Blackbox (im-)possibility of selective decommitments

Dennis Hofheinz (CWI, Amsterdam)

Abstract

The *selective decommitment problem* can be described as follows: assume an adversary receives a number of commitments and then may request openings of, say, half of them. Do the unopened commitments remain secure? Although this question arose more than twenty years ago, no satisfactory answer could be presented so far. We answer the question in several ways:

1. If simulation-based security is desired (i.e., if we demand that the adversary's output can be simulated by a machine that does not see the unopened commitments), then security is *not achievable* via blackbox reductions to standard cryptographic assumptions.
2. If only indistinguishability of the unopened commitments from random commitments is desired, then security is *not achievable* for perfectly binding commitment schemes, via blackbox reductions to standard cryptographic assumptions. However, statistically hiding schemes *do* achieve security in this sense, using a blackbox reduction.

Our results give an almost complete picture when and how security under selective openings can be achieved. Applications of our results include:

- Essentially, an encryption scheme *must* be non-committing in order to achieve provable security against an adaptive adversary.
- There are inherent limitations for "commit-choose-open" style zero-knowledge proofs.

On the technical side, we develop a technique to show very general impossibility results for blackbox proofs.

Keywords: cryptography, commitments, zero-knowledge, blackbox separations.

1 Introduction

Consider an adversary A that observes ciphertexts sent among parties in a multi-party cryptographic protocol. At some point, A may decide, based on the information he already observed, to corrupt, say, half of the parties. By this, A learns the secret keys of these parties, which allows him to open some of the observed ciphertexts. The question is: do the unopened ciphertexts remain secure? Since most encryption schemes actually constitute *commitments* to the respective messages, we can rephrase the question as what is known as the *selective decommitment problem*: assume A receives a number of commitments and then may request openings of half of them. Do the unopened commitments remain secure? According to Dwork et al. [10], this question arose already more than twenty years ago in the context of Byzantine agreement, but it is still relatively poorly understood. In particular, standard cryptographic techniques (e.g., guessing which commitments are opened, or hybrid arguments) fail to show that "ordinary" commitment security against a static adversary guarantees security under selective openings.[1] Even worse: no commitment scheme is known to be secure under selective openings.

Our work. We answer the selective decommitment problem in several ways. First, we consider what happens if "security of the unopened commitments" means that we require the existence of a simulator S, such that S essentially achieves what A does, only without seeing the unopened commitments in the first place. We call a commitment scheme which is secure in this sense *simulatable under selective openings.* We show that the answer to the selective decommitment problem is negative here: That is, we show that

[1] For instance, the probability to correctly guess an $n/2$-sized subset of n commitments is too small, and a hybrid argument would require some independence among the commitments, which we cannot assume in general.

no commitment scheme can be proven simulatable under selective openings via blackbox reductions from standard asssumptions. In particular, not even a statistically hiding commitment scheme can be proven simulatable under selective openings in a blackbox way.

We proceed to consider what happens if "security" means that A cannot distinguish the messages inside the unopened commitments from independent[2] messages. We call a commitment scheme which is secure in this sense *indistinguishable under selective openings.* We show that no perfectly binding commitment scheme can be proven indistinguishable under selective openings, via blackbox reductions from standard assumptions. However, we also show that statistically hiding commitment schemes *are* indistinguishable under selective openings. Hence, the existence of a simulator is a much harder requirement than indistinguishability.

Technically, we derive blackbox impossibility results in the style of Impagliazzo and Rudich [15], but we can deriver stronger claims, similar to Dodis et al. [9]. Concretely, we prove impossibility via $\forall\exists$semi-blackbox proofs from *any* computational assumption that can be formalized as an oracle $\mathcal{X}$ and a corresponding security propety $\mathcal{P}$ which the oracle satisfies. For instance, to model one-way permutations, $\mathcal{X}$ could be a truly random permutation and $\mathcal{P}$ could be the one-way game in which a PPT adversary tries to invert a random image. We emphasize that, disturbingly, our impossibility claim holds even if $\mathcal{P}$ models security under selective openings. In that case, however, a reduction will necessarily be non-blackbox, see Appendix D for a discussion.

