Steganalysis of Embedding in Two Least-Significant Bits

Andrew D. Ker, *Member, IEEE*

***Abstract*—This paper proposes steganalysis methods for extensions of least-significant bit (LSB) overwriting to both of the two lowest bit planes in digital images: there are two distinct embedding paradigms. The author investigates how detectors for standard LSB replacement can be adapted to such embedding, and how the methods of "structural steganalysis," which gives the most sensitive detectors for standard LSB replacement, may be extended and applied to make more sensitive purpose-built detectors for two bit plane steganography. The literature contains only one other detector specialized to detect replacement multiple bits, and those presented here are substantially more sensitive. The author also compares the detectability of standard LSB embedding with the two methods of embedding in the lower two bit planes: although the novel detectors have a high accuracy from the steganographer's point of view, the empirical results indicate that embedding in the two lowest bit planes is preferable (in some cases, highly preferable) to embedding in one.**

***Index Terms*—Steganography, structural steganalysis, two least-significant bit (LSB) embedding.**

I. Introduction

REPLACEMENT of least-significant bits (LSBs) in digital images is an extremely simple form of information hiding. For the nonexpert steganographer, its ease of embedding, high capacity, and visual imperceptibility may prove attractive. However, it is now known that there are particular flaws which make steganalysis (detection) of this embedding method much easier than that of other additive steganography.

The aim of this paper is to consider the extension to replacement of the two LSBs. Such embedding is still visually imperceptible, of even higher capacity, and still extremely simple. But there exist parallel "structural" weaknesses of such embedding, which allows us to extend the most sensitive detectors for LSB replacement to detect embedding in two bit planes; we will develop and benchmark such detectors. One might ask why a steganographer would want to extend the weak LSB embedding method to more bit planes. It will be shown that, at least as far as the detectors presented here are concerned, it is actually somewhat better (harder to detect) to embed in two bit planes than in one. Therefore, if one must embed by replacement of bits (for example, if steganography software is not available; see [1] for a commandline LSB replacement program and some explanation of scenarios in which LSB embedding may be necessary), it would be better to use two bit planes than one.

Manuscript received July 4, 2006; revised September 29, 2006. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Jessica J. Fridrich.

The author is with the Computing Laboratory, Oxford University, Oxford OX1 3QD, U.K. (e-mail: adk@comlab.ox.ac.uk).

Digital Object Identifier 10.1109/TIFS.2006.890519

The format of this paper is as follows. Next, we shall describe methods for embedding in the least and two least-significant bit planes. In Section II, we recap the structural framework of [2], presenting a modification of the standard LSB detector from [3] in a form suitable for extension to multiple bit-plane replacement. Methods for application of conventional LSB replacement detectors to the detection of replacement of more than one bit plane are outlined in Section III. Section IV explains the extended structural framework and derives novel detectors specialized for replacement of two bit planes. The new detectors are benchmarked in Section V and conclusions appear in Section VI.

A. Embedding in Two LSBs

There are two obvious ways to embed a payload by overwriting the least and second-least significant bits of a cover. For definiteness, we also describe the standard LSB embedding method, which is folklore to steganographers. Throughout this paper, we assume that the cover object has N bytes (more generally, words of some fixed bit length) and that the embedded payload is of proportionate length p, with $0 \leq p \leq 1$, where the proportion is of the available capacity.

For embedding in the LSB only, Np bits are embedded in the cover by selecting Np pixels and replacing the LSB of each pixel by the corresponding bit of the payload. The author will abbreviate this embedding method as LSB embedding.

For embedding in the two LSBs, $2Np$ bits are embedded in the cover by selecting Np pixels and replacing both of the two LSBs of each pixel with two corresponding bits of the payload. The author will abbreviate this as 2LSB embedding.

As an alternative method of using two LSBs, $2Np$ bits can be embedded in the cover by selecting Np pixels and replacing only the second-LSBs of each pixel with a corresponding bits of the payload, then repeating with a new selection of pixels of which only the LSB is used. Therefore, changes occur in the least and second-LSB planes independently. The author will abbreviate this as I2LSB embedding (the I signifying the independence of the effects on the two lowest bit planes).

In each case, the selection of pixels and order of embedding is generated from a secret key shared by the recipient. They can therefore reconstruct the pixels used and recover the payload by reading off the relevant bits from the stego object.

Each method has its advantages and disadvantages. The standard LSB embedding can only embed half as much data as the others, but the distortion to the cover is limited to pixel value changes of, at most, one, with on average $Np/2$ pixels changed. The 2LSB and I2LSB embedding methods carry twice the payload of plain LSB embedding and the 2LSB method changes fewer pixels than I2LSB at the cost of the average distortion

TABLE I
EMBEDDED PAYLOAD AND STEGO NOISE PROFILE FOR THE METHODS OF EMBEDDING IN LEAST AND TWO LSBS

	payload (bits)	stego noise probability 0	± 1	± 2	± 3
LSB	Np	$1-\frac{p}{2}$	$\frac{p}{4}$	0	0
2LSB	$2Np$	$1-\frac{3p}{4}$	$\frac{3p}{16}$	$\frac{p}{8}$	$\frac{p}{16}$
I2LSB	$2Np$	$1-p+\frac{p^2}{4}$	$\frac{3p}{4}-\frac{p^2}{16}$	$\frac{p}{4}-\frac{p^2}{8}$	$\frac{p^2}{16}$

being higher; both 2LSB and I2LSB embedding change more pixels than standard LSB embedding. The stego noise profile of each embedding method is summarized in Table I. It is easy to verify that stego noise at this level remains quite imperceptible visually, even under maximum embedding, as long as the cover is not pathological.

II. STRUCTURAL STEGANALYSIS OF LSB EMBEDDING REVIEWED

LSB replacement is a weak form of steganography because of the structure it contains: when replacing an LSB, an even cover value may be incremented or unchanged, but never decremented; conversely for an odd cover value. There is no detection power in this property when individual pixels are considered, but the same property in pairs of pixels does give quite sensitive detectors for the presence of even small quantities of data hidden by LSB replacement. The first detectors making implicit use of such properties include [4] and [5]. The first work to describe the structure explicitly was [6], with an improved method appearing in [3]. Subsequently, the structural property was fully generalized to groups of two or more pixels, in [2] (with [7] explaining some difficulties in an application to groups of size four); this also included a clean exposition involving matrices, more easily generalizable than [6]. (Reference [8] also contains a similar exposition, but a different treatment of cover assumptions.) The exposition format of [2] will be used to re-present the method of [3] in a clear way suitable for extension; some of the notation must be slightly altered.

Although steganography detectors that will detect embedding in multiple bit planes as a side-effect of their intended application already exist, such as [9], they are very weak when compared with structural detectors. To our knowledge, the literature contains only one other statistical detector specifically designed for multiple bit replacement, which appears in [10]; this is developed by extending the steganalysis method now known as "WS" [11], which is the only LSB replacement detector not easily placed into the structural framework of [2] which has nearly as good performance. As such, the detector of [10] is quite different than those presented here, and makes different cover assumptions; it also works for embedding in more than two bit planes, but does not accurately detect I2LSB embedding. Its performance will be examined briefly, in Section V where it is shown to be weaker than the structural detectors, except for near-maximal embedded payloads.

