1}(x) \propto x^{1/\beta}$,and we have $\delta = n\frac{\beta}{\beta+1}\left(\frac{k}{n}\right)^{\frac{\beta+1}{\beta}}$, which gives $k \propto n^{\frac{1}{\beta+1}}$ by solving for k.

References

[1] M. Aloraini, M. Sharifzadeh, and D. Schonfeld. Quantized gaussian JPEG steganography and pool steganalysis. *IEEE Access*, 10:38031–38044, 2022.

[2] P. Bas. Steganography via cover-source switching. In *2016 IEEE International Workshop on Information Forensics and Security (WIFS)*, pages 1–6, December 4-7 2016.

[3] P. Bas, T. Filler, and T. Pevný. Break our steganographic system – the ins and outs of organizing BOSS. In T. Filler, T. Pevný, A. Ker, and S. Craver, editors, *Information Hiding, 13th International Conference*, volume 6958 of Lecture Notes in Computer Science, pages 59–70, Prague, Czech Republic, May 18–20, 2011.

[4] R. Böhme. *Improved Statistical Steganalysis Using Models of Heterogeneous Cover Signals.* PhD thesis, Faculty of Computer Science, Technische Universität Dresden, Germany, 2008.

[5] M. Boroumand, M. Chen, and J. Fridrich. Deep residual network for steganalysis of digital images. *IEEE Transactions on Information Forensics and Security*, 14(5):1181–1193, May 2019.

[6] J. Butora, Y. Yousfi, and J. Fridrich. How to pretrain for steganalysis. In D. Borghys and P. Bas, editors, *The 9th ACM Workshop on Information Hiding and Multimedia Security*, Brussels, Belgium, 2021. ACM Press.

[7] R. Cogranne, Q. Giboulot, and P. Bas. ALASKA–2: Challenging academic research on steganalysis with realistic images. In *IEEE International Workshop on Information Forensics and Security*, New York, NY, December 6–11, 2020.

[8] E. Dworetzky and J. Fridrich. Explaining the bag gain in batch steganography. *IEEE Transactions on Information Forensics and Security*, 18:3031–3043, 2023.

[9] Q. Giboulot, P. Bas, and R. Cogranne. Multivariate side-informed Gaussian embedding minimizing statistical detectability. *IEEE Transactions on Information Forensics and Security*, 17:1841–1854, 2022.

[10] Q. Giboulot, R. Cogranne, and P. Bas. Detectability-based JPEG steganography modeling the processing pipeline: The noise-content trade-off. *IEEE Transactions on Information Forensics and Security*, 16:2202–2217, 2021.

[11] Q. Giboulot and J. Fridrich. Payload scaling for adap-

tive steganography: An empirical study. *IEEE Signal Processing Letters*, 26(9):1339–1343, July 2019.

[12] V. Holub, J. Fridrich, and T. Denemark. Universal distortion design for steganography in an arbitrary domain. *EURASIP Journal on Information Security, Special Issue on Revised Selected Papers of the 1st ACM IH and MMS Workshop*, 2014:1, 2014.

[13] J. R. Janesick. *Scientific Charge-Coupled Devices*, volume Monograph PM83. Washington, DC: SPIE Press - The International Society for Optical Engineering, January 2001.

[14] A. D. Ker. Batch steganography and pooled steganalysis. In J. L. Camenisch, C. S. Collberg, N. F. Johnson, and P. Sallee, editors, *Information Hiding, 8th International Workshop*, volume 4437 of Lecture Notes in Computer Science, pages 265–281, Alexandria, VA, July 10–12, 2006. Springer-Verlag, New York.

[15] A. D. Ker. The square root law in stegosystems with imperfect information. In R. Böhme and R. Safavi-Naini, editors, *Information Hiding, 12th International Conference*, volume 6387 of Lecture Notes in Computer Science, pages 145–160, Calgary, Canada, June 28–30, 2010. Springer-Verlag, New York.

[16] A. D. Ker. On the relationship between embedding costs and steganographic capacity. In M. Stamm, M. Kirchner, and S. Voloshynovskiy, editors, *The 5th ACM Workshop on Information Hiding and Multimedia Security*, Philadelphia, PA, June 20–22, 2017. ACM Press.

[17] A. D. Ker. The square root law of steganography. In M. Stamm, M. Kirchner, and S. Voloshynovskiy, editors, *The 5th ACM Workshop on Information Hiding and Multimedia Security*, Philadelphia, PA, June 20–22, 2017. ACM Press.

[18] A. D. Ker and Tomas Pevný. Batch steganography in the real world. In J. Dittmann, S. Craver, and S. Katzenbeisser, editors, *Proceedings of the 14th ACM Multimedia & Security Workshop*, pages 1–10, Coventry, UK, September 6–7, 2012.

[19] B. Li, M. Wang, and J. Huang. A new cost function for spatial image steganography. In *Proceedings IEEE, International Conference on Image Processing, ICIP*, Paris, France, October 27–30, 2014.

[20] T. Mingxing and V. L. Quoc. EfficientNet: Rethinking model scaling for convolutional neural networks. In *Proceedings of the 36th International Conference on Machine Learning (ICML)*, volume 97, pages 6105–6114, June 9–15, 2019.

[21] T. Pevný and I. Nikolaev. Optimizing pooling function for pooled steganalysis. In *IEEE International Workshop on Information Forensics and Security*, pages 1–6, Rome, Italy, November 16–19, 2015.

[22] V. Sedighi, R. Cogranne, and J. Fridrich. Content-adaptive steganography by minimizing statistical detectability. *IEEE Transactions on Information Forensics and Security*, 11(2):221–234, 2016.

[23] V. Sedighi, R. Cogranne, and J. Fridrich. Practical strategies for content-adaptive batch steganography and pooled steganalysis. In *Proceedings IEEE, International Conference on Acoustics, Speech, and Signal Processing*, March 5–9, 2017.

[24] Y. Yousfi, J. Butora, E. Khvedchenya, and J. Fridrich. ImageNet pre-trained CNNs for JPEG steganalysis. In *IEEE International Workshop on Information Forensics and Security*, New York, NY, December 6–11, 2020.

[25] Y. Yousfi, E. Dworetzky, and J. Fridrich. Detector-informed batch steganography and pooled steganalysis. In J. Butora, B. Tondi, and C. Veilhauer, editors, *The 10th ACM Workshop on Information Hiding and Multimedia Security*, Santa Barbara, CA, 2022. ACM Press.

[26] A. Zakaria, M. Chaumont, and G. Subsol. Pooled steganalysis in JPEG: how to deal with the spreading strategy? In *IEEE International Workshop on Information Forensics and Security (WIFS)*, pages 1–6, 2019.

Author Biography

Eli Dworetzky is currently pursuing his PhD in electrical and computer engineering at Binghamton University. His research currently focuses on image steganography and steganalysis. He received an MS in computer engineering from Binghamton University in 2021.

Edgar Kaziakhmedov received M.S. degree in Applied Mathematics and Physics from Moscow Institute of Physics and Technology, Moscow, in 2020. He is currently pursuing PhD degree in electrical engineering at Binghamton University. His research areas lie within digital image steganalysis and steganography, neural network based image processing and digital media forensics.

Jessica Fridrich is Distinguished Professor of Electrical and Computer Engineering at Binghamton University. She received her PhD in Systems Science from Binghamton University in 1995 and MS in Applied Mathematics from Czech Technical University in Prague in 1987. Her main interests are in steganography, steganalysis, and digital image forensics. Since 1995, she has received 23 research grants totaling over $13 mil that lead to more than 230 papers and 7 US patents.