Hard Communication Channels for Steganography

Sebastian Berndt[1] **and Maciej Liśkiewicz**[2]

1 **University of Lübeck, Lübeck, Germany**
berndt@tcs.uni-luebeck.de
2 **University of Lübeck, Lübeck, Germany**
liskiewi@tcs.uni-luebeck.de

Abstract

This paper considers steganography – the concept of hiding the presence of secret messages in legal communications – in the computational setting and its relation to cryptography. Very recently the first (non-polynomial time) steganographic protocol has been shown which, for any communication channel, is provably secure, reliable, and has nearly optimal bandwidth. The security is unconditional, i.e. it does not rely on any unproven complexity-theoretic assumption. This disproves the claim that the existence of one-way functions and access to a communication channel oracle are both necessary and sufficient conditions for the existence of secure steganography in the sense that secure and reliable steganography exists independently of the existence of one-way functions. In this paper, we prove that this equivalence also does not hold in the more realistic setting, where the stegosystem is polynomial time bounded. We prove this by constructing (a) a channel for which secure steganography exists if and only if one-way functions exist and (b) another channel such that secure steganography implies that *no* one-way functions exist. We therefore show that security-preserving reductions between cryptography and steganography need to be treated very carefully.

1998 ACM Subject Classification E.3 Data Encryption

Keywords and phrases provable secure steganography, cryptographic assumptions, pseudorandom functions, one-way functions, signature schemes

Digital Object Identifier 10.4230/LIPIcs.ISAAC.2016.16

1 Introduction

Digital steganography has recently received substantial interest in modern computer science since it allows secret communication without revealing its presence. Currently, using freely available steganographic software, one party is able to spread secret messages over widely accessible services, such as photo-sharing websites, camouflaging the presence of the messages in legal communications. Although the uploads and views by other users can be recorded and analyzed it is fairly difficult to distinguish the altered documents containing a secret message from those of millions of the other ordinary documents. For more details on applied steganography see the textbook [16] or the current survey [38] and the literature therein. For applications of steganography in other areas, like covert computation, broadcasting, or anonymous communication see e.g. [6, 7, 14, 18, 24, 35].

A common computational model for secret-key steganography, also used in this paper, was introduced by Hopper, Langford, and von Ahn [21, 22, 23]. Independently, Katzenbeisser and Petitcolas [25] provided a similar formulation. In this setting, a *stegosystem* is defined as a pair of probabilistic algorithms, called encoder and decoder, which share a secret-key. The aim of the encoder (often called Alice or the steganographer) is to hide a secret message in a document and to send it to the decoder (Bob) via a public *channel* $\mathcal{C}$, which is completely

monitored by an *adversary* (Warden or steganalyst). The channel is modeled by a *cover-document* sampler that can be queried adaptively, in a black-box manner and the adversary's task is to distinguish those from altered ones called *stego-documents*.

To hide a secret message m, the encoder can take sample cover-documents, based on past communication, and manipulate them to embed m. The decoder, receiving stego-documents, should be able to decode the hidden message correctly. The stegosystem is called *reliable* if the decoder succeeds with high probability. The adversary is a probabilistic algorithm with access to additional knowledge about the channel. A stegosystem is *secure* if no adversary of *polynomial* time complexity is able to distinguish with significant probability between cover- and stego-documents generated by the stegosystem's encoder. This implies in general that the distributions of cover-documents and stego-documents have to be fairly close in a complexity-theoretic sense. The *insecurity* of a stegosystem is the advantage of the best adversary to distinguish between cover- and stego-documents. Thus, a stegosystem is secure if its insecurity is sufficiently small, i.e. negligible in the security parameter κ defining the length of the shared secret-key.

The security and reliability are necessary attributes of any reasonable stegosystem. Additionally, the system should be efficient in terms of the transmission *rate* (payload), i.e. the number of bits transmitted per single stego-document should be as high as possible. The stegosystems used in practice (not necessary *provable* secure in the computational model) typically achieve a rate of $\sqrt{n}$ [26], where $n := n(\kappa)$ denotes the length of a single document that is polynomial in κ. A longstanding conjecture, the *Square Root Law of Steganographic Capacity* [15, 27] says that a rate of the form $(1-\varepsilon)\sqrt{n}$ is always achievable in the information-theoretic setting.

Importantly, in the definition of the computational model Hopper, Langford and von Ahn [21, 22, 23] do not bound the running time of the stegosystem, while the time complexity of the adversary is required to be bounded by a polynomial. For this setting we have shown very recently the strongest possible result; namely, that there exists a *universal* stegosystem which for *any* channel is secure, reliable and achieves almost optimal rate. Recall, that a system is called *universal*[1] if the encoding method does not rely on knowledge of the distribution for the channel $\mathcal{C}$ except that its min-entropy is sufficiently large.

▶ **Theorem 1** ([4], Informal). *There exists a universal (non-polynomial time) stegosystem $\mathcal{S}$ that is unconditionally secure and reliable. Moreover $\mathcal{S}$ is rate-efficient.*

This disproves the widely circulated result claimed in [21, 23] that the existence of one-way functions and access to a communication channel oracle are both necessary and sufficient conditions for the existence of secure steganography (see e.g. the textbook [16] for a discussion). In fact, secure and reliable (non-polynomial) steganography exists independently of the existence of one-way functions.

In this paper we investigate a more reasonable setting in which the stegosystem's running time is bounded by a polynomial and study provably secure steganography and its relation to cryptography. We prove that, despite strong connections, polynomial time steganography is *not* cryptography. More precisely we show that, similarly as in the case of non-polynomial time steganography, the equivalence between the existence of one-way functions and the existence of secure, reliable, and rate-efficient (polynomial time) steganography does not hold.

[1] In the literature universal stegosystems are also called "black-box".

1.1 Previous Works

As we discuss in [4], a commonly heard argument for the premise that steganography is cryptography goes as follows: Let m and m' be two different secret messages and s and s' be stego-documents which embed m, resp. m'. If the distributions of s and s' are indistinguishable from the distribution of the cover-documents, then by the triangle-inequality, the distributions of s and s' are also indistinguishable. Hence, a secure stegosystem is also a secure cryptosystem.