The outline of the structural detectors, suggested in [2], is as follows. They are quantitative in that they estimate the proportionate length of payload. First, we define a macroscopic property of stego images which depends on the proportionate amount of hidden data p, a vector $\mathbf{B}(p)$; we will derive how $\mathbf{B}(p)$ depends on p and $\mathbf{B}(0)$ and then invert so that, given a stego image, we can hypothesize a value for p and compute what this would imply for $\mathbf{B}(0)$. In the second stage, we give a model for cover images, expressed in terms of $\mathbf{B}(0)$. Then, we can find the value of p which leads to a value of $\mathbf{B}(0)$ closest to the model: this is the estimator for p.

A. Macroscopic Property

As in [2], we will use calligraphic letters $(\mathcal{X})$ for sets, uppercase letters (X) for random variables, and lowercase letters (x) for constants and realizations of random variables. For our purposes, the cover will be fixed, and the payload random. Suppose that a digital image consists of a series of N samples with values $s_1, s_2, \ldots, s_N$ in the range $0 \ldots 2M+1$ (typically, $M = 127$). A sample pair is a pair of sample locations (j,k) for some $1 \le j \ne k \le N$. Let $\mathcal{P}$ be a set of sample pairs; we will use the set of all pairs of horizontally or vertically adjacent pixels (as in [6]). When the image is in color, we pool the sample pairs from each color channel. Consider some subsets of $\mathcal{P}$

$$\mathcal{C}_m = \{(j,k) \in \mathcal{P} \,|\, \lfloor s_k/2 \rfloor = \lfloor s_j/2 \rfloor + m\}$$
$$\mathcal{B}^0_m = \{(j,k) \in \mathcal{P} \,|\, s_k = s_j + m, \text{with } s_j \text{ even}\}$$
$$\mathcal{B}^1_m = \{(j,k) \in \mathcal{P} \,|\, s_k = s_j + m, \text{with } s_j \text{ odd}\}$$

in the first of these, $-M \le m \le M$; for the second, $-2M \le m \le 2M+1$; and for the third, $-2M+1 \le m \le 2M$.

The sets $\mathcal{C}_m$ we call trace sets: they do not involve the LSBs of the pairs, so any pair in $\mathcal{C}_m$ must remain there after LSB overwriting. The sets $\mathcal{B}^0_m$ and $\mathcal{B}^1_m$ we call trace subsets;[1] it is simple to check that each $\mathcal{C}_m$ is partitioned into $\mathcal{B}^0_{2m}, \mathcal{B}^0_{2m+1}, \mathcal{B}^1_{2m-1}, \mathcal{B}^1_{2m}$, and that LSB replacement moves sample pairs among the four trace subsets of each trace set. In [2], $\mathcal{B}^0_m$ was called $\mathcal{E}_m$ ("even") and $\mathcal{B}^1_m$ was called $\mathcal{O}_m$ ("odd"), but we will later want to extend the breakdown of trace subsets to modulo 4 arithmetic and the notation $\mathcal{B}^0_m$ ("a binary number ending in 0, followed by a number m higher") will do so nicely. Note that these sets are not quite equivalent to those called $\mathcal{X}_m$ and $\mathcal{Y}_m$ used by Dumitrescu *et al.* in [6]: their definition is symmetrical in the order of the pairs but introduces a special case at $m = 0$ which causes complications for extended analysis, as explained in [2].

Now assume that the payload of length pN is of the form of a random bitstream (it suffices to be uncorrelated with the cover) embedded using LSB replacement (the selection of pixels used is also assumed uncorrelated with cover or payload). Suppose that a sample pair lies in trace set $\mathcal{C}_m$. Each sample in the pair is altered independently with probability $(p/2)$, moving the sample pair among the trace subsets of $\mathcal{C}_m$ according to the transition diagram (Fig. 1).

The author counts the size of the trace subsets: let b^i_m represent the number of sample pairs in $\mathcal{B}^i_m$ before embedding, and the random variable B'^i_m be the number after such random embedding (we shall later use b'^i_m for a realization of B'^i_m). Consider, for example, B'^0_{2m}. Sample pairs can be in $\mathcal{B}^0_{2m}$, after embedding, in four ways: either having been in $\mathcal{B}^0_{2m}$ before and

[1] This is a simplification of the terminology of [6] which we introduced in [2].

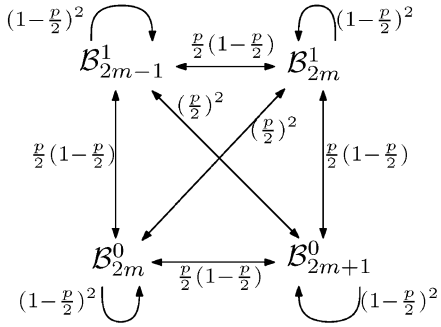

Fig. 1. Transitions between the trace subsets of C_m, when proportion p of maximum payload is embedded by LSB replacement.

remaining there (and, on average, a proportion $(1-(p/2))^2$ of the b^0_{2m} pairs in this position should remain), having been in $\mathcal{B}^0_{2m+1}$ ($p/2(1-(p/2))$ of b^0_{2m+1} will do so), having been in $\mathcal{B}^1_{2m-1}$ before and moving to $\mathcal{B}^0_{2m}$ ($p/2(1-(p/2))$ of the b^1_{2m-1} will do so), or having been in $\mathcal{B}^1_{2m}$ ($(p/2)^2$ of b^1_{2m} will do so). Thus

$$E\left[B'^0_{2m}\right] = \left(1-\frac{p}{2}\right)^2 b^0_{2m} + \frac{p}{2}\left(1-\frac{p}{2}\right) b^0_{2m+1} + \frac{p}{2}\left(1-\frac{p}{2}\right) b^1_{2m-1} + \left(\frac{p}{2}\right)^2 b^1_{2m}.$$

This calculation can be repeated for each $B'^0_{2m+1}, B'^1_{2m-1}, B'^1_{2m}$ to get four linear equations. They can be expressed in vector form if we write $\mathbf{b_m} = (b^0_{2m}, b^0_{2m+1}, b^1_{2m-1}, b^1_{2m})^T$ and similarly $\mathbf{B'_m}$ (and later $\mathbf{b'_m}$ for the realization of $\mathbf{B'_m}$). Then

$$E[\mathbf{B'_m}] = M\mathbf{b_m} \tag{1}$$

where

$$M = \begin{pmatrix} \left(1-\frac{p}{2}\right)^2 & \frac{p}{2}\left(1-\frac{p}{2}\right) & \frac{p}{2}\left(1-\frac{p}{2}\right) & \left(\frac{p}{2}\right)^2 \\ \frac{p}{2}\left(1-\frac{p}{2}\right) & \left(1-\frac{p}{2}\right)^2 & \left(\frac{p}{2}\right)^2 & \frac{p}{2}\left(1-\frac{p}{2}\right) \\ \frac{p}{2}\left(1-\frac{p}{2}\right) & \left(\frac{p}{2}\right)^2 & \left(1-\frac{p}{2}\right)^2 & \frac{p}{2}\left(1-\frac{p}{2}\right) \\ \left(\frac{p}{2}\right)^2 & \frac{p}{2}\left(1-\frac{p}{2}\right) & \frac{p}{2}\left(1-\frac{p}{2}\right) & \left(1-\frac{p}{2}\right)^2 \end{pmatrix}.$$

Observe that M can be written as a Kronecker tensor product $M = P \otimes P$ where

$$P = \begin{pmatrix} 1-\frac{p}{2} & \frac{p}{2} \\ \frac{p}{2} & 1-\frac{p}{2} \end{pmatrix}. \tag{2}$$

This is no surprise as it reflects the independence of LSB change in the samples of each pair.

