

Adaptive Steganography in the Noisy Channel with Dual-Syndrome Trellis Codes

Christy Kin-Cleaves and Andrew D. Ker
Department of Computer Science, Oxford University
Oxford, United Kingdom

Christy.Kin-Cleaves@cs.ox.ac.uk, adk@cs.ox.ac.uk

Abstract

Adaptive steganography aims to reduce distortion in the embedding process, typically using Syndrome Trellis Codes (STCs). However, in the case of non-adversarial noise, these are a bad choice: syndrome codes are fragile by design, amplifying the channel error rate into unacceptably-high payload error rates. In this paper we examine the fragility of STCs in the noisy channel, and consider how this can be mitigated if their use cannot be avoided altogether. We also propose an extension called Dual-Syndrome Trellis Codes, that combines error correction and embedding in the same Viterbi process, which slightly outperforms a straightforward combination of standard forward error correction and STCs.

1. Introduction

Steganography, unlike digital watermarking, does not typically strive for robustness. The adversary is a passive detector, and the communication channel noiseless [5]. This models real-world problems of communicating a covert message inside, for example, an image shared on a social media platform.

However, some image-sharing websites and social media platforms recompress uploaded images, or add visible watermarks. This introduces a non-adversarial element of noise into the channel. Unfortunately even a small amount of noise is often catastrophic for so-called *adaptive* algorithms [9], which are the current state-of-art in steganography. Such algorithms employ Syndrome Trellis Codes (STCs) [4, 3], which carry the payload \mathbf{m} as the syndrome of a block-structured parity-check matrix \mathbb{H} . The STC algorithm allows the embedder to minimize an additive distortion heuristic while solving $\mathbf{m} = \mathbb{H}\mathbf{y}$, where \mathbf{y} represents the binary or ternary remainder of the pixels or DCT coefficients in the stego object. Unless \mathbb{H} is extremely sparse, one error in the channel causes many errors in the syndrome, so

that even small amounts of noise in the channel lead to high bit error rates in the payload.

To illustrate this phenomenon, we measured the error rate of the Least Significant Bit (LSB) of DCT coefficients in JPEG files that have been decompressed and recompressed with a matching quality factor and identical JPEG library. This represents the least-damaging form of recompression. Although most of the JPEG process is invertible, the final decompression step of clamping and rounding pixel values means that compression is not an idempotent operation. Errors are more numerous in JPEGs with a higher quality factor where the quantization is finer, as rounding noise is more likely to tip a coefficient into the next quantization bin.

In Fig. 1 we show the bit error rate induced by this recompression, for JPEG quality factors between 50 and 100, measured on the 512×512 images in the BOSSBase [1] set. We also show the bit error rate for payload if STCs are used to embed a payloads of 26214 and 52429 bits (0.1 and 0.2 bits per pixel, respectively). The constraint height of the STC was $h = 8$; as we shall see later, this is a major factor in the fragility of STC embedding. Note also that the *shorter* payload is more fragile. However, there are few alternatives to STCs if we want to reduce distortion.

In this paper we examine the relationship between the error rate of a binary symmetric channel (BSC) and the error rate in the syndrome, and examine methods for error correction in adaptive steganographic algorithms. We emphasise that, unlike digital watermarking methods, we consider only non-adversarial noise. There is relatively little literature on this problem and we show that it has unexpected lessons for adaptive steganography: where possible, STCs should be *avoided* when the channel is noisy, but if distortion must be reduced then there are ways to mitigate STC fragility: lower the constraint height and parity-check density, and use only as much cover as necessary to meet a distortion constraint.

Later in the paper we introduce Dual-Syndrome Trellis Codes (DSTC) as an extension to the STC algorithm. One syndrome is used for error correction, in much the same

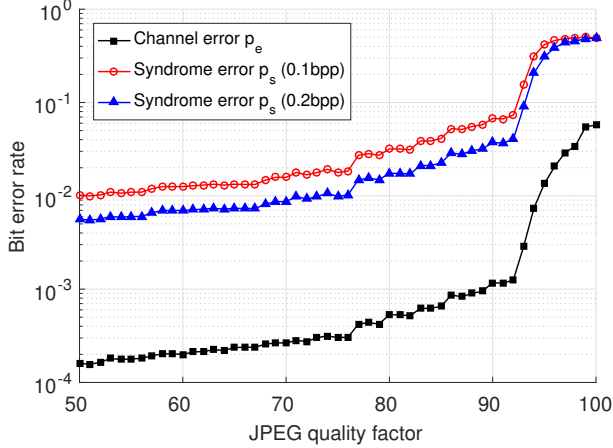


Figure 1. Bit error rates p_e for coefficient LSBs in recompressed JPEGs, as the quality factor varies, and corresponding error rates of adaptive steganography.

way as convolutional codes, and the other for payload. By modifying the trellis, the Viterbi algorithm can be used for both embedding and decoding.