While the argument concerning the triangle-inequality is true, the argument ignores the channel oracle. If the channel documents are e.g. natural digital pictures, the cryptosystem simulating the stegosystem needs access to samples of those documents. But an efficient sampler for this channel seems highly unlikely. Thus, this reasoning is wrong and in fact we show in [4] that (non-polynomial time) steganography exists independently of the existence of one-way functions. Below we discuss known results in this direction.

In contrast to the non-polynomial case, *universal* steganography is very limited when requiring polynomial running time. In [10], Dedić et al. proved that for every stegosystem $\mathcal{S}$ with security parameter κ (describing the length of the secret key) which hides $\lambda := \lambda(\kappa)$ bits, takes $q := q(\kappa)$ samples per stego-document and runs in time $p := p(\kappa)$ there exists a channel $\mathcal{C}(\kappa)$ of min-entropy $\mathsf{pol}(\kappa)$ such that

$$\mathbf{InSec}(\kappa) + \mathbf{UnRel}(\kappa) \;\geq\; \frac{1}{2} - \frac{e \cdot q}{2^{\lambda}} - \Psi(p, \kappa) - o(1). \tag{1}$$

Here, $\mathbf{InSec}(\kappa)$ denotes the insecurity (against polynomial time bounded wardens) and $\mathbf{UnRel}(\kappa)$ the unreliability of $\mathcal{S}$ on $\mathcal{C}(\kappa)$, and Ψ describes a term caused by the insecurity of the pseudorandom function used in the construction of $\mathcal{C}(\kappa)$. From this result we get that if restricted to polynomial time steganography, Theorem 1 does not hold unless one-way functions do not exist:

▶ **Theorem 2** ([10], Informal statement). *Assuming one-way functions exist there exists no secure and reliable universal* polynomial time *stegosystem of rate* $\omega(\log \kappa)$.

Interestingly, the logarithmic bound on the bandwidth above is sharp. Due to Hopper et al. [23] and Dedić et al. [10] we know that the existence of one-way functions implies the existence of a secure and reliable universal (polynomial time) stegosystem of rate $\mathcal{O}(\log \kappa)$.

Theorem 2 shows a very important property, interesting in itself: when requiring polynomial time, the applicability of universal steganography is very limited. Due to this reason it makes sense to consider the security of a stegosystem $\mathcal{S}$ only for a specific channel or for channels of a specific family, and do not to require its security for all possible channels. This is also a common approach in practical steganography where a system has to satisfy security properties for a specific channel, like e.g. natural images in JPEG-format, but its security for texts, audio signals, TCP/IP transmission packages, etc. is irrelevant. For this setting the relationship between steganography and cryptography remains unsolved. Particularly, it is not known whether for any channel $\mathcal{C}(\kappa)$ there exists a secure, reliable, and rate-efficient (polynomial time) stegosystem for $\mathcal{C}(\kappa)$. The question remains open both for unconditional security and under some unproven assumptions like the existence of one-way functions.

Note that the lower bound (1) above does not allow to answer this question. To prove their result, Dedić et al. [10] show that for every (polynomial time) stegosystem $\mathcal{S}$ there exists a channel $\mathcal{C}(\kappa)$ that satisfies inequality (1). However, every channel $\mathcal{C}(\kappa)$ of [10] has a secure, reliable and rate-efficient (polynomial time) stegosystem (for a proof see e.g. [31]). Also the following lower bound provided by Hopper et al. [23] does not suffice to solve this

problem. They show that for any function $q(\kappa)$ bounded by a polynomial in κ there exists a channel $\mathcal{C}(\kappa)$ such that for every (polynomial time) stegosystem $\mathcal{S}$ of query complexity $q(\kappa)$ which hides $\lambda(\kappa)$ bits per document it is true

$$\mathbf{InSec}(\kappa) + \mathbf{UnRel}(\kappa) \;\geq\; 1 - q/2^{\lambda} - 2^{-\kappa}. \tag{2}$$

In case $\lambda(\kappa) \in \omega(\log \kappa)$ the right-hand side of the inequality (2) is big, meaning that $\mathcal{S}$ is insecure or unreliable, but again in in this situation one can construct a (polynomial time) stegosystem $\mathcal{S}'$ of query complexity $q(\kappa) + 1$ that is secure, reliable and rate-efficient on $\mathcal{C}(\kappa)$.

Hence both of these lower bounds prove that every stegosystem that hides $\omega(\log \kappa)$ bits is insecure or unreliable on *some* channel from a channel family $\mathcal{F}$. On the other hand, for all of those channels, one can construct a secure and reliable stegosystem. Hence, the insecurity or unreliability of the stegosystem on those channels comes from the fact that the stegosystem must work for *all* channels in $\mathcal{F}$ and not necessarily from the complexity of a single channel.

1.2 Our Contributions

We prove that polynomial time bounded, provably secure, reliable, and rate-efficient steganography is independent of cryptographic assumptions, such as the existence of one-way functions. This is a consequence of the following results.

▶ **Theorem 3** (Informal). *Assuming one-way functions exist there exists a channel $\mathcal{C}(\kappa)$ such that for $\mathcal{C}(\kappa)$ no secure and reliable polynomial time stegosystem of rate $\omega(\log \kappa)$ is possible.*

The logarithmic bound on the bandwidth above is sharp unless one-way functions do not exist. One can conclude even more, namely that if Theorem 3 holds for rate $\mathcal{O}(\log \kappa)$, no one-way functions exists. More formally, we have the following:

▶ **Corollary 4.** *If proposition (a) is true:*

(a) *Assuming one-way functions exist there exists a channel $\mathcal{C}(\kappa)$ such that for $\mathcal{C}(\kappa)$ no secure and reliable polynomial time stegosystem of rate $\mathcal{O}(\log \kappa)$ is possible;*

then one-way functions do not exist.

To see this, again from [23] and [10] we know that: (b) *If one-way functions exist then for every channel $\mathcal{C}(\kappa)$ there exists secure and reliable polynomial time stegosystem of rate $\mathcal{O}(\log \kappa)$*. Thus, proving the proposition (a) in Corollary 4 would be possible only if one-way functions do not exist – only in this case both of the proposition (a) and (b) are true. Clearly, current research is far from proving anything like proposition (a).