At this stage, we appeal to the Law of Large Numbers, assuming that in an observed stego object, the sizes of the trace subsets $b^0_{2m}, \ldots, b^1_{2m}$ are approximately equal to their expectations. Therefore, (1) can be inverted to obtain estimators for the cover image, given a stego image plus knowledge of p. But the inverse of M has rational polynomial entries, and the substitution $q = 1/(1-p)$ makes for a simpler expression

$$\hat{\mathbf{b}}_\mathbf{m} \approx (Q \otimes Q)\mathbf{b'_m} \tag{3}$$

where

$$Q = \frac{1}{2}\begin{pmatrix} 1+q & 1-q \\ 1-q & 1+q \end{pmatrix}. \tag{4}$$

In this way, we have estimated properties of the cover image $\hat{\mathbf{b}}_\mathbf{m}$ in terms of properties of the stego image $\mathbf{b'_m}$ and p (which is in one–one correspondence with q), possibly with some small error introduced by assuming that the observed vector is close to its expectation.

B. Cover Model and the Detector It Induces

The second part of the structural framework requires an assumption about cover images. As explained in [7], this can become rather difficult when the structure involved is more than simply pairs of pixels. But here, we only have pairs and only one sensible assumption, the same assumption (mutatis mutandis) used in [3] and [6]

$$b^0_m \approx b^1_m \text{ for each } m. \tag{5}$$

In [7], such assumptions are termed symmetries, because they are symmetrical in form and should hold, in natural images, because of symmetry of parity structure: we do not expect any correlation between pixel differences and pixel parity, in a continuous-tone image.

However, these assumptions are not useful, for all m, in discriminating cover from stego images, for example, from (1)

$$\begin{aligned} E\left[B'^0_{2m} - B'^1_{2m}\right] &= \left(1-\frac{p}{2}\right)^2 \left(b^0_{2m} - b^1_{2m}\right) + \left(\frac{p}{2}\right)^2 \left(b^1_{2m} - b^0_{2m}\right) \\ &= (1-p)\left(b^0_{2m} - b^1_{2m}\right). \end{aligned}$$

This shows that the assumption $b^0_{2m} \approx b^1_{2m}$ remains valid even after steganography. Indeed, it will tend to hold more tightly as p increases. This indicates that the symmetry $b^0_{2m} \approx b^1_{2m}$ is not useful for discrimination between cover and stego objects. There is no such problem with $b^0_{2m+1} \approx b^1_{2m+1}$, best verified empirically (we omit to do so here) and it is only these assumptions which we will use in the detector.

Finally, we make an estimate of q by finding, via (3), which value most closely matches the cover assumptions. There are a number of ways to measure the "distance" from assumptions, and in this work, we will use the simple technique of [3] which proposes the sum–square deviation $S(q) = \sum_m (\hat{b}^0_{2m+1} - \hat{b}^1_{2m+1})^2$. From (3), we read off each term

$$\begin{aligned} 4\left(\hat{b}^0_{2m+1} - \hat{b}^1_{2m+1}\right) &= (1-q^2)\left(b'^0_{2m} + b'^1_{2m} - b'^0_{2m+2} - b'^1_{2m+2}\right) \\ &\quad + (1+q)^2\left(b'^0_{2m+1} - b'^1_{2m+1}\right) \\ &\quad + (1-q)^2\left(b'^0_{2m-1} - b'^1_{2m+3}\right). \end{aligned}$$

This allows us to express $S(q)$ as a quartic in q, of which the minimum may be found by differentiating. There may be up to

three turning points, in which case it is necessary to substitute back into $S(q)$ to find which is the global minimum.[2]

This detector, although presented using a much shorter notation, is almost equivalent to that in [3]: the difference is due to the asymmetrical definition of trace subset. It will be named the couples/least squares method (LSM) detector for LSB embedding ("couples" indicating that pairs of pixels are considered and "LSM" for the least squares method dubbed in [3]). This is the detector we shall extend to work with 2LSB and I2LSB embedding. It was preferred to extend this detector rather than the original sample pairs detector of [6] because of its broadly better performance, and than the triplet-based detector of [2] or the quadruplet detector of [7] because they have already complicated cover assumptions which become more difficult under extension to two bit planes.

This is not to say that the least-squares detector is without drawbacks. Most notably, its performance for large p, particularly $p > 0.8$, is very poor. (This seems to be omitted in [3].) As noted in [2], this is because the matrix P becomes ill-conditioned for large p (indeed, it is not invertible when $p = 1$) and the discontinuity in the substitution $q = 1/(1-p)$ is symptomatic of this. More precisely, the condition number of the matrix in (3) is q^2 which is equal to $(1-p)^{-2}$. The condition number bounds the norm of errors in the output of a linear transformation relative to errors in the input. Recall that we have approximated, at (3), $E[\mathbf{B}'_\mathbf{m}]$ by $\mathbf{b}'_\mathbf{m}$. Errors in this approximation may therefore be magnified by as much as $(1-p)^{-2}$ into errors in $\hat{\mathbf{b}}_\mathbf{m}$, and this can be substantial for p close to 1.

The author has suggested in [2] that this problem should be avoided by applying such steganalysis methods after other more robust, but less accurate methods, have indicated that p is not close to 1. Such a "screening" method is an easy way to concentrate the performance of a detector where it is best but we shall ignore it in this paper. Our detectors will not be "screened" and we should not be discouraged by poor performance for high values of p. In any case, the interesting performance is for small p, where discrimination of stego objects from cover objects is difficult.

III. Conventional LSB Steganalysis for 2LSB/I2LSB Embedding

Before extending the structural techniques to two bit planes, we consider how detectors for LSB embedding can be applied to detect 2LSB and I2LSB embedding.

An obvious method is to delete the LSB in a stego image, leaving an image with one fewer bit plane. The effects of the 2LSB or I2LSB embedding on the original now appear to be to ordinary LSB replacement. Therefore, a simple way to detect 2LSB or I2LSB embedding is to apply the method of the previous section (or any other detector for LSB replacement) to an image after the LSB plane is deleted: its estimate for p should be a good one.

[2]In [8] it is shown that one should always take the smallest root of a polynomial somewhat analogous to our $S'(q)$, but this depends on further assumptions about the nature of the covers. The author prefers not to make additional assumptions, especially since the method of checking each turning point to find the global minimum is computationally cheap.

This halving of the dynamic range may have some effect on the accuracy of the detector, and we have thrown away a lot of information by deleting half of the payload. So we ask whether we can also estimate the part of the payload embedded in the LSB using conventional methods. Imagine that 2LSB or I2LSB embedding is a two-stage process, first embedding in the second-LSB, and then in the LSB; the second stage is conventional LSB embedding (except that, under 2LSB embedding, the locations of the payload are not truly independent). If the first stage does not break the model for cover images used by the LSB detector, then the stego image after the first stage is still ripe for conventional LSB steganalysis and we could form another estimate for p in the standard way. In this way, we could make two estimates—one for the payload in the second-least bit plane and one for the payload in the least—which may be valuable to combine in order to reduce overall error.