Applications. We apply our results to the adaptively secure encryption example mentioned in the beginning, and to a special class of interactive proof systems. First, we comment that an adaptively secure encryption scheme must be non-committing, or rely on nonstandard techniques. Namely, whenever a committing (i.e., ciphertexts commit to messages) encryption scheme is adaptively secure, then it also is, interpreted as a commitment scheme, simulatable under selective openings. Our impossibility results show that hence, a committing encryption scheme cannot be proven adaptively secure via blackbox reductions from standard assumptions.

Second, we apply our results to "commit-choose-open" style interactive proof systems. Dwork et al. [10] prove that if the underlying commitment scheme of such a proof system is simulatable under selective openings, then the proof system is (weakly) zero-knowledge, even under parallel composition. Unfortunately, our impossibility result shows that commitment schemes simulatable under selective openings are very hard to construct. However, we show that if the underlying commitment scheme is indistinguishable under selective openings, then the proof system is witness-indistinguishable (a relaxation of zero-knowledge), also under parallel composition. Since we show that statistically hiding commitment schemes are in fact indistinguishable under selective openings, this demonstrates the usefulness of our definition.

Related work. The selective decommitment problem arises in particular in the encryption situation described above, and hence was recognized and mentioned in a number of works before (e.g., [4, 1, 5, 8, 6]). However, these works solved the problem by using (and, in fact, inventing) non-committing encryption, which circumvents the underlying commitment problem. (In a sense, we show that a non-committing property is a necessity in the encryption case.)

Dwork et al. [10] is, to the best of our knowledge, the only work that explicitly studies the selective decommitment problem. They prove that a commitment scheme which is simulatable under selective openings would have tremendous applications. In particular, it would imply the parallel composability of a certain class of zero-knowledge protocols, something very surprising in the light of Goldreich and Krawczyk [12]. (Unfortunately, we show in our work that such a scheme cannot be easily constructed.) They proceed to give positive results for substantially relaxed selective decommitment problems (essentially, they prove security when standard techniques can be applied, i.e., when the set of opened commitments can be guessed, or when the messages are independent). However, they leave the general question open.

Organization. After fixing some notation, we present in Section 3 our impossibility result for the simulation-based security definition of Dwork et al. [10]. We give an indistinguishability-based security definition, along with possibility and impossibility results in Section 4. In Appendix B and Appendix C, we consider applications of our results to encryption and interactive proof systems. We discuss the role of the computational

[2]"independent" can of course only mean "independent, conditioned on the already opened messages"

assumption in our impossibility results in Appendix D. Postponed proofs are found in Appendix E.

2 Preliminaries

Notation. Throughout the paper, $k \in \mathbb{N}$ denotes a security parameter. With growing k, attacks should be become harder, but we also allow schemes to be of complexity which is polynomial in k. A function $f = f(k)$ is called negligible if it vanishes faster than the inverse of any polynomial. That is, f is negligible iff $\forall c\, \exists k_0\, \forall k > k_0 : |f(k)| < k^{-c}$. If f is not negligible, we call f non-negligible. We say that f is overwhelming iff $1-f$ is negligible. We write $[n] := \{1, \ldots, n\}$. If $M = (M_i)_i$ is an indexed set, then we write $M_I := (M_i)_{i \in I}$.

Commitment schemes. In the spirit of Dwork et al. [10], we focus on noninteractive commitment schemes, but only to ease presentation. We stress that all our results also hold for interactive schemes (in which committing and opening are interactive processes). We will comment at the appropriate places on this.