Theorem 3 is the main technical achievement of this paper. We complement our result by showing a channel for which the existence of one-way functions implies the existence of a secure, reliable, and rate-efficient polynomial time stegosystem. Constructions of similar channels are known in the steganography community however, for the sake of correctness and completeness we formulate and prove a suitable result in our paper:

▶ **Theorem 5** (Informal). *There exists a channel $\mathcal{C}(\kappa)$ such that if one-way functions exist then secure, reliable, and rate-efficient polynomial time stegosystem for $\mathcal{C}(\kappa)$ exists.*

The proofs of the theorems are constructive. Interestingly, the channel $\mathcal{C}(\kappa)$ satisfying Theorem 3 is specified by a cryptographic signature scheme protocol that is widely used in practice. While $\mathcal{C}(\kappa)$ per se is artificial, its close relative, the channel of cryptographic signed emails on the internet, is widely used. In this work we prove also that there exist more such hard channels satisfying the conditions of Theorem 3. In fact we show that any channel

which can express the signature scheme belongs to this family. Our construction is inspired by the technique used in the work of De et al. [9] which apply this method to show that it is not possible to uniformly generate satisfying assignments to a 3-CNF formula if one is given polynomial many samples of satisfying assignments. The channels satisfying the conditions of Theorem 5 are channels that can be sampled by an algorithm in polynomial time.

1.3 Relevant Work

The running time of universal steganography was improved by Kiayias et al. in [28] by using t-wise independent family of functions instead of a pseudorandom function to choose the corresponding documents from the channel. They also showed that a key length of $(1+o(1))n$ is sufficient to achieve information-theoretic security of $2^{-n/log^{O(1)}(n)}$ for message length n.

Van Le and Kurosawa [29] used arithmetic coding techniques to improve upon the rates of the universal systems proposed in [23] and [10]. In order to achieve this they assume that the system has access to additional knowledge on the channel. Their work thus does not fit into the model introduced by Hopper et al. [23].

Von Ahn and Hopper [36] gave the first complexity-theoretic definitions of *public-key steganography*, where the running time of the stegosystem is polynomial time bounded. Their work was extended by Backes and Cachin [2], who introduced stronger security definitions and presented a universal non-rate-efficient stegosystem for one of their definitions. Hopper [20] then proceeded by proving that every so-called efficiently sampleable channel has a non-rate-efficient stegosystem that achieves the strongest security definition.

Universal stegosystems have also been studied in the information-theoretic setting, where the information-theoretic distance between the distribution of the channel documents and the distribution of the stego-documents must be bounded. The first information-theoretic definitions of steganography were given by Cachin [5]. Wang and Moulin [37] presented a whole framework to study the optimal embedding rate of information-theoretic perfect stegosystems. For more information on this see e.g. [8, 15, 33].

The paper is organized as follows: The next section contains the basic definitions regarding stegosystems, their security and the cryptographic primitives we make use of. The proof of Theorem 3 and its extension can be found in Section 3, while Theorem 5 is proved in Section 4. Finally, we conclude our paper and discuss the future work in Section 5.

2 Preliminaries and Definitions

We say that an algorithm A has *oracle access* to a probability distribution D (denoted as A^D), if A can sample an element d according to D in unit time. The elements are sampled independently. If D is parameterized by $\rho_1, \rho_2, \ldots, \rho_k$, we write $A^{D(\rho_1,\ldots,\rho_k)}$ to describe the situation, where all of the parameters are fixed. If D is allowed to choose the parameter ρ_i itself, this is denoted by a dot, as in $A^{D(\rho_1,\ldots,\rho_{i-1},\cdot,\rho_{i+1},\ldots,\rho_k)}$. More generally, we also use dots in the parameters of an algorithm to indicate that this parameter may be chosen freely.

If one tries to hide the transfer of a secret message via unsuspicious communication, one first needs to define a model for this type of communication. This is done via the notion of a *channel* $\mathcal{C}$ on an alphabet Σ.

▶ **Definition 6.** A *channel* $\mathcal{C}$ on the alphabet Σ is a function taking an $n \in \mathbb{N}$ and a history $h \in (\Sigma^n)^*$ to a probability distribution on Σ^n, denoted by $\mathcal{C}_{h,n}$.

Note that we do not require the distributions $\mathcal{C}_{h,1}, \mathcal{C}_{h,2}, \ldots$ to be polynomial time constructible, as the typical channels in use may be of high complexity, e.g., pictures or poems.

As usual, a communication channel has a certain capacity, that is bounded by the entropy of the channel. The *min-entropy* $\mathcal{H}(D)$ of a probability distribution D is defined as $\mathcal{H}(D) := \min_{d \in \mathrm{supp}(D)}\{-\log D(d)\}$. The *min-entropy* $\mathcal{H}(\mathcal{C}_n)$ of a channel $\mathcal{C}$ with respect to $n \in \mathbb{N}$ is then defined as $\mathcal{H}(\mathcal{C}_n) = \min_h\{\mathcal{H}(\mathcal{C}_{h,n})\}$. The number of bits embeddable into a single document is bounded by $\mathcal{H}(\mathcal{C}_n)$ (see e.g. [22] for a proof).

To give a sound formal treatment, we parameterize the behaviour of all parties by the security parameter κ – the length of the secret key k. We therefore say that a function $f\colon \mathbb{N} \to [0,1]$ is *negligible*, if for every c and all sufficiently large n, it holds that $f(n) < n^{-c}$.

Informally, a stegoencoder SE has access to samples of $\mathcal{C}$ and *embeds* a message m into a sequence of documents $d_1, \ldots, d_\ell$, thereby producing a sequence $d_1^*, \ldots, d_\ell^*$. The goal of SE is that no efficient algorithm can distinguish the distributions of $d_1, \ldots, d_\ell$ and $d_1^*, \ldots, d_\ell^*$.