However, it is, in general, not the case that random embedding in the second-LSB plane preserves the cover model of Section II. A fully rigorous proof of this can be constructed using the structural technology of the following section, but the algebra is extremely messy and we therefore stick to a simplified explanation here, for I2LSB embedding only, and point to experimental results for verification. Let us make the assumption that the second-LSB plane is uncorrelated with any other. Suppose that proportion $p/2$ of these bits is randomly flipped—this mimics the first stage of I2LSB embedding as described before. Count the sizes of the trace subsets after this, and call their expectations $\tilde{b}^0_m$ and $\tilde{b}^1_m$. Now by a calculation similar to that in the previous section, we can derive

$$\begin{aligned}\tilde{b}^0_{4m+1} - \tilde{b}^1_{4m+1} &= (1-p)^2\left(b^0_{4m+1} - b^1_{4m+1}\right) \\ &\quad + p(1-p)\left(b^0_{4m+3} - b^1_{4m+3} + b^0_{4m-1} - b^1_{4m-1}\right) \\ &\quad + \frac{1}{2}p^2\left(2b^0_{4m+1} - b^1_{4m-3} - b^1_{4m+5}\right) \qquad (6)\end{aligned}$$

with a similar equation for $\tilde{b}^0_{4m+3} - \tilde{b}^1_{4m+3}$. The first two terms vanish under the assumption (5), but the third term does not, in general, equal zero. Some empirical investigations can check that although usually fairly small, $2b^0_{4m+1} - b^1_{4m-3} - b^1_{4m+5}$ differs significantly from zero in natural images. Therefore, the assumption (5) is broken by embedding in the second-LSB, and we can no longer expect the method of Section II to give correct answers for the size of payload in the LSB plane when embedding is also carried out in the second-LSB plane. The error induced is substantial and its extent can be seen in Section V.

IV. Extending Structural Steganalysis to Two Bit Planes

Purpose-built detectors for 2LSB and I2LSB embedding, based on the structure of replacement of two least significant bits, are now developed. First, we must change the definition of trace sets and subsets to recognize the two lowest bit planes

$$\begin{aligned}\mathcal{C}_m &= \{(j,k) \in \mathcal{P} \,|\, \lfloor s_k/4 \rfloor = \lfloor s_j/4 \rfloor + m\} \\ \mathcal{B}^{00}_m &= \{(j,k) \in \mathcal{P} \,|\, s_k = s_j + m, \text{with } s_j \equiv 0 \pmod 4\} \\ \mathcal{B}^{01}_m &= \{(j,k) \in \mathcal{P} \,|\, s_k = s_j + m, \text{with } s_j \equiv 1 \pmod 4\}\end{aligned}$$

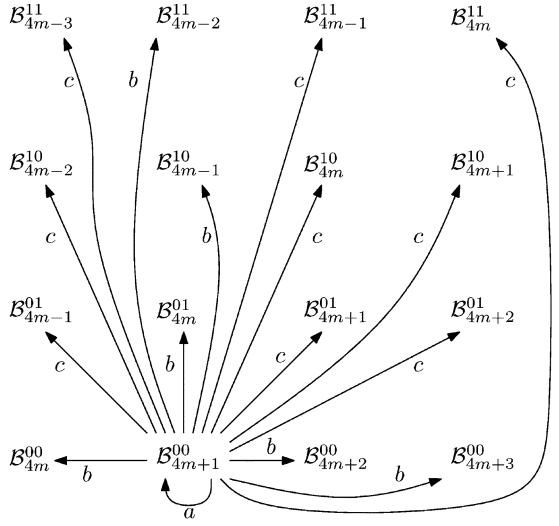

Fig. 2. Transitions between the trace subsets of $\mathcal{C}_m$, when proportion p of the maximum payload is embedded by 2LSB replacement. The transition labeled a has probability $(1-(3p/4))^2$, those labeled b have probability $(p/4)(1-(3p/4))$, and those labeled c have probability $(p/4)^2$.

and similarly for $\mathcal{B}^{10}_m$ and $\mathcal{B}^{11}_m$. As required, $\mathcal{C}_m$ is preserved by, and the trace subsets $\mathcal{B}^i_m$ sensitive to, operations on the two least significant bit planes. The trace sets $\mathcal{C}_m$ are each partitioned into 16 trace subsets which we shall consider in the following order:

$$\mathcal{B}^{00}_{4m}, \mathcal{B}^{00}_{4m+1}, \mathcal{B}^{00}_{4m+2}, \mathcal{B}^{00}_{4m+3}, \mathcal{B}^{01}_{4m-1}, \mathcal{B}^{01}_{4m}, \mathcal{B}^{01}_{4m+1}, \mathcal{B}^{01}_{4m+2},$$
$$\mathcal{B}^{10}_{4m-2}, \mathcal{B}^{10}_{4m-1}, \mathcal{B}^{10}_{4m}, \mathcal{B}^{10}_{4m+1}, \mathcal{B}^{11}_{4m-3}, \mathcal{B}^{11}_{4m-2}, \mathcal{B}^{11}_{4m-1}, \mathcal{B}^{11}_{4m}.$$

Count the sizes of the trace subsets and form vectors of length 16: $\mathbf{b_m} = (b^{00}_{4m}, \ldots, b^{11}_{4m})^T$ of the sizes before embedding, and $\mathbf{B'_m} = (B'^{00}_{4m}, \ldots, B'^{11}_{4m})^T$ for the random vector (with realisation $\mathbf{b'_m}$) of corresponding sizes after embedding of a message, assumed random.

From this point the two embedding methods must be treated separately, because they induce different transitions on the trace subsets.

A. Structural Steganalysis of 2LSB Embedding

A diagram, analogous for Fig. 1 for 2LSB embedding, is now constructed. Under 2LSB embedding, a message of length $2pN$ causes each pixel to be selected with probability p, and if selected, then there is probability (1/4) that it is not changed, and (1/4) that it is changed to each of the three other pixels with the same most-but-two significant bits.

Consider, for example, a sample pair in the trace subset $\mathcal{B}^{00}_{4m+1}$. With probability $(1-(3p/4))^2$, neither pixel value is altered, leaving the pair in the same trace subset. With probability $(p/4)(1-(3p/4))$, the first value is unchanged and the second changed from $4y+1$ to each of $4y$, $4y+2$, or $4y+3$, moving the pair into $\mathcal{B}^{00}_{4m}$, $\mathcal{B}^{00}_{4m+2}$, or $\mathcal{B}^{00}_{4m+3}$. With the same probability, the first value is changed, from $4x$ to each of $4x+1$, $4x+2$, or $4x+3$, and the second unchanged: this moves the sample pair into $\mathcal{B}^{01}_{4m}$, $\mathcal{B}^{10}_{4m-1}$, or $\mathcal{B}^{11}_{4m-2}$. The final case is that both values in the pair are changed and this will move the sample pair into any of the nine remaining trace subsets of $\mathcal{C}_m$, each with probability $(p/4)^2$. These transitions are displayed in Fig. 2. There are symmetrical transitions from each other trace subset. Fig. 2 has been laid out so that a change in the first sample value is represented by a vertical move, and a change in the second value by a horizontal move.