A *(noninteractive) commitment scheme* $(\mathsf{Com}, \mathsf{Ver})$ is a pair of algorithms, such that for any $M \in \{0,1\}^k$,[3] algorithm $\mathsf{Com}(M)$ outputs a commitment *com* to M along with a decommitment *dec*. $\mathsf{Ver}(com, dec)$ outputs either a message M' (and we require $M' = M$ for honestly generated (com, dec)), or $\perp$. We require that $(\mathsf{Com}, \mathsf{Ver})$ is *hiding* in the sense that no PPT adversary A can distinguish, with non-negligible advantage, commitments to M_0 from commitments to M_1, where A may initially choose M_0 and M_1. If this holds for all (not necessarily PPT) A, then we call $(\mathsf{Com}, \mathsf{Ver})$ *statistically hiding*. Additionally, we require that $(\mathsf{Com}, \mathsf{Ver})$ is *binding*. This means that no PPT A can output, with non-negligible probability, (com, dec_1, dec_2) such that $\mathsf{Ver}(com, dec_1) = M_1 \neq M_2 = \mathsf{Ver}(com, dec_2)$. If for all A, the probability to output such (com, M_1, M_2) is even 0, then we call $(\mathsf{Com}, \mathsf{Ver})$ *perfectly binding*. More detailed definitions are in Appendix A.

Note that perfectly binding implies that *any* commitment *com* can only be opened to at most one value M. Perfectly binding commitment schemes can be achieved from any one-way permutation (e.g., Blum [3]). On the other hand, statistically hiding implies that for any $M_1, M_2 \in \{0,1\}^k$, the statistical distance between the respective commitments com_1 and com_2 is negligible. One-way functions suffice to implement statistically hiding commitment schemes (Haitner and Reingold [14]).

3 A simulation-based definition

Consider the following real security game: adversary A gets, say, n commitments, and then may ask for openings of some of them. The security notion of [10] requires that for any such A, there exists a simulator S that can approximate A's output. More concretely, for any relation R, we require that $R(M, out_A)$ holds about as often as $R(M, out_S)$, where $M = (M_i)_{i\in[n]}$ are the messages in the commitments, out_A is A's output, and out_S is S's output. Formally, we get the following definition (where henceforth, $\mathcal{I}$ will denote the set of "allowed" opening sets):

Definition 3.1 (Simulatable under selective openings/SIM-SO-COM)**.** *Let $n = n(k) > 0$, and let $\mathcal{I} = (\mathcal{I}_n)_n$ be a family of sets such that each $\mathcal{I}_n$ is a set of subsets of $[n]$. A commitment scheme $(\mathsf{Com}, \mathsf{Ver})$ is* simulatable under selective openings *(short* SIM-SO-COM*) iff for every PPT n-message distribution $\mathcal{M}$, every PPT relation R, and every PPT adversary $A = (A_1, A_2)$, there is a PPT simulator $S = (S_1, S_2)$, such that $\mathsf{Adv}^{\textsf{sim-so}}_{\mathsf{Com},\mathcal{M},A,S,R}$ is negligible. Here*

$$\mathsf{Adv}^{\textsf{sim-so}}_{\mathsf{Com},\mathcal{M},A,S,R} := \Pr\left[\mathsf{Exp}^{\textsf{sim-so-real}}_{\mathsf{Com},\mathcal{M},A,R} = 1\right] - \Pr\left[\mathsf{Exp}^{\textsf{sim-so-ideal}}_{\mathcal{M},S,R} = 1\right],$$

where $\mathsf{Exp}^{\textsf{sim-so-real}}_{\mathsf{Com},\mathcal{M},A,R}$ proceeds as follows:

1. *sample messages $M = (M_i)_{i\in[n]} \leftarrow \mathcal{M}$,*
2. *compute (de-)commitments $(com_i, dec_i) \leftarrow \mathsf{Com}(M_i)$ for $i \in [n]$,*
3. *run $(s, I) \leftarrow A_1(1^k, (com_i)_{i\in[n]})$ to get state information s and a set $I \in \mathcal{I}$,*

[3] We restrict to k-bit messages to ease presentation.