The outline of the paper is as follows. In section 2 we survey existing literature on robust steganography. In section 3 we explore how parameters such as cover size, payload size, and constraint height affect bit error rates of payload when embedding using STCs; we conclude that STCs should be avoided if possible, but that it will not always be possible. In section 4 we introduce DSTCs; there is no comparable prior art in the literature, but we benchmark their performance against a method that uses forward error correction (FEC) inside the syndrome, in section 5. We conclude in section 6.

1.1. Notation

We shall denote vectors as lowercase boldface \mathbf{x} , with the i^{th} element of \mathbf{x} denoted x_i . Matrices are written in uppercase double struck: \mathbb{H} , indexed as \mathbb{H}_{ij} . We will only consider binary payloads and codes, and all vector arithmetic will be mod 2.

2. Prior art

There is a wealth of literature on digital watermarking [2], aiming for robustness against an active attacker, but there has been little written about steganography robust to non-adversarial noise, and almost all predates modern adaptive embedding.

Marvel [11] builds on an early type of steganography that requires the receiver to know the cover object. In their system, called ‘SSIS’, the receiver estimates the cover from the stego object using image restoration techniques. Within their hiding mechanism, low-rate Reed Solomon (RS) codes are used to correct the errors of imperfect restoration. This

is not robust steganography as such, but it illustrates the most basic idea of applying forward error correction to a payload before embedding it into a channel that might be noisy. A similar architecture is employed by ‘YASS’ [14], where noise is introduced by using ‘big blocks’ that do not align with JPEG blocks. The errors are corrected using a Repeat-Accumulate (RA) code.

Zhang and Wang [16] use cyclic codes to both improve embedding efficiency of, and to introduce error correction to, the resulting stego objects. Predating the technology of STCs, the distortion heuristic is that all changes have equal cost. The authors make clear that there is a trade-off between embedding rate and robustness, because increasing the embedding rate forces codewords to be closer together. Ref. [13] proposes a similar idea with a distortion heuristic that is not constant: some blocks are not used for embedding, in which case their symbols are erased. Either RS or RA codes are used to remove errors caused by the selective avoidance of high-cost blocks.

Separately from the practical applications of steganography, Hopper [8] formalise a complexity-theoretic approach to steganography. This work is most famous for its analysis of the *rejection sampler*, a low-bandwidth but provably perfectly-secure steganography scheme; less famously, the same paper also proposes a modification to work in noisy channels. The authors identify an upper-bound on the steganographic capacity of noisy channels, but the abstract model is far from the practice of steganography in digital images. Rejection sampling is also computationally expensive, so we do not consider it further.

The most recent work on steganography in noisy channels is [12]. Published just prior to the introduction of STCs, it uses syndrome coding with the parity-check matrix of a (7, 4) Hamming code. Like YASS, some blocks are avoided; furthermore, the matrix embedding allows the number of embedding changes to be reduced. This is equivalent to costs of 1 or ∞ . To reduce errors, RA codes are using *inside the syndrome*, i.e. forward error correction is applied prior to embedding. A soft decoder is used.

3. The robustness/distortion trade-off

We restrict our attention to binary codes. Let \mathbf{m} be the message, of l bits, to be hidden in cover \mathbf{x} of n LSBs. This is embedded using adaptive steganography to create a stego object whose LSBs are \mathbf{y} of length n . After transmission, the stego object is corrupted to \mathbf{y}' , decoded as message \mathbf{m}' .

Let us examine why the bit error rate of STC adaptive steganography is so high in Fig. 1. According to the STC algorithm, $\mathbf{m} = \mathbb{H}\mathbf{y}$ and $\mathbf{m}' = \mathbb{H}\mathbf{y}'$ ¹.

Suppose that the errors in \mathbf{y}' are due to a BSC with error

¹Recall that, with STCs, \mathbb{H} is constructed from copies of $\hat{\mathbb{H}}$ as in the top part of Fig. 5.

rate p_e ². We define p_s , the *syndrome error rate*, as the bit error rate of $\mathbb{H}\mathbf{y}$ and derive it in terms of p_e , l , n , and the average number of 1s in each column of \mathbb{H} , denoted a .