As before, the expected size of each trace subset after embedding is a linear combination of the sizes before embedding. Analogous to (1), we can express this in vector form. Completing all of the transitions in Fig. 2, one can check that the matrix can again be written as a tensor product (which reflects independence between horizontal and vertical transitions)

$$E[\mathbf{B'_m}] = (P' \otimes P')\mathbf{b_m} \tag{7}$$

where

$$P' = \begin{pmatrix} 1-\frac{3p}{4} & \frac{p}{4} & \frac{p}{4} & \frac{p}{4} \\ \frac{p}{4} & 1-\frac{3p}{4} & \frac{p}{4} & \frac{p}{4} \\ \frac{p}{4} & \frac{p}{4} & 1-\frac{3p}{4} & \frac{p}{4} \\ \frac{p}{4} & \frac{p}{4} & \frac{p}{4} & 1-\frac{3p}{4} \end{pmatrix}.$$

Now invert, again appealing to the law of large numbers. The substitution $q = 1/(1-p)$ is still helpful. This time, the inverse equation is

$$\hat{\mathbf{b}}_\mathbf{m} \approx (Q' \otimes Q')\mathbf{B'_m} \tag{8}$$

where

$$Q' = \frac{1}{4}\begin{pmatrix} 1+3q & 1-q & 1-q & 1-q \\ 1-q & 1+3q & 1-q & 1-q \\ 1-q & 1-q & 1+3q & 1-q \\ 1-q & 1-q & 1-q & 1+3q \end{pmatrix}.$$

Having completed the analysis of the effects of embedding on trace subsets, we must now formulate a set of cover assumptions. The equivalent to (5) would be

$$b^{00}_m \approx b^{01}_m \approx b^{10}_m \approx b^{11}_m \text{ for each } m$$

but again we find that some of these symmetries do not discriminate covers from stego objects. For example, when $p=1$, then the matrix P' consists of a single repeated entry (1/4); therefore, $P' \otimes P'$ is made up only of the entry (1/16). Equation (7) implies that all of the entries in the vector $E[\mathbf{B'_m}]$ are equal. This means that any cover assumption of the form $\mathcal{B}^i_m \approx \mathcal{B}^j_n$ will also hold for maximally embedded stego images whenever $\mathcal{B}^i_m$ and $\mathcal{B}^j_n$ are within the same trace subset.[3] This immediately helps us restrict our attention to the following ten symmetries:

$$\begin{aligned} b^{00}_n &\approx b^{01}_n, \quad \text{for } n \equiv 3 \pmod 4 \\ b^{00}_n &\approx b^{10}_n, \quad \text{for } n \equiv \{2,3\} \pmod 4 \\ b^{00}_n &\approx b^{11}_n, \quad \text{for } n \equiv \{1,2,3\} \pmod 4 \\ b^{01}_n &\approx b^{10}_n, \quad \text{for } n \equiv 2 \pmod 4 \\ b^{01}_n &\approx b^{11}_n, \quad \text{for } n \equiv \{1,2\} \pmod 4 \\ b^{10}_n &\approx b^{11}_n, \quad \text{for } n \equiv 1 \pmod 4. \end{aligned} \tag{9}$$

[3]Note that this does not guarantee that the cover assumption is unbroken in all stego images; only that it holds under maximal embedding as well as no embedding. Nonetheless, this at least indicates that, given our process of minimizing the sum–square error, we should not include such symmetries.

There may be some further filtering of the symmetries: it seems somehow strange to include the assumptions $b^{00}_{4m+2} \approx b^{11}_{4m+2}$ and $b^{00}_{4m+2} \approx b^{10}_{4m+2}$ but not $b^{10}_{4m+2} \approx b^{11}_{4m+2}$.[4]

Now it is possible to construct the estimator for p: as before, we minimize the sum–square deviation $S(q) = \sum_m (\hat{b}^{00}_{4m+1} - \hat{b}^{11}_{4m+1})^2 + \cdots$ where we include some or all of the symmetries in (9). For example, (8) gives

$$4\left(\hat{b}^{00}_{4m+1} - \hat{b}^{11}_{4m+1}\right) = A(1+3q)^2 + B(1-q)(1+3q)^2 + C(1-q)^2$$

where

$$\begin{aligned}
A &= b'^{00}_{4m+1} - b'^{11}_{4m+1} \\
B &= b'^{00}_{4m} + b'^{00}_{4m+2} + b'^{00}_{4m+3} + b'^{01}_{4m} + b'^{10}_{4m-1} + b'^{11}_{4m-2} \\
&\quad - b'^{00}_{4m+4} - b'^{01}_{4m+3} - b'^{10}_{4m+2} - b'^{11}_{4m+2} - b'^{11}_{4m+3} - b'^{11}_{4m+4} \\
C &= b'^{01}_{4m-1} + b'^{01}_{4m+1} + b'^{01}_{4m+2} + b'^{10}_{4m-2} + b'^{10}_{4m} \\
&\quad + b'^{10}_{4m+1} + b'^{11}_{4m-3} + b'^{11}_{4m-1} + b'^{11}_{4m} \\
&\quad - b'^{00}_{4m+5} - b'^{00}_{4m+6} - b'^{00}_{4m+7} - b'^{01}_{4m+4} - b'^{01}_{4m+5} \\
&\quad - b'^{01}_{4m+6} - b'^{10}_{4m+3} - b'^{10}_{4m+4} - b'^{10}_{4m+5}.
\end{aligned}$$

Each deviation from a cover symmetry has the same form. This again allows us to express the overall sum–square deviation from all of the cover symmetries as a quartic in q. Again, differentiate and find the global minimum to estimate q and, hence, p. The resulting cubic equation is not displayed which must be solved in order to find these turning points because it is immensely long to write down, but it can be determined by elementary if rather tedious calculation from (8) and (9). This estimator is called of p the 2couples detector of 2LSB replacement.

Recall from Section II that the couples/LSM detector suffers from poor performance when p is close to 1, because of small deviations between $E[\mathbf{B}'_\mathbf{m}]$ and $\mathbf{b}'_\mathbf{m}$. This effect is again evident here. The condition number of the system (8) is $q^2 = (1-p)^{-2}$ is the same as for the standard couples/LSM detector. One therefore expects to see a degradation in performance for the same values of p as in the couples/LSM case, but we note in passing that the deviation may be a little larger in the case of 2couples because each trace set is divided into more trace subsets, which are therefore smaller, and the law of large numbers approximation will be correspondingly less precise. Theoretical analysis of such errors is beyond the scope of this work and we defer to the experimental evidence of Section V.

B. Structural Steganalysis of I2LSB Embedding

When the embedding in the two-least bit planes is independent, the transitions between trace subsets is different. But they can be determined in exactly the same way and the resulting diagram (indicating the transitions from $\mathcal{B}^{00}_{4m+1}$) is shown in Fig. 3. There is a high degree of symmetry not only between horizontal and vertical transitions but also between square blocks of four trace subsets.