▶ **Definition 7.** A *stegosystem* $\mathcal{S}$ for the polynomial time constructible *message space* $\{\mathcal{M}_\kappa\}_{\kappa\in\mathbb{N}}$ with *document length* $n\colon \mathbb{N} \to \mathbb{N}$ and *output length* $\ell\colon \mathbb{N} \to \mathbb{N}$ is a pair of probabilistic, polynomial time Turing machines (PPTMs) $[SE, SD]$ with the following functionality upon security parameter κ:

- The *encoding algorithm* SE takes as input a key $k \in \{0,1\}^\kappa$, a message $m \in \mathcal{M}_\kappa$, a history h and a *state information* $s \in \{0,1\}^*$ and produces a document d and state information $s' \in \{0,1\}^*$ by having access to $\mathcal{C}_{h,n}$. By $SE^{\mathcal{C}}(k,m,h)$, we denote the outcome of:

  Steganographic Encoding $SE^{\mathcal{C}}(k,m,h)$

  1. $s := \lambda$; *// initialize the first state as the empty string*
  2. **for** $i = 1, \ldots, \ell$:
  3. $\quad (d_i, s) \leftarrow SE^{\mathcal{C}_{h,n(\kappa)}}(k,m,h,s)$;
  4. $\quad h := hd_i$; *// concatenate h with the new document*
  5. **return** $d_1, \ldots, d_\ell$

  Note that SE is only allowed to get samples for the $i+1$-th document, after it produced the i-th document. For the sake of simplicity, we sometimes write $SE^{\mathcal{C}}(k,m,h)_i$ to denote the i-th document d_i.
- The *decoding algorithm* SD takes as input a key $k \in \{0,1\}^\kappa$ and a sequence of documents $d_1, \ldots, d_\ell$ and outputs a message m'.

The *sampling complexity* $q(\kappa)$ of SE is the number of calls of SE to its sampling oracle. The *transmission rate* $b(\kappa)$ is defined as $b(\kappa) := {}^{\log|\mathrm{supp}(\mathcal{M}_\kappa)|}/_{\ell(\kappa)} \leq n(\kappa)$.

The key $k \in \{0,1\}^\kappa$ is shared by SE and SD before the embedding process. Clearly, SD should be able to reconstruct the original message with high probability. We say that $\mathcal{S}$ is *ρ-reliable*, if the maximum probability of an error (i.e. $SD(k, SE^{\mathcal{C}}(k,m,h)) \neq m$) is bounded by $1 - \rho(\kappa)$ for every message m and every history h. If $\mathcal{S}$ is ρ-reliable for a negligible ρ, we call $\mathcal{S}$ *reliable*. In addition to this, SE wants to embed as much information as possible into a document in order to reduce the overhead of the transmission. We say that $\mathcal{S}$ is *rate-efficient*, if there is constant $\alpha > 0$ such that $b(\kappa) \geq \mathcal{H}(\mathcal{C}_{n(\kappa)})^\alpha$ for all κ (we thus embed n^α bits per document with entropy n).

2.1 Security of a Stegosystem

A *warden* W is a PPTM that should decide whether the communication parties use steganography or not. In order to do so, W chooses a history and a message and presents this to a *challenge oracle* CH which, on key k, message m and history h outputs a sequence of $\ell(\kappa)$

documents $d_1, \ldots, d_\ell$. This sequence is either the output of the stegosystem $SE^{\mathcal{C}}(k, m, h)$ for a uniformly chosen key k or the ℓ-fold output $\mathcal{C}_\ell(m, h)$ of the channel with the distribution $\mathcal{C}_\ell(m, h) \sim (\mathcal{C}_{h,n(\kappa)}, \mathcal{C}_{hd_1,n(\kappa)}, \ldots, \mathcal{C}_{hd_1d_2\cdots d_{\ell-1},n(\kappa)})$.

The goal of W is now to reconstruct whether the challenge oracle is $SE^{\mathcal{C}}(k, \cdot, \cdot)$ (it outputs 0 or »Stego«) or $\mathcal{C}_\ell(\cdot, \cdot)$ (it outputs 1 or »not Stego«). More precisely, we consider the following experiment for an chosen hiddentext attack (CHA):

Steganographic Security $\text{CHA}_\kappa(W, \mathcal{S}, \mathcal{C})$

1. $k \leftarrow \{0,1\}^\kappa$;
2. $b \leftarrow \{0,1\}$;
3. **if** $b = 0$ **then** $CH := SE^{\mathcal{C}}(k, \cdot, \cdot)$ **else** $CH := \mathcal{C}_\ell(\cdot, \cdot)$;
4. $b' \leftarrow W^{\mathcal{C}_{\cdot,n(\kappa)}, CH(\cdot,\cdot)}(1^\kappa)$; *// W chooses h and m for CH*
5. **if** $b = b'$ **then return** 1 **else return** 0

Note that the warden has several liberties: It may choose the history for the channel oracle (the stegosystem can only work with its given history), the history submitted to the challenge oracle and the message submitted to the challenging oracle.

As W is able to chose the message (or the hiddentext), we say that the *chosen hiddentext attack (*CHA*) advantage* $\mathbf{Adv}^{\text{CHA}}_{W,\mathcal{S},\mathcal{C}}(\kappa)$ of W on the stegosystem $\mathcal{S}$ on channel $\mathcal{C}$ is given as

$$\mathbf{Adv}^{\text{CHA}}_{W,\mathcal{S},\mathcal{C}}(\kappa) = 2 \cdot \left|\Pr[\text{CHA}_\kappa(W, \mathcal{S}, \mathcal{C}) = 1] - 1/2\right|,$$

where the probabilities are taken over the random choice of k and the randomness of CH, W and the channel. The *random hiddentext attack (*RHA*) advantage* $\mathbf{Adv}^{\text{RHA}}_{W,\mathcal{S},\mathcal{C}}(\kappa)$ of a warden W is defined similarly with the difference that the messages given to the challenge oracle CH are chosen randomly instead of adversarially. This is a much weaker security requirement than CHA-security. Finally, for $\text{X} \in \{\text{CHA}, \text{RHA}\}$, the X-*insecurity* $\mathbf{InSec}^{\text{X}}_{\mathcal{S},\mathcal{C}}(q, t, \kappa)$ of a stegosystem $\mathcal{S}$ on the channel $\mathcal{C}$ is defined as

$$\mathbf{InSec}^{\text{X}}_{\mathcal{S},\mathcal{C}}(q, t, \kappa) = \max_W\{\mathbf{Adv}^{\text{X}}_{W,\mathcal{S},\mathcal{C}}(\kappa)\}.$$

The maximum is taken over all wardens W that make an expected number of $q(\kappa)$ queries and run in expected time $t(\kappa)$. We say that $\mathcal{S}$ is X-ϵ-*secure*, if $\mathbf{InSec}^{\text{X}}_{\mathcal{S},\mathcal{C}}(q, t, \kappa) \leq \epsilon(\kappa)$ for all polynomials q and t and X-*secure* if it is X-negl-secure for a negligible function negl.