There is a elementary argument as follows. On average there will be np_e errors in \mathbf{y}' . Suppose that one error affects bit i of \mathbf{y} ; this adds column i of \mathbb{H} to \mathbf{m} , which corrupts on average a bits of \mathbf{m} . Although some errors cancel out, a first-order approximation to the number of errors in \mathbf{m}' is therefore anp_e , and so the error rate of the syndrome is

$$p_s = p_e \frac{an}{l} + O(p_e^2). \quad (1)$$

A more sophisticated argument assumes that each entry of \mathbb{H}_{ij} are iid Bernoulli with parameter a/l (we ignore the block structure that is present in STCs). The probability that entry i, j both contains a 1 and was effected by an error is $q = ap_e/l$. Thus the probability that $m'_j \neq m_j$ is the probability that bit j is subject to an odd number of flips, which is the probability that $X \sim \text{Bi}(n, q)$ is odd. Expanding $(q \pm (1-q))^n$ using the binomial theorem, this probability is $\frac{1}{2}(1 - (1-2q)^n)$, so

$$p_s = \frac{1 - (1 - 2ap_e/l)^n}{2} = p_e \frac{an}{l} + O(p_e^2). \quad (2)$$

Setting aside other error-correction techniques, this suggests three elementary ways to increase robustness of adaptive syndrome coding, all of which decrease p_s : reduce n (do not use the entire cover); reduce a (decrease the constraint height of the STC); increase l (pad out the message in some way).

3.1. Benchmarking the trade-off

For the experiments in this paper, we used 1000 of the 512×512 grayscale images from BOSSBase [1], JPEG compressed with quality factor 75. Broadly informed by Fig. 1, we will simulate a noisy channel by flipping LSBs of the stego DCT coefficients at error rate $p_e = 10^{-3}$. For adaptive embedding, we use the J-UNIWARD [7] costs, and we simulate payloads of length $l \in \{2000, 4000, 8000, 12000, \dots, 40000\}$ bits.

We will examine the robustness/distortion trade-off empirically, displaying the results in plots such as Fig. 2: on log scales, average total J-UNIWARD cost (distortion) on the x -axis and the bit error rate of the syndrome p_s (later the corrected bit error rate p_b) on the y -axis. For various STC parameters, later augmented by forward error correction, we plot the average distortion $D(\mathbf{x}, \mathbf{y})$ and average error rate for each payload; naturally, increasing payloads

cause more distortion. The best robustness/distortion trade-off is found with a graph to the left and below the others, but we must also take into account the absolute amount of distortion (a low bit error rate is not secure if the embedding is easily detectable).

To relate distortion to detectability, we computed the average distortion costs of payloads corresponding to 0.1 and 0.4 bits per nonzero DCT coefficient, when embedding using binary STCs. These are 8.4×10^5 and 6.8×10^6 , respectively. We plot these vertical lines on each distortion-error chart. In [6], on BOSSBase images, Holub found that these payload rates were detectable by JRM features [10] with error rates of 0.4632 and 0.2376, respectively. This suggests that average distortion of 10^7 is almost certainly too high, and below 10^6 probably sufficiently low for mostly-undetectable embedding.

3.2. Using a subset of the cover

Equation (1) shows that the syndrome error rate can be reduced by reducing n ; the STC is applied to only $n' \leq n$ locations of the cover (and the other $n - n'$ left unchanged). In order to convey the entire message we require $n' \geq l$; in the extreme case $n' = l$ the STC becomes trivial and we resort to an operation such as LSB Matching. Lowering n' , we would expect to improve the robustness/distortion trade-off by reducing fragility. We parameterize this by

$$n' = l + r(n - l), \quad r \in [0, 1], \quad (3)$$

and plot the distortion-robustness trade-off in Fig. 2. Here the constraint height of the STC is $h = 1$: see the following section for why we made this choice. The line $r = 0$ is below the others: the clear message is that *for best robustness/distortion trade-off, do not use STCs at all*. However, not using STCs does have a distortion penalty, and for the case $r = 0$ even the smallest message of 2000 bits causes distortion of approximately $3 \cdot 10^6$, which is probably detectable by existing methods. If the embedder is constrained by both payload length and distortion, they *must* use an adaptive method to reduce the latter.

3.3. Reducing the constraint height

The number of ones in a column of \mathbb{H} is determined by the constraint height and the density of the generator matrix $\hat{\mathbb{H}}$, which we denote η . In the notation of Eq. 1, $a = h\eta$.