[4]A number of possible sets of assumptions were evaluated experimentally, and it was concluded that it does not make a great deal of difference whether all ten symmetries are used or a subset, but including any of the other six symmetries results in much poorer performance. Therefore, the decision was made to use all ten of the symmetries in (9).

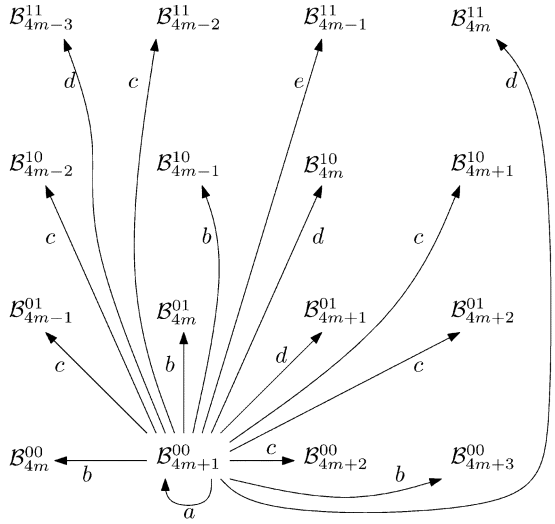

Fig. 3. Transitions between the trace subsets of $\mathcal{C}_m$, when proportion p of the maximum payload is embedded by I2LSB replacement. The transition labeled a has probability $(1-(p/2))^4$, those labeled b have probability $(p/2)(1-(p/2))^3$, those labeled c have probability $(p/2)^2(1-(p/2))^2$, those labeled d have probability $(p/2)^3(1-(p/2))$, and that labeled e has probability $(p/2)^4$.

Considering the symmetry between these transitions (or more prosaically writing down the matrix associated with Fig. 3 and checking), we see that the trace subsets, before and after embedding, are related by the fourfold Kronecker product

$$\mathrm{E}[\mathbf{B}'_\mathbf{m}] = (P \otimes P \otimes P \otimes P)\mathbf{b}_\mathbf{m} \tag{10}$$

where P is as in (2). The approximate inverse equation is

$$\hat{\mathbf{b}}_\mathbf{m} \approx (Q \otimes Q \otimes Q \otimes Q)\mathbf{B}'_\mathbf{m} \tag{11}$$

where Q is as in (4).

The cover assumptions are the same for 2LSB steganalysis.[5] Again, we seek to minimize the sum–square deviation $S(q) = \sum_m (\hat{b}^{00}_{4m+1} - \hat{b}^{11}_{4m+1})^2 + \cdots$ but this time S is a polynomial of higher degree. For example, (11) gives

$$\begin{aligned}
&16\left(\hat{b}^{00}_{4m+1} - \hat{b}^{11}_{4m+1}\right) \\
&\quad = A(1+q)^4 + B(1-q)(1+q)^3 + C(1-q)^2(1+q)^2 \\
&\qquad + D(1-q)^3(1+q) + E(1-q)^4
\end{aligned}$$

where

$$\begin{aligned}
A &= b'^{00}_{4m+1} - b'^{11}_{4m+1} \\
B &= b'^{00}_{4m} + b'^{00}_{4m+3} + b'^{01}_{4m} + b'^{10}_{4m-1} \\
&\quad - b'^{01}_{4m+3} - b'^{10}_{4m+2} - b'^{11}_{4m+2} - b'^{11}_{4m+3} \\
C &= b'^{00}_{4m+2} + b'^{01}_{4m-1} + b'^{01}_{4m+2} + b'^{10}_{4m-2} + b'^{10}_{4m+1} + b'^{11}_{4m-2} \\
&\quad - b'^{00}_{4m+4} - b'^{01}_{4m+4} - b'^{01}_{4m+5} - b'^{10}_{4m+3} - b'^{10}_{4m+4} - b'^{11}_{4m+4} \\
D &= b'^{01}_{4m+1} + b'^{10}_{4m} + b'^{11}_{4m-3} + b'^{11}_{4m} \\
&\quad - b'^{00}_{4m+5} - b'^{00}_{4m+6} - b'^{01}_{4m+6} - b'^{10}_{4m+5} \\
E &= b'^{11}_{4m-1} - b'^{00}_{4m+7}.
\end{aligned}$$

[5]Again, experiments show that we might as well take all ten symmetries in (9).

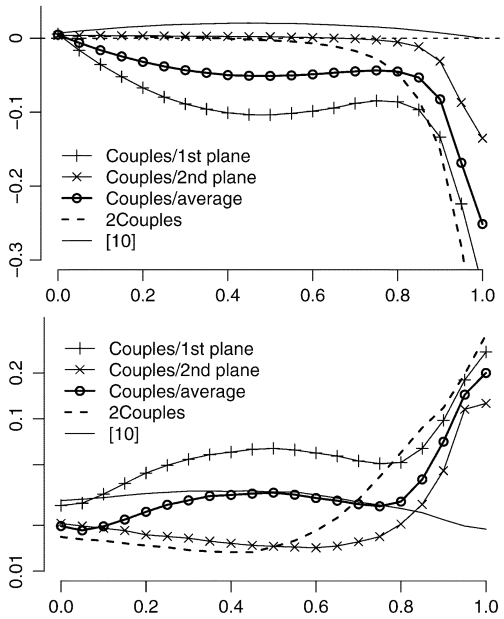

Fig. 4. Random messages inserted by 2LSB embedding into 3000 color never-compressed bitmap images. The x-axis is the proportionate payload. Above, observed bias (median error) of estimators; below, interquartile range (log axis).

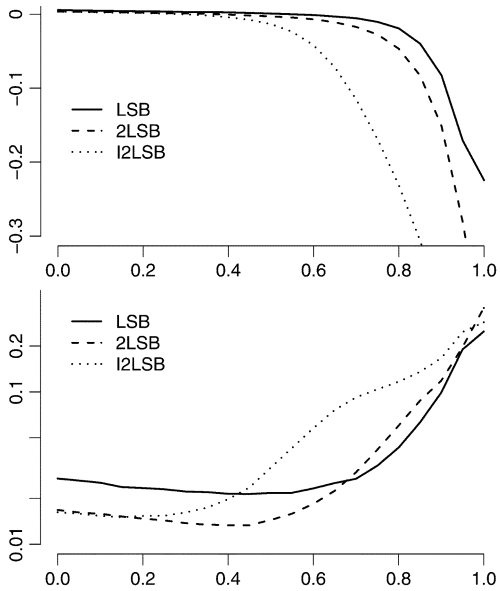

Fig. 5. Above, observed bias; below, interquartile range (log axis), when each estimator is applied to the corresponding method of steganography. The x-axis is the proportionate payload. Data from 3000 color never-compressed bitmaps.