2.2 Cryptographic Primitives

We recall briefly the definitions of the following three cryptographic primitives and the known relationships between them. For exact definitions see e.g. the literature quoted below.

One-Way Function. A polynomial time computable function $F\colon \{0,1\}^* \to \{0,1\}^*$ is called a *one-way function*, if every algorithm (inverter) upon input $F(x)$ fails to produce an element x' such that $F(x') = F(x)$.

Signature Scheme. A *signature scheme* $\mathcal{SIG}$ consists of a probabilistic *key-generation* algorithm G, that produces a secret key and a public key, a probabilistic *signing* algorithm S, that takes the secret key, a message and produces a signature for the message and a deterministic *verifying* algorithm V, that takes the public key and tests whether a message-signature pair is valid. An attacker either gets random valid message-signature pairs (random-message attack (RMA)) or can produce valid signatures for chosen messages (chosen-message attack (CMA)). Its goal is to produce a fresh message-signature pair.

Symmetric Encryption Scheme. A *symmetric encryption scheme* $\mathcal{SES}$ consists of an *encryption* algorithm *ENC*, which takes a secret key and a plain text and produces a cyphertext. This cyphertext can be decoded by the *decryption* algorithm *DEC* with the help of the same secret key. An attacker is given access to an oracle, which either encrypts a message chosen by the attacker (the *real* message) or gives a totally random cyphertext (the *random* message). The goal of the attacker is to distinguish those cases. We denote the *advantage* of an attacker A to distinguish real messages from random ones (ROR) on a symmetric encryption scheme $\mathcal{SES}$ with key length κ by $\mathbf{Adv}^{\text{ROR}}_{\mathcal{SES},A}(\kappa)$. Also, the probability that the decrypted message does not equal the original message must be negligible.

There is a deep connection between those primitives, as all of them are equivalent to each other. The groundbreaking works [3, 13, 17, 19, 32] imply the following:

▶ **Theorem 8** (informal). *One-Way functions exists ⇔ RMA-secure signature schemes exists ⇔ CMA-secure signature schemes exists ⇔ secure symmetric encryption schemes exist*

In Section 3, we construct an RMA-forger on a special signature scheme $\widehat{\mathcal{SIG}}$, that is "complete" for all signature schemes, i.e., if $\widehat{\mathcal{SIG}}$ is insecure, every signature scheme is insecure. The construction of such a complete signature scheme relies on the following theorem of Levin which states the existence of a complete one-way function $\widehat{F}$:

▶ **Theorem 9** (Levin [30]). *The function $\widehat{F}$ is a one-way function iff one-way functions exist.*

Combining Theorem 8 and Theorem 9, we get the following corollary needed to construct the "complete" signature scheme $\widehat{\mathcal{SIG}}$:

▶ **Corollary 10.** *The signature scheme $\widehat{\mathcal{SIG}}$ is RMA-secure iff one-way functions exist.*

3 A Channel $\mathcal{C}$ such that Efficient Steganography on $\mathcal{C}$ Does Imply the Non-existence of One-way Functions

The main result of this section, Corollary 14, says that for the widely used channel specified by a signature scheme protocol, secure and efficient steganography implies that one-way functions do not exist. Then we show that our construction can be generalized for more channels. We will only work with RHA-secure stegosystems in this section, as impossibility results upon this weaker notion imply the same results for CHA-secure stegosystems.

Our first technical goal is to formalize the following intuition: A secure and reliable stegosystem for a channel $\mathcal{C}$ must (a) have negligible probability of producing documents outside of $\text{supp}(\mathcal{C}_{h,n})$ and (b) be able to generate new documents out of the sampled documents. These properties have been formulated first in [10] for universal stegosystems.

We start with showing that the probability that the output of a secure stegosystem is not in the support of the channel is small (under the assumption that Warden can efficiently test whether a document belongs to the support of the channel). Before, let us introduce an auxiliary notion of a *membership-testable channel with confidence parameter* ν: We say that $\mathcal{C}$ is membership-testable with confidence parameter ν if there exists a probabilistic polynomial time algorithm, call it Test, which takes a polynomial number $\vec{x} = x_1, x_2, \ldots, x_q$ of documents such that $\mathcal{C}_{x_1 x_2 \ldots x_{i-1}}(x_i) > 0$ for every $i \geq 1$ and a document x and it either returns 1 or 0 such that the probability $\Pr_{\vec{x} \leftarrow \text{supp}(\mathcal{C}_{\varnothing,n})}[\text{Test}(\vec{x}, x) = 1]$ is $\geq 1 - \nu$, if $x \in \text{supp}(\mathcal{C}_{\vec{x},n})$ and $\leq \nu$ otherwise.

▶ **Lemma 11.** *Let $\mathcal{S} = [SE, SD]$ be a stegosystem for the message space $\{\mathcal{M}_n\}_{n\in\mathbb{N}}$ with document length n and output length ℓ for the channel $\mathcal{C}$ such that $\mathcal{S}$ is* RHA-ϵ*-secure.*

Furthermore, let $\mathcal{C}$ be membership-testable with parameter ν. Then for all $\kappa \in \mathbb{N}$, $m \in \operatorname{supp}(\mathcal{M}_\kappa)$, histories h, and all $i = 1, \ldots, \ell(\kappa)$, it holds for $d_i = SE^{\mathcal{C}}(k, m, h)_i$ that $\Pr_{k \leftarrow \{0,1\}^\kappa}[d_i \notin \operatorname{supp}(\mathcal{C}_{hd_1 d_2 \ldots d_{i-1}, n(\kappa)})] \leq \epsilon(\kappa) + 2\nu$.

Next, we will prove that, as long as the support of $\mathcal{C}_{h,n}$ is large enough, a reliable stegosystem needs to produce non-seen examples of $\operatorname{supp}(\mathcal{C}_{h,n})$. Intuitively, we need to embed $|\operatorname{supp}(\mathcal{M}_n)| \approx 2^n$ messages (hereby creating at least 2^n different documents) while we only have access to $\mathsf{pol}(n)$ example documents. Note that for a rate-efficient polynomial time stegosystem, the term $\frac{\log |\operatorname{supp}(\mathcal{M}_\kappa)|}{\ell(\kappa)} = b(\kappa)$ is of the form κ^α for a $\alpha > 0$ and thus the term $\frac{q(\kappa)^{\ell(\kappa)}}{|\operatorname{supp}(\mathcal{M}_\kappa)|} = \frac{q(\kappa)^{\ell(\kappa)}}{2^{b(\kappa)\cdot\ell(\kappa)}} = \left(\frac{q(\kappa)}{2^{b(\kappa)}}\right)^{\ell(\kappa)} \approx \left(\frac{\mathsf{pol}(\kappa)}{2^{\kappa^\alpha}}\right)^{\ell(\kappa)}$ is negligible.