To examine these parameters, we tested STCs with constraint heights $h \in \{1, 2, 3, 5, 8\}$, with $\eta \in \{0.25, 0.5, 0.75\}$ for $h = 1$, and the same other parameters as for Fig. 2, showing the results in Fig. 3. Eq. 1 predicts that reducing h or η results in lower p_s ; our understanding of STCs as approaching random linear codes as $h \rightarrow \infty$, and zero columns of $\hat{\mathbb{H}}$ as unused locations in the cover, predicts that reducing h or η results in higher distortion. This is what we observe, and furthermore that reducing h or η is a

²Errors in digital media might occur in bursts, but we can mitigate this by visiting the cover and stego objects in a keyed pseudorandom order. Later, we will also assume that errors in the syndrome are also independent, as the syndrome can be visited in a random order.

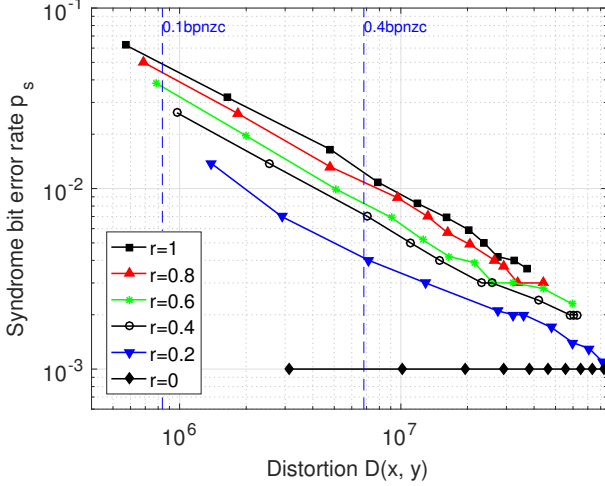


Figure 2. p_s vs $D(x, y)$ for varying values of r .

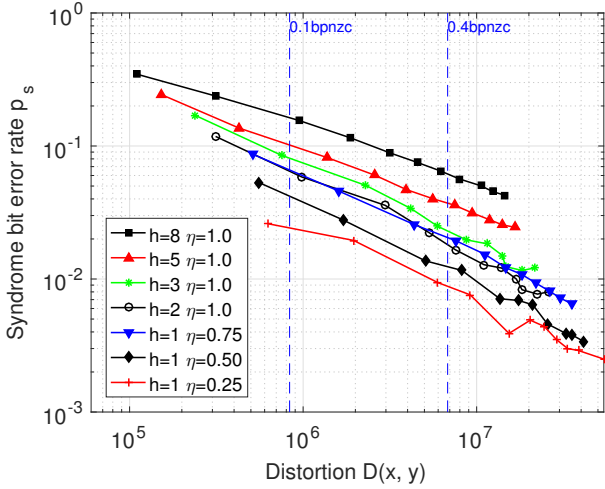


Figure 3. p_s vs $D(x, y)$ for varying values of h and η .

better trade-off between distortion and syndrome error rate (in that error rates are lower for each corresponding distortion), but if detectability is a concern then there is a limit to the embedder's ability to thin out the matrix \mathbb{H} .

3.4. Increasing the message length

We have determined that the most robust STCs are used as little as possible, and with a constraint height of $h = 1$ so that they become block codes (the trellis is trivial).

The final way to reduce p_s , from (1), is to increase l . Like the other options, this inevitably increases distortion. A simple method is to pad the payload with random garbage, but if the message is to be expanded it makes sense to use the extra bits for forward error correction. We now introduce \tilde{m} of length $\tilde{l} \geq l$ as the result of $\text{FEC}_{\text{Encode}}$: some forward error correction encoding function on the message m . In Fig 4 we show how a practical stego system using forward error correction could work, where forward error

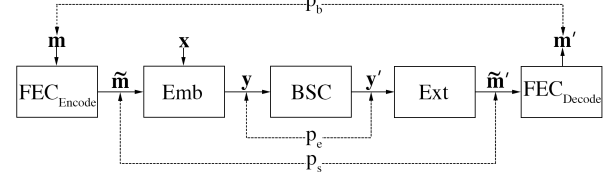


Figure 4. Notation for layering FEC within steganography.