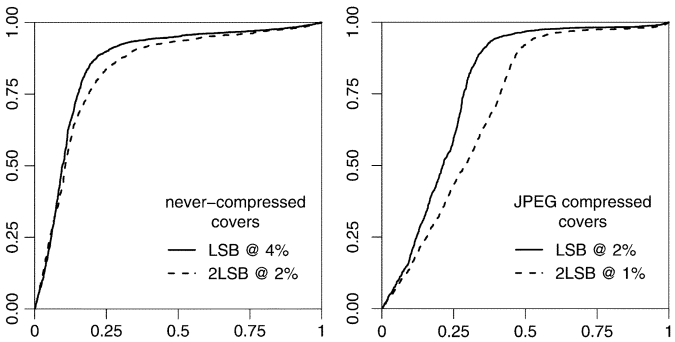

Fig. 6. ROC curves (x-axis representing false positive rate and y-axis representing true positive rate) comparing LSB embedding detected using couples/LSM, and 2LSB embedding detected using 2couples, with the same size payload. Left, results from 3000 never-compressed color images. Right, from 3000 color images subject to JPEG compression before embedding.

The function $S(q)$ is a polynomial of degree eight, so differentiating may find up to seven turning points and we substitute back into S to find the global minimum. The author calls this estimator of p the I2couples detector of I2LSB replacement.

Once again, we can expect poor performance for values of p close to 1. This time, however, the condition number of the matrix in (11) is $q^4 = (1-p)^{-4}$. Therefore, we expect the poor performance to appear more rapidly as p increases.

V. Experimental Results

The steganalysis estimators are now benchmarked, using large sets of cover images, simulating steganography of various kinds, and comparing the values produced by the estimators with the true embedded data rate. Estimator bias will be measured by the sample median, and estimator spread by the sample interquartile range (in preference to mean and standard deviation because of the results of [12] which indicate that the error distributions can be so heavy tailed that sample standard deviation may not even converge). Our primary test set of covers is 3000 never-compressed bitmaps, downloaded from http://photogallery.nrcs.usda.gov; originally very high resolution color images, we reduced them in size to approximately 640×450 pixels. These images were also converted to grayscale and/or subjected to JPEG compression (quality factor 90) so as to benchmark the performance of the detectors on these different types of cover—we shall see that the type of cover makes a big difference.

First, we perform some experiments to verify the results of Section III. For embedding rates $p \in \{0, 0.05, 0.1, \ldots, 1\}$, we inserted random messages by 2LSB embedding in each of the 3000 covers, and tested the performance of the following estimators: applying the standard couples/LSM detector, applying the standard detector after deleting the lowest bit plane, averaging the previous two estimates, and applying the 2couples detector. Also tested was the other 2LSB detector in the literature [10]. The results are shown in Fig. 4.

It is apparent that all of the structural detectors' performance is poor for (the uninteresting case of) p close to 1, manifesting substantial negative bias and large spread. It has already been mentioned that this would, in practice, be avoided by "screening." The new 2couples detector is the superior method for the interesting case of $p < 0.5$, although application of the couples/LSM method to an image with the lowest bit plane deleted is, in fact, a better performer for $p > 0.5$. But the standard couples/LSM method is not a good estimator for 2LSB embedding, suffering immediately from significant negative bias and large deviations; its effect is also to ruin the average of two applications of the standard method, with and without the least bit-plane removed. This verifies our calculations in Section III, showing that embedding in the second-least bit plane damages the assumptions of couples analysis in the least

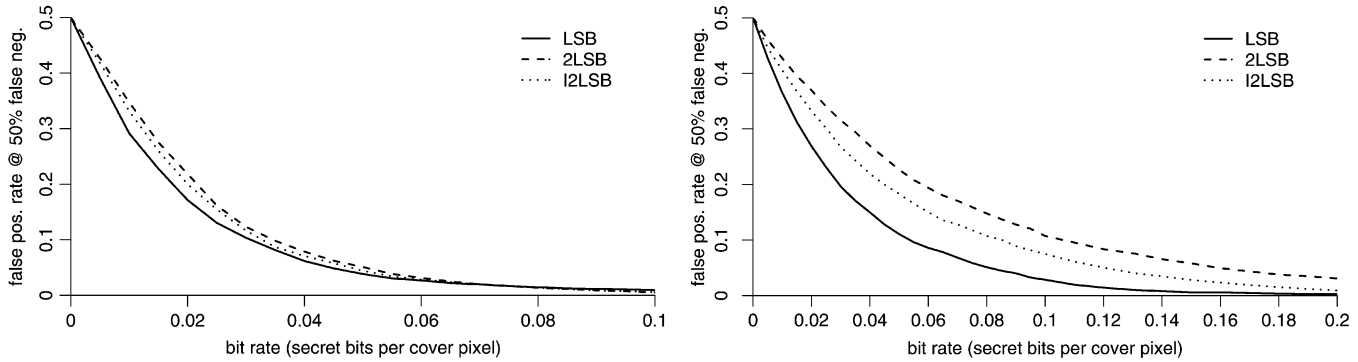

Fig. 7. Observed false positive rate at which the false negative rate is 50%, testing each steganography method against the corresponding steganalysis method. Left, results from 3000 never-compressed grayscale images. Right, from 3000 color images subject to JPEG compression before embedding.

bit plane. (These experiments were repeated for I2LSB embedding but the charts are very similar and we do not include them here; the conclusion is the same: standard couples steganalysis is not appropriate for I2LSB embedding either.)

Observe that the detector of [10] is weak, except for very large payloads. For payloads below $p = 0.5$, estimator errors are 2–3 times larger for the detector of [10] than our novel 2couples detector. This accords with the observations of [13] and [12], where analogous detectors for plain LSB steganography were carefully benchmarked. Further experiments, for which a chart is not displayed, show that the performance of the detector in [10] is relatively much worse when the cover images had been previously subjected to JPEG compression. It appears that WS-based detectors cannot match the performance of structural detectors based on [2] in the (interesting case of) detection of small payloads, but they do not suffer from poor performance near $p = 1$ and, therefore, have a place in the estimation of near-maximal payloads.

Experiments are now undertaken which compare the performance of each LSB, 2LSB, and I2LSB embedding when the appropriate structural detector—couples/LSM, 2couples, I2couples—is applied. Fig. 5 shows the corresponding results. Note that the I2couples detector shows poor performance more rapidly as p grows; this is in accordance with the theory of the previous section, because this detector is based on a linear system with a higher condition number. Note also that the 2couples and I2couples detectors, for $p < 0.4$, are substantially more accurate in estimating the value of p.

But we should not be misled: although it appears that the estimation of LSB steganography is harder than 2LSB or I2LSB (by the methods presented here), that is not the case when the relative sizes of the payload are considered: a value of p for 2LSB or I2LSB embedding represents twice as much data as the same value for LSB. In Fig. 6, we make an equivalent payload comparison of LSB and 2LSB steganography, detected by the couples/LSM and 2couples detectors, focusing now on the binary question of whether any data are embedded. Such performance is measured by receiver operating characteristic (ROC) curves, which display how false positive and false negative rates vary as sensitivity—in this case, a threshold for the diagnosis of steganography—is altered. Two such ROC curves are displayed, for two types of cover objects and selecting an interesting embedding rate for each.