▶ **Lemma 12.** *Let $\mathcal{S} = [SE, SD]$ be a ρ-reliable stegosystem for the message space $\{\mathcal{M}_n\}_{n\in\mathbb{N}}$ with sample complexity c and output length ℓ for the channel $\mathcal{C}$. Then for every κ, the probability that the encoder SE produces a cover-document, which was not provided by the channel oracle, is at least $1 - \rho(\kappa) - \frac{q(\kappa)^{\ell(\kappa)}}{|\operatorname{supp}(\mathcal{M}_\kappa)|}$.*

The proofs of the lemmas above are similar to those presented in [10], thus we skip them.

We will now combine the two lemmas in order to construct an attacker to a signature scheme. For a signature scheme $\mathcal{SIG} = [G, S, V]$, define the channel $\mathcal{C}_{\mathcal{SIG}}$ with probability distributions $\mathcal{C}_{h,n}$ as follows: If h is the empty history $\varnothing$, the probability distribution $\mathcal{C}_{\varnothing,n}$ is the uniform distribution on all public keys generated by $G(1^n)$. If $(pk, sk) \in \operatorname{supp}(G(1^n))$, the probability distribution $\mathcal{C}_{pk,n}$ is then created by the following experiment:

Distribution of $\mathcal{C}_{pk,n}$

1. $m \leftarrow \mathcal{M}_n^{\text{sig}}$; $\sigma \leftarrow S(sk, m)$; **return** (m, σ)

Furthermore, for every $i \geq 1$ and every series of valid (with respect to (pk, sk)) message-signature pairs $(m_1, \sigma_1), (m_2, \sigma_2) \ldots$ the distribution $\mathcal{C}_{pk(m_1,\sigma_1)(m_2,\sigma_2)\ldots(m_i,\sigma_i),n}$ is also equal to $\mathcal{C}_{pk,n}$. Note that $\mathcal{C}_{\mathcal{SIG}}$ is membership-testable with confidence parameter 0 due to the public key. A similar technique was used by Dwork et al. [12] and later by Ullman [34] in the context of differential privacy [11]. They prove that a certain class of databases exists such that any algorithm for a given set of counting queries is either not differentially private or inaccurate.

▶ **Theorem 13.** *Let $\mathcal{SIG} = [G, S, V]$ be a signature scheme. If there exists a polynomial time stegosystem $\mathcal{S} = [SE, SD]$ for $\mathcal{C}_{\mathcal{SIG}}$ for the message space $\{\mathcal{M}_n\}_{n\in\mathbb{N}}$ with rate b, output length ℓ and sampling complexity q such that $\mathcal{S}$ is* RHA-ϵ*-secure and ρ-reliable on $\mathcal{C}_{\mathcal{SIG}}$, then there exists an efficient forger on $\mathcal{SIG}$ with advantage at least $1 - \epsilon(\kappa) - \rho(\kappa) - \frac{q(\kappa)^{\ell(\kappa)}}{|\operatorname{supp}(\mathcal{M}_\kappa)|}$ for every κ.*

Combining Theorem 13 and Corollary 10 with $\widehat{\mathcal{SIG}}$, we obtain the following result that directly implies Theorem 3.

▶ **Corollary 14.** *The existence of a secure, reliable and rate-efficient polynomial time stegosystem on the channel $\mathcal{C}_{\widehat{\mathcal{SIG}}}$ implies that one-way functions do not exist.*

In the rest of this section we show that the proof of Theorem 13 can be generalized to more channels if they can express the signature scheme. Examples for such channels include satisfying assignments of 3-CNF formulas or satisfying assignments of monotone 2-CNF formulas. Our construction is inspired by the work of De et al. [9] who used a similar

technique to show that it is not possible to uniformly generate satisfying assignments to a 3-CNF formula if one is given polynomial many samples of satisfying assignments.

Let $\mathcal{SIG} = [G, S, V]$ be a signature scheme and $in(\kappa)$ be an upper bound on the size of every message-signature pair constructed by the signing algorithm S on security parameter κ. Let $\boldsymbol{B}$ be a function class of Boolean functions such that there is a polynomial time invertible Levin reduction $[A, B, C]$ from CIRCUIT-SAT (see e.g. [1] for a formal definition) to $\boldsymbol{B}$. Such a reduction transforms a circuit $\mathfrak{C}$ into a function $f := A(\mathfrak{C})$ and a satisfying assignment β of $\mathfrak{C}$ into a value $x := B(\mathfrak{C}, \beta)$ with $f(x) = 1$. Furthermore, every x' with $f(x') = 1$ can be transformed into a satisfying assignment $\beta' := C(f, x')$ of $\mathfrak{C}$. Moreover let γ: $A(\text{CIRCUIT-SAT}) \to \{0,1\}^*$ be a polynomial time encoding of the functions generated by the reduction such that $red(\kappa)$ is an upper bound on $|\gamma(A(\mathfrak{C}))|$, if $\mathfrak{C}$ has $in(\kappa)$ input gates. Furthermore, let $\mathcal{C}$ be a channel with probability distributions $\mathcal{C}_{h,\kappa}$ defined as follows. For the empty history $\varnothing$, the distribution $\mathcal{C}_{\varnothing,\kappa}$ is the uniform distribution on $\gamma(A(\{\mathfrak{C} \mid \mathfrak{C} \text{ has } in(\kappa) \text{ input gates}\})) \subseteq \{0,1\}^{red(\kappa)}$. For every history $h_0 = \gamma(A(\mathfrak{C}))$ the probability distribution $\mathcal{C}_{h_0,\kappa}$ is the uniform distribution on documents $x \in \{0,1\}^{in(\kappa)}$ with $A(\mathfrak{C})(x) = 1$. Furthermore, for every $i \geq 1$ and every series of documents $x_1, x_2, \ldots \in \{0,1\}^{in(\kappa)}$ with $A(\mathfrak{C})(x_j) = 1$ for every j, the probability distribution $\mathcal{C}_{h_0 x_1 x_2 \ldots x_i,\kappa}$ is also the uniform distribution on the documents $x \in \{0,1\}^{in(\kappa)}$ with $A(\mathfrak{C})(x) = 1$. Moreover, assume $\mathcal{C}$ is membership-testable with confidence parameter ν.