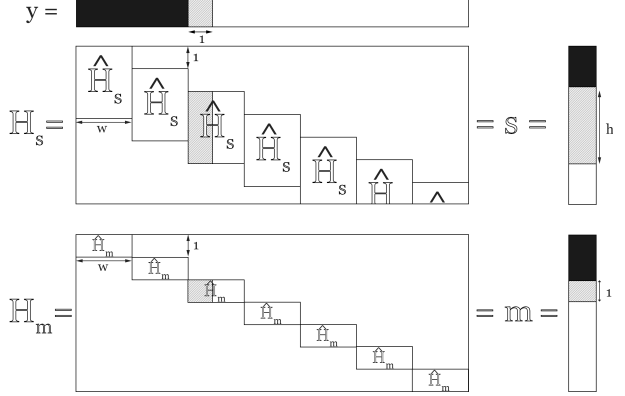


Figure 5. Partial DSTC encoding, darkened regions showing the set portion of y , shaded regions showing how a single cover value can affect the result of h coefficients in s the FEC syndrome, and 1 coefficient in m the message syndrome.

correction pads the message. As well as the syndrome error rate p_s , we also have p_b , the message bit error rate after error correction. We reiterate that error correction inside syndrome has two advantages: the decoding process removes some errors, but also the increased syndrome length *reduces* the syndrome error rate, thanks to (1). We will benchmark one example of FEC (RS codes) in Sect. 5. RS codes are a maximum distance separable code, a (n, k) RS code with n information bits, and k parity bits can correct $(n - k)/2$ errors. Therefore, as k increases, the length of the RS code increases, as does the error-correcting performance [15].

4. Dual-Syndrome Trellis Codes

We propose to develop STCs into *Dual-Syndrome* Trellis Codes, that solve for two syndromes simultaneously. We formalise the embedding as

$$\text{Emb}(x, m) = \underset{\substack{\mathbb{H}_s y = s \\ \mathbb{H}_m y = m}}{\text{argmin}} D(x, y), \quad (4)$$

where s is the shared syndrome of length l , \mathbb{H}_s and \mathbb{H}_m are matrices generated by tiling sub-matrices $\hat{\mathbb{H}}_s$, and $\hat{\mathbb{H}}_m$ respectively.

We define $\hat{\mathbb{H}}_m \in \{0, 1\}^w$ as the $1 \times w$ vector, used to generate \mathbb{H}_m , containing η_m ones on average in each column. $\hat{\mathbb{H}}_s$ is the $h \times w$ matrix used to generate \mathbb{H}_s , containing η_s ones on average in each columns. As the matrix $\hat{\mathbb{H}}_m$ is of height one, the value of m_i is the mod2 accumulation of

$\eta_m \backslash \eta_s$	$p_s \times 10^{-2}$				$D(\mathbf{x}, \mathbf{y}) \times 10^6$			
	0.2	0.4	0.6	0.8	0.2	0.4	0.6	0.8
0.2	0.19	0.22	0.24	0.40	36.66	36.21	36.35	46.00
0.4	0.26	0.24	0.28	0.35	26.20	31.38	31.30	39.15
0.6	0.44	0.41	0.43	0.39	18.35	19.40	20.69	28.30
0.8	0.57	0.57	0.41	0.39	17.37	18.37	20.75	29.19

Figure 6. p_s and $D(\mathbf{x}, \mathbf{y})$ for varying values of η_m , and η_s .

the w columns of $\hat{\mathbb{H}}_m$ multiplied by the respective values of \mathbf{y} . To encode this, we increase the number of states from 2^h to 2^{h+1} , allowing each state to have a representation for the partial syndrome setting the m_i to 0 or 1. Thus, the h LSBs of the state label represent the syndromes equivalence for setting the value of s_i , and the MSB represents the value of m_i . Upon coding w coefficients we prune, only accepting states which LSB matches the shared syndrome s_i value, as this value will not change any more, and the MSB matching the value of the secret message m_i .

As shown in (4) \mathbf{y} is encoded to match both the shared syndrome \mathbf{s} , and secret message \mathbf{m} . However due to the noisy channel Bob receives \mathbf{y}' , the perturbed stego object. To correct \mathbf{y}' we use an STC to solve $\mathbb{H}_s \mathbf{y}' = \mathbf{s}$. By using the hamming distance distortion function $D_H(\mathbf{y}', \mathbf{y}'')$, we are using the Viterbi algorithm to find the closest code-word \mathbf{y}'' , which has the correct syndrome \mathbf{s} :

$$\mathbf{y}'' = \underset{\mathbb{H}_s \mathbf{y}'' = \mathbf{s}}{\operatorname{argmin}} D_H(\mathbf{y}', \mathbf{y}''), \quad (5)$$

we then get the corrected secret message $\mathbf{m}' = \mathbb{H}_m \mathbf{y}''$.