Observe that, given a payload of fixed size, the discrimination of cover objects from stego objects is slightly harder (less reliable) when the embedding is done using 2LSB than LSB replacement. Ideally, this would be verified for a range of embedding rates, and for I2LSB embedding as well, but we cannot display ROC curves for every possible rate. Instead, we settled on the metric of the false positive rate when the false negative rate is 50% (used and justified in part in [1]), displaying how this varies with the embedding rate (this time measured in bits per cover pixel, fairly to compare LSB with 2LSB and I2LSB embedding, at intervals of 0.005) in Fig. 7. Two such charts are shown, illustrating the cases of never-compressed grayscale and previously JPEG-compressed color images. From the steganographer's point of view, the least detectable embedding method is 2LSB embedding. I2LSB is intermediate and LSB is the most detectable. For never-compressed grayscale images, the difference is quite small (although some simple bootstraps show that it is nevertheless statistically significant for $p \in [0.005, 0.05]$) and very substantial for color JPEGs.

The experiments were repeated on a wide range of sets of cover images (including one taken directly from digital cameras in raw format, never subject to image-processing operations) with very consistent results: the detector of [10] is substantially outperformed by the structural detectors except for near-maximal payloads (for which the discrimination problem is extremely easy in any case); from the steganographer's point of view, there is a small advantage in using 2LSB embedding in never-compressed images and a large advantage in covers previously subject to JPEG compression (including when the covers are substantially reduced in size after compression). In no case did we find it significantly preferable to use LSB embedding over 2LSB embedding.

VI. Conclusions and Further Work

After making a clear distinction between 2LSB and I2LSB embedding, we have extended the structural framework of [2] to produce novel detectors specialized for these embedding methods. They have been demonstrated to be more sensitive than the other such detector in the literature. The various embedding methods have been benchmarked against their respective detectors: as close as possible we have been comparing them on a like-for-like basis, with all types of steganography detected by structural detectors using pairs of

pixels, a least-squares estimator, and cover assumptions of the same type. Therefore, we believe that this is a fair comparison, and that it appears that 2LSB embedding is genuinely superior (slightly so in never-compressed covers; substantially so in previously compressed covers) to LSB embedding. That is not to say that either method is sensible if an alternative to bit overwriting is available.

Of course, we cannot say for sure that the "steganographic capacity" is higher under 2LSB embedding because there might be other detectors which make 2LSB embedding the easier one to detect. However, it is now very clear that structural detectors for LSB replacement are comfortably the most sensitive. Although it is possible that, for example, structural steganalysis of 2LSB replacement using an extended triples method might prove easier than of LSB replacement using the triples detector of [2], this seems unlikely. In any case, an extension to more than pairs of pixels is likely to run in to complexities with the cover assumptions.

Another natural progression would be the steganalysis of embedding in three or more bit planes. Perhaps "3LSB" embedding is more secure than 2LSB embedding? The author urges caution: as a reductio ad absurdum argument, consider simple replacement of whole cover bytes by a stego payload: there is no structure in this embedding, and structural steganalysis will fail. But, of course, such embedding is easily perceptible. At some point, the structure of multiple bit-plane embedding will be redundant, and other detection methods, such as the additive-noise steganalysis of [14] will be more appropriate.

One issue we have not addressed here, postponing it to future work, is how to determine which steganalysis method to apply. Given a stego image, should one estimate the amount of hidden data using couples, 2couples, or I2couples? There are a number of ways to make this determination and we would prefer to find a method with some optimality. This requires closer consideration of the nature of errors in our cover assumptions.

REFERENCES

[1] A. Ker, "Improved detection of LSB steganography in grayscale images," in *Proc. 6th Information Hiding Workshop*, Heidelberg, Germany, 2004, vol. 3200, pp. 97–115, ser. Springer Lecture Notes Computer Science.

[2] ——, "A general framework for the structural steganalysis of LSB replacement," in *Proc. 7th Information Hiding Workshop*, Heidelberg, Germany, 2005, vol. 3727, pp. 296–311, ser. Springer Lecture Notes Computer Science.

[3] P. Lu, X. Luo, Q. Tang, and L. Shen, "An improved sample pairs method for detection of LSB embedding," in *Proc. 6th Information Hiding Workshop*, Heidelberg, Germany, 2004, vol. 3200, pp. 116–127, ser. Springer Lecture Notes Computer Science.

[4] J. Fridrich, M. Goljan, and R. Du, "Detecting LSB steganography in color and grayscale images," *IEEE Multimedia, Special Issue on Security*, vol. 8, no. 4, pp. 22–28, Oct.–Dec. 2001.

[5] J. Fridrich, M. Goljan, and D. Soukal, "Higher-order statistical steganalysis of palette images," in *Proc. SPIE, Security and Watermarking of Multimedia Contents V*, E. J. Delp III and P. W. Wong, Eds., 2003, vol. 5020, pp. 178–190.

[6] S. Dumitrescu, X. Wu, and Z. Wang, "Detection of LSB steganography via sample pair analysis," *IEEE Trans. Signal Process.*, vol. 51, no. 7, pp. 1995–2007, Jul. 2003.

[7] A. Ker, "Fourth-order structural steganalysis and analysis of cover assumptions," in *Proc. SPIE, Security, Steganography and Watermarking of Multimedia Contents VIII*, E. J. Delp III and P. W. Wong, Eds., 2006, vol. 6072, pp. 25–38.

[8] S. Dumitrescu and X. Wu, "A new framework of LSB steganalysis of digital media," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pt. 2, pp. 3936–3947, Oct. 2005.

[9] K. Sullivan, O. Dabeer, U. Madhow, B. Manjunath, and S. Chandrasekaran, "LLRT based detection of LSB hiding," in *Proc. IEEE Int. Conf. Image Processing*, 2003, vol. 1, pp. 497–500.

[10] X. Yu, T. Tan, and Y. Wang, "Extended optimization method of LSB steganalysis," in *Proc. IEEE Int. Conf. Image Processing*, 2005, vol. 2, pp. 1102–1105.

[11] J. Fridrich and M. Goljan, "On estimation of secret message length in LSB steganography in spatial domain," in *Proc. SPIE, Security, Steganography, and Watermarking of Multimedia Contents VI*, E. J. Delp III and P. W. Wong, Eds., 2004, vol. 5306, pp. 23–34.

[12] R. Böhme and A. Ker, "A two-factor error model for quantitative steganalysis," in *Proc. SPIE, Security, Steganography and Watermarking of Multimedia Contents VIII*, E. J. Delp III and P. W. Wong, Eds., 2006, vol. 6072, pp. 59–74.

[13] R. Böhme, "Assessment of steganalytic methods using multiple regression models," in *Proc. 7th Information Hiding Workshop*, Heidelberg, Germany, 2005, vol. 3727, pp. 278–295, ser. Springer Lecture Notes Computer Science.

[14] M. Goljan, J. Fridrich, and T. Holotyak, "New blind steganalysis and its implications," in *Proc. SPIE, Security, Steganography and Watermarking of Multimedia Contents VIII*, E. J. Delp III and P. W. Wong, Eds., 2006, vol. 6072, pp. 1–13.

Andrew D. Ker (M'06) was born in Birmingham, U.K., in 1976. He received the B.A. degree in mathematics and computer science and the D.Phil. degree in computer science from Oxford University, Oxford, U.K., in 1997 and 2001, respectively.

Previously, he was a Junior Research Fellow with University College, Oxford. Currently, he is a Royal Society University Research Fellow with the Computing Laboratory, Oxford University. His initial work was in the foundations of computer science and his research interests are steganography and steganalysis.

Dr. Ker is a member of the SPIE, The International Society for Optical Engineering.