▶ **Theorem 15.** *Let $\mathcal{SIG}$ be a signature scheme and let $\mathcal{C}$ be a channel as defined above. Assume $\mathcal{S}$ is a polynomial time stegosystem for the message space $\{\mathcal{M}_n\}_{n\in\mathbb{N}}$ with transmission rate b, output length ℓ and sampling complexity q for $\mathcal{C}$ such that $\mathcal{S}$ is* RHA-ϵ*-secure and ρ-reliable on $\mathcal{C}$. Then for every κ, there is a polynomial forger for $\mathcal{SIG}(\kappa)$ with advantage at least $1 - \epsilon(\kappa) - 2\nu - \rho(\kappa) - \frac{q(\kappa)^{\ell(\kappa)}}{|\operatorname{supp}(\mathcal{M}_\kappa)|}$.*

4 A Channel $\mathcal{C}$ such that Efficient Steganography on $\mathcal{C}$ Does Imply the Existence of One-way Functions

We will now show a channel $\mathcal{C}$ such that secure and reliable steganography on it implies the existence of one-way functions (this will follow from the theorem below and Theorem 8). The channel is assumed to be *efficiently sampleable*, i.e. such for which a polynomial time algorithm simulating sampling from $\mathcal{C}$ exists. Then a straightforward argument implies the following equivalences between steganography and cryptography.

▶ **Theorem 16.** *Let $\mathcal{C}$ be a channel with $\mathcal{C}_{h,n} = \mathcal{C}_{h',n} := \mathcal{C}_n$ for all histories h, h' and assume $\mathcal{C}$ is efficiently sampleable. If there exists a secure, reliable, and rate-efficient (polynomial time) stegosystem $\mathcal{S} = [SE, SD]$ for the channel $\mathcal{C}$ with message space $\{\mathcal{M}_n\}_{n\in\mathbb{N}}$, then there exists a secure symmetric encryption scheme $\mathcal{SES}$ for the plaintexts $\{\mathcal{M}_n^{\text{plain}}\}_{n\in\mathbb{N}}$ with $\mathcal{M}_n^{\text{plain}} = \mathcal{M}_n$ and cyphertexts $\{\mathcal{M}_n^{\text{cypher}}\}_{n\in\mathbb{N}}$ with $\mathcal{M}_\kappa^{\text{cypher}} = \mathcal{C}_{n(\kappa)}^{\ell(\kappa)}$.*

▶ **Theorem 17.** *Let $\mathcal{SES}$ be a secure symmetric encryption scheme on plaintexts $\{\mathcal{M}_n^{\text{plain}}\}_{n\in\mathbb{N}}$ and cyphertexts $\{\mathcal{M}_n^{\text{cypher}}\}_{n\in\mathbb{N}}$. Let $\mathcal{C}$ be a channel with the documents $\operatorname{supp}(\mathcal{M}_n^{\text{cypher}})$ and $\mathcal{C}_{h,n} = \mathcal{M}_n^{\text{cypher}}$ for every h. There exists a secure, reliable, and rate-efficient (polynomial time) stegosystem $\mathcal{S}$ for $\mathcal{C}$ with message space $\{\mathcal{M}_n\}_{n\in\mathbb{N}}$ with $\mathcal{M}_n = \mathcal{M}_n^{\text{plain}}$.*

Thus, reasonable steganography on e.g. the channel $\mathcal{C}_n$ that is the uniform distribution on $\{0,1\}^n$, is equivalent to the existence of one-way functions. This proves Theorem 5.

5 Conclusions and Further Work

We have proved that steganography and cryptography are somehow orthogonal to each other. To show this statement, we constructed a specific channel based upon secure signature schemes and proved that every rate-efficient stegosystem on this channels breaks the security of the signature scheme. By using a universal one-way function due to Levin, we were then able to show that the existence of such a rate-efficient stegosystem implies that one-way functions do not exist. This is a generalization of the result of Dedić et al. [10], who only proved the existence of a *family* of channels $\mathcal{F}$ such that the existence of a rate-efficient stegosystem that works for *every* channel in $\mathcal{F}$ implies the non-existence of one-way functions. We thus proved that there is a channel $\mathcal{C}_1$ such that rate-efficient steganography on $\mathcal{C}_1$ implies the non-existence of one-way functions. On the other hand, we also gave a simple channel $\mathcal{C}_2$ and proved that rate-efficient steganography on $\mathcal{C}_2$ implies the existence of one-way functions.

The existence of those channels thus implies that statements of the form "Steganography is Cryptography" or "Steganography implies Cryptography" are wrong in this universality. Furthermore, it proves that the communication channel is a fundamental object in steganography and can not be ignored. In order to explore the fascinating connection between steganography and cryptography, it would be interesting to broaden our understanding of the influence of the communication channels. The work of Liśkiewicz et al. [31] already showed that knowledge or ignorance about some aspect of the channels has a significant impact on the steganographic setting.

References

1 Sanjeev Arora and Boaz Barak. *Computational complexity: a modern approach*. Cambridge University Press, 2009.

2 Michael Backes and Christian Cachin. Public-key steganography with active attacks. In *Theory of Cryptography Conference (TCC)*, pages 210–226. Springer, 2005.

3 Mihir Bellare, Anand Desai, Eron Jokipii, and Phillip Rogaway. A concrete security treatment of symmetric encryption. In *Foundations of Computer Science (FOCS)*, pages 394–403. IEEE, 1997.

4 Sebastian Berndt and Maciej Liśkiewicz. Provable secure universal steganography of optimal rate. In *Information Hiding and Multimedia Security (IH&MMSec)*. ACM, 2016.

5 Christian Cachin. An information-theoretic model for steganography. *Information and Computation*, 192(1):41–56, 2004. doi:10.1016/j.ic.2004.02.003.

6 Nishanth Chandran, Vipul Goyal, Rafail Ostrovsky, and Amit Sahai. Covert multi-party computation. In *Foundations of Computer Science (FOCS)*, pages 238–248. IEEE, 2007.