4.1. Finding optimal values of η_s , η_m , \mathbf{s} , and h

In section 3.3 we showed that increasing the sparsity of the \mathbb{H} reduces p_s at a cost of increasing the distortion. In DSTCs we have two independent matrices \mathbb{H}_s and \mathbb{H}_m . To find the optimal values of the sparsity of $\hat{\mathbb{H}}_s$, and $\hat{\mathbb{H}}_m$, we performed DSTC embedding with η_s and $\eta_m \in \{0.2, \dots, 0.8\}$, using the same cover set and channel error as previous experiments. Syndrome error rates and distortion are displayed in Fig. 6, from this we selected $\eta_m = 0.4$, and $\eta_s = 0.2$.

During our experiments, we observed a curious phenomenon: in cover objects, syndromes are non-uniform, often favouring zeros. We can attribute this to JPEG compression creating more zeros than nonzeros. In traditional STCs, the steganographer has no choice over the syndrome – it is the message – but in DSTCs we could choose \mathbf{s} to require fewest changes. Since \mathbb{H}_s is secret, \mathbf{s} need not be, and it seems safe to choose $\mathbf{s} = \mathbf{0}$, slightly reducing distortion.

DSTCs allow for a constraint height to be used as would be in a regular STC. In practise, we found that although codewords produced with a $h > 1$ had a lower distortion, they had worse correcting performance. When $h = 1$, the codeword that satisfies the constraint (4) is similar to that which solves just $\mathbb{H}_s \mathbf{y} = \mathbf{s}$, therefore, in the correction step

\mathbf{y}' is corrected to \mathbf{y}'' , a codeword close to \mathbf{y} . When $h > 1$, the codeword which satisfies (4) is less similar to the one which satisfies just $\mathbb{H}_s \mathbf{y} = \mathbf{s}$, meaning our correction step is less effective.

5. Performance

As the aims of this paper are to create robust adaptive stego objects, the balance of distortion minimisation and robustness is critical. For our experiments we took the same image set as used earlier and encoded them both with DSTC methods as proposed in section 4, and forward error correction in the syndrome as described in section 3.4.

Using DSTCs, we embedded messages, and RS ($k = 232$) FEC-encoded messages. For comparison we used RS codes ($k \in \{168, 200\}$) before embedding using a traditional STC. We chose the values of k for STCs + RS codes, to be similar in average distortion to our proposed method, so that we could compare the effect of RS codes in the Syndrome against our proposed method. We also included the case $r = 0$ ($l = n$, plain overwriting), with no error correction. All stego objects were subjected to a BSC with $p_e = 10^{-3}$, and STCS used $h = 1$, and $r \in \{0.75, 1.0\}$.

In Figs. 8 and 9 we show how our proposed method is able to generate more robust stego objects per the amount of distortion we must take on, for smaller messages. Considering our findings in Sect. 3, a distortion over 10^7 is most likely very detectable, and for payloads below this threshold our proposed method outperforms STCs + RS codes in the syndrome. In Fig. 7 we see that for longer payloads, STCs + RS codes outperforms our proposed method, but this is not much use if the stego objects are easily detectable. We can attribute this fall off as typical of FEC: as the code length increases, the bit-error rate approaches zero.

6. Conclusion

STCs are the dominant method for adaptive steganography, but they have very poor robustness to non-adversarial noise. We have explored ways to mitigate their fragility, a topic that has received little attention in the literature.

Adding robustness is bound to increase distortion, but robustness and adaptive steganography are a difficult combination: a code with low distortion requires codewords to be close together (so that a stego object can be close to the cover), yet high robustness needs codewords that are far apart. The Dual-Syndrome code attempts to balance these aims, and slightly outperforms the layering of FEC inside an STC with low constraint height, for small messages.

If recompression cannot be avoided (and its erasures cannot be predicted, which is a tacit assumption of this paper), it may be necessary to try adaptive steganography that does not use syndrome codes at all. This is a possible direction for further research.