7 Chongwon Cho, Dana Dachman-Soled, and Stanisław Jarecki. Efficient concurrent covert computation of string equality and set intersection. In *Topics in Cryptology-CT-RSA*, pages 164–179. Springer, 2016.

8 Pedro Comesaña and Fernando Pérez-González. On the capacity of stegosystems. In *Multimedia & Security (MMSec)*, pages 15–24. ACM, 2007.

9 Anindya De, Ilias Diakonikolas, and Rocco A. Servedio. Learning from satisfying assignments. In *Symp. on Discrete Algorithms (SODA)*, pages 478–497. ACM-SIAM, 2015.

10 Nenad Dedić, Gene Itkis, Leonid Reyzin, and Scott Russell. Upper and lower bounds on black-box steganography. *Journal of Cryptology*, 22(3):365–394, 2009.

11 Cynthia Dwork. Differential privacy. In *Automata, Languages and Programming (ICALP)*, pages 1–12. Springer, 2006.

12 Cynthia Dwork, Moni Naor, Omer Reingold, Guy N. Rothblum, and Salil P. Vadhan. On the complexity of differentially private data release: efficient algorithms and hardness results. In *Symp. on Theory of Computing (STOC)*, pages 381–390, 2009.

13 Shimon Even, Oded Goldreich, and Silvio Micali. On-line/off-line digital signatures. *Journal of Cryptology*, 9(1):35–67, 1996.

14 Nelly Fazio, Antonio R Nicolosi, and Irippuge Milinda Perera. Broadcast steganography. In *Topics in Cryptology–CT-RSA 2014*, pages 64–84. Springer, 2014.

15 Tomás Filler, Andrew D. Ker, and Jessica J. Fridrich. The square root law of steganographic capacity for markov covers. In *Media Forensics and Security I, part of the IS&T-SPIE Electronic Imaging Symposium*, page 725408, 2009.

16 Jessica Fridrich. *Steganography in digital media: principles, algorithms, and applications*. Cambridge University Press, 2009.

17 Oded Goldreich, Shafi Goldwasser, and Silvio Micali. How to construct random functions. *Journal of the ACM (JACM)*, 33(4):792–807, 1986.

18 Vipul Goyal and Abhishek Jain. On the round complexity of covert computation. In *Symp. on Theory of Computing (STOC)*, pages 191–200. ACM, 2010.

19 Johan Håstad, Russell Impagliazzo, Leonid A. Levin, and Michael Luby. A pseudorandom generator from any one-way function. *SIAM Journal on Computing*, 28(4):1364–1396, 1999.

20 Nicholas Hopper. On steganographic chosen covertext security. In *Automata, Languages and Programming (ICALP)*, volume 3580, pages 311–323. Springer, 2005.

21 Nicholas J Hopper. Toward a theory of steganography. Technical report, Technical Report CMU-CS-04-157, Carnegie Mellon Univ., 2004.

22 Nicholas J. Hopper, John Langford, and Luis von Ahn. Provably secure steganography. In *Advances in Cryptology (CRYPTO)*, pages 77–92. Springer, 2002. doi:10.1007/3-540-45708-9_6.

23 Nicholas J. Hopper, Luis von Ahn, and John Langford. Provably secure steganography. *Computers, IEEE Transactions on*, 58(5):662–676, 2009.

24 Sune K Jakobsen and Claudio Orlandi. How to bootstrap anonymous communication. In *Conf. on Innovations in Theoretical Computer Science (ITCS)*, pages 333–344. ACM, 2016.

25 Stefan Katzenbeisser and Fabien A.P. Petitcolas. Defining security in steganographic systems. In *Electronic Imaging 2002*, pages 50–56. SPIE, 2002.

26 Andrew D. Ker, Patrick Bas, Rainer Böhme, Rémi Cogranne, Scott Craver, Tomáš Filler, Jessica Fridrich, and Tomáš Pevnỳ. Moving steganography and steganalysis from the laboratory into the real world. In *Information Hiding and Multimedia Security (IH&MMSec)*, pages 45–58. ACM, 2013.

27 Andrew D. Ker, Tomás Pevný, Jan Kodovský, and Jessica J. Fridrich. The square root law of steganographic capacity. In *Multimedia Security (MMSec)*, pages 107–116. ACM, 2008.

28 Aggelos Kiayias, Alexander Russell, and Narasimha Shashidhar. Key-efficient steganography. In *Information Hiding (IH)*, pages 142–159. Springer, 2012.

29 Tri Van Le and Kaoru Kurosawa. Bandwidth optimal steganography secure against adaptive chosen stegotext attacks. In *Information Hiding (IH)*, pages 297–313. Springer, 2006.

30 Leonid A. Levin. One way functions and pseudorandom generators. *Combinatorica*, 7(4):357–363, 1987.

31 Maciej Liśkiewicz, Rüdiger Reischuk, and Ulrich Wölfel. Security levels in steganography – insecurity does not imply detectability. *Electronic Colloquium on Computational Complexity (ECCC)*, 22(10), 2015.

32 Michael Luby and Charles Rackoff. How to construct pseudorandom permutations from pseudorandom functions. *SIAM Journal on Computing*, 17(2):373–386, 1988.

33 Boris Ryabko and Daniil Ryabko. Constructing perfect steganographic systems. *Information and Computation*, 209(9):1223–1230, 2011.

34 Jonathan Ullman. Answering $n^{2+o(1)}$ counting queries with differential privacy is hard. In *Symp. on Theory of Computing (STOC)*, pages 361–370. ACM, 2013.

35 Luis von Ahn, Nicholas Hopper, and John Langford. Covert two-party computation. In *Symposium on Theory of Computing (STOC)*, pages 513–522. ACM, 2005.

36 Luis von Ahn and Nicholas J Hopper. Public-key steganography. In *Advances in Cryptology (EUROCRYPT)*, pages 323–341. Springer, 2004.

37 Ying Wang and Pierre Moulin. Perfectly secure steganography: Capacity, error exponents, and code constructions. *Information Theory, IEEE Transactions on*, 54(6):2706–2722, 2008.

38 Elżbieta Zielińska, Wojciech Mazurczyk, and Krzysztof Szczypiorski. Trends in steganography. *Communications of the ACM*, 57(3):86–95, 2014.