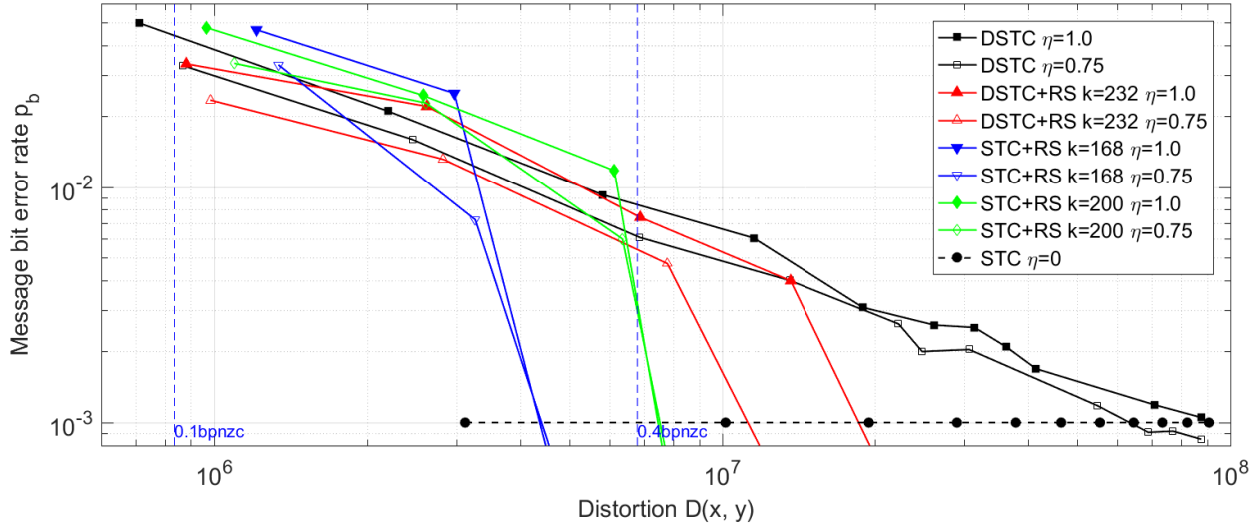


Figure 7. Comparison of all the methods explored in this paper. Average distortion $D(\mathbf{x}, \mathbf{y})$ vs. bit error rate of corrected payload (p_b), for combinations of STC, DSTC, and RS codes, $l \in \{2000, 4000, 8000, 12000, \dots, 40000\}$ (left to right).

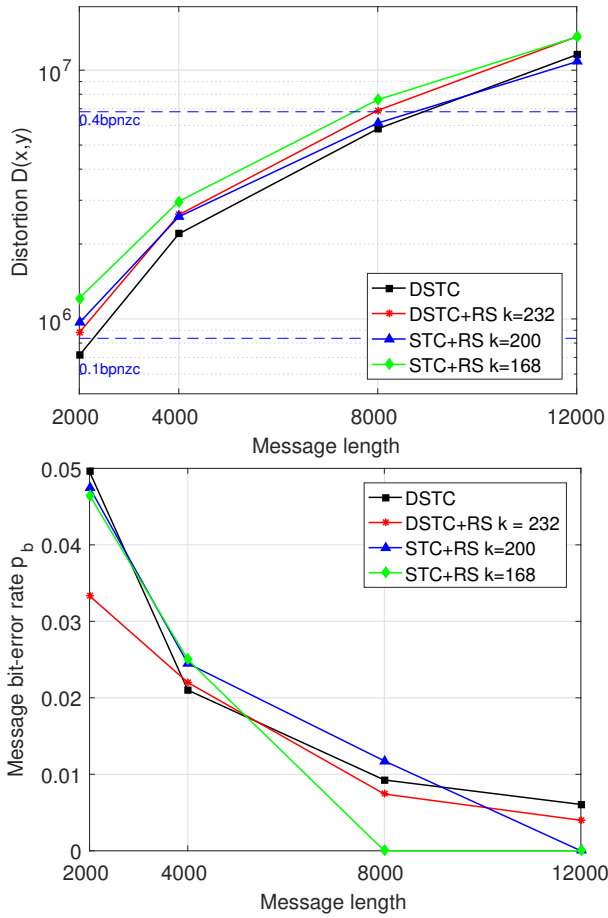


Figure 8. Comparison of results for cover utilisation $r = 1.0$.

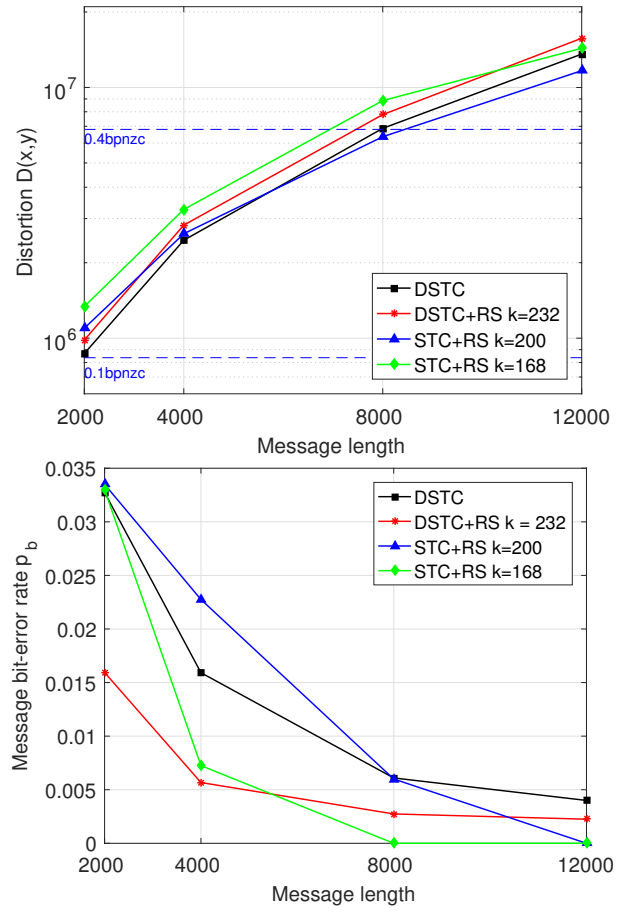


Figure 9. Comparison of results for cover utilisation $r = 0.75$.

References

- [1] P. Bas, T. Filler, and T. Pevny. ‘Break Our Steganographic System’: The Ins and Outs of Organizing BOSS. *Information Hiding*, pages 59–70, 2011. 1, 3
- [2] I. Cox, M. Miller, J. Bloom, J. Fridrich, and T. Kalker. *Digital Watermarking and Steganography*. Morgan Kaufmann, 2007. 2
- [3] T. Filler, J. Judas, and J. Fridrich. Minimizing Embedding Impact in Steganography Using Trellis-Coded Quantization. In *Media Forensics and Security*, page 754105, 2010. 1
- [4] T. Filler, J. Judas, and J. Fridrich. Minimizing Additive Distortion in Steganography Using Syndrome-Trellis Codes. *IEEE Transactions on Information Forensics and Security*, 6(3):920–935, 2011. 1
- [5] J. Fridrich. *Steganography in Digital Media: Principles, Algorithms, and Applications*. Cambridge University Press, 2009. 1
- [6] V. Holub and J. Fridrich. Challenging the Doctrines of JPEG Steganography. In *Media Watermarking, Security, and Forensics*, 2014. 3
- [7] V. Holub, J. Fridrich, and T. Denemark. Universal Distortion Function for Steganography in an Arbitrary Domain. *EURASIP Journal on Information Security*, 2014(1), 2014. 3
- [8] N. Hopper, L. von Ahn, and J. Langford. Provably Secure Steganography. *IEEE Transactions on Computers*, 58(5):662–676, 2009. 2
- [9] A. Ker, P. Bas, R. Böhme, R. Cogranne, S. Craver, T. Filler, J. Fridrich, and T. Pevny. Moving Steganography and Steganalysis From the laboratory into the Real World. In *Proceedings of the first ACM workshop on Information hiding and multimedia security*, pages 45–58. ACM, 2013. 1
- [10] J. Kodovský and J. Fridrich. Steganalysis of JPEG Images Using Rich Models. In *Media Watermarking, Security, and Forensics 2012*, volume 8303, page 83030A. International Society for Optics and Photonics, 2012. 3
- [11] L. Marvel, C. Boncelet, and C. Retter. Spread Spectrum Image Steganography. *IEEE Transactions on image processing*, 8(8):1075–1083, 1999. 2
- [12] A. Sarkar, U. Madhow, and B. Manjunath. Matrix Embedding With Pseudorandom Coefficient Selection and Error Correction for Robust and Secure Steganography. *IEEE Transactions on Information Forensics and Security*, 5(2):225–239, 2010. 2
- [13] K. Solanki, N. Jacobsen, U. Madhow, B. Manjunath, and S. Chandrasekaran. Robust Image-Adaptive Data Hiding Using Erasure and Error Correction. *IEEE Transactions on image processing*, 13(12):1627–1639, 2004. 2
- [14] K. Solanki, A. Sarkar, and B. Manjunath. Yass: Yet Another Steganographic Scheme That Resists Blind Steganalysis. In *International Workshop on Information Hiding*, pages 16–31. Springer, 2007. 2
- [15] S. Wicker and V. Bhargava. *Reed-Solomon Codes and Their Applications*. John Wiley & Sons, 1999. 4
- [16] X. Zhang and S. Wang. Stego-encoding with Error Correction Capability. *IEICE Transactions on Fundamentals*

of Electronics, Communications and Computer Sciences, 88(12):3663–3667, 2005. 2