4. run $out_A \leftarrow A_2(s, (dec_i)_{i \in I})$*,*
5. output 1 iff $R(M, out_A)$*.*

On the other hand, $\mathsf{Exp}^{\text{sim-so-ideal}}_{\mathcal{M},S,R}$ *proceeds as follows:*

1. sample messages $M = (M_i)_{i \in [n]} \leftarrow \mathcal{M}$*,*
2. run $(s, I) \leftarrow S_1(1^k)$ *to get state information* s *and a set* $I \in \mathcal{I}$*,*
3. run $out_S \leftarrow S_2(s, (M_i)_{i \in I})$*,*
4. output 1 iff $R(M, out_S)$*.*

For interactive commitments, the $\mathsf{Exp}^{\text{sim-so-real}}_{\mathcal{M},A,R}$ experiment concurrently performs n commitment processes with A in step 3, and $|I|$ decommitment processes in step 4. Note that we opted not to give auxiliary input to the adversary. Such an auxiliary input is a common tool in cryptographic definitions to ensure some form of composability. Not giving the adversary auxiliary input only makes our negative results stronger. We stress, however, that our positive result (4.6) holds also for adversaries with auxiliary input.

Formalization of computational assumptions. Our first main result states that SIM-SO-COM security cannot be achieved via blackbox reductions from standard assumptions. In order to be able to consider such standard assumptions in a general way that allows to make statements even in the presence of "relativizing" oracles, we make the following definition:

Definition 3.2 (Property of an oracle)**.** *Let* $\mathcal{X}$ *be an oracle. Then a* property $\mathcal{P}$ *of* $\mathcal{X}$ *is a (not necessarily PPT) machine* $\mathcal{P}$ *that, after arbitrarily interacting with* $\mathcal{X}$ *and another machine* A*, finally outputs a bit* b*. For an adversary* A *(that may interact with* $\mathcal{X}$ *and* $\mathcal{P}$*), we define* A*'s* advantage against $\mathcal{P}$ *as*

$$\mathsf{Adv}^{\mathcal{P}}_A := \Pr\left[\mathcal{P} \textit{ outputs } b = 1 \textit{ after an interaction with } A\right] - 1/2.$$

Now $\mathcal{X}$ *is said to* satisfy *property* $\mathcal{P}$ *iff for all PPT adversaries* A*, we have that* $\mathsf{Adv}^{\mathcal{P}}_A$ *is negligible.*

Our definition is similar in spirit to "hard games" as used by Dodis et al. [9], but more general. We emphasize that $\mathcal{P}$ can *only* interact with $\mathcal{X}$ and A, but not with possible additional oracles. (See Appendix D for further discussion of properties of oracles, in particular their role in our proofs.) Intuitively, $\mathcal{P}$ acts as a challenger in the sense of a cryptographic security experiment. That is, $\mathcal{P}$ tests an adversary A whether A can "break" $\mathcal{X}$ in the intended way. We give an example, where "breaking" means "breaking $\mathcal{X}$'s one-way property".

Example. If $\mathcal{X}$ is a random permutation of $\{0,1\}^k$, then the following $\mathcal{P}$ models $\mathcal{X}$'s one-way property: $\mathcal{P}$ acts as a challenger that challenges A to invert a randomly chosen $\mathcal{X}$-image. Concretely, $\mathcal{P}$ initially chooses a random $Y \in \{0,1\}^k$ and sends Y to A. Upon receiving a guess $X \in \{0,1\}^k$ from A, $\mathcal{P}$ checks if $\mathcal{X}(X) = Y$. If yes, then $\mathcal{P}$ terminates with output $b = 1$. If $\mathcal{X}(X) \neq Y$, then $\mathcal{P}$ tosses an unbiased coin $b' \in \{0,1\}$ and terminates with output $b = b'$.

We stress that we only gain generality by demanding that $\Pr[1 \leftarrow \mathcal{P}]$ is close to $1/2$ (and not, say, negligible). In fact, this way indistinguishability-based games (such as, e.g., the indistinguishability of ciphertexts of an ideal cipher $\mathcal{X}$) can be formalized very conveniently. On the other hand, cryptographic games like the one-way game above can be formulated in this framework as well, by letting the challenger output $b = 1$ with probability $1/2$ when A fails.

On the role of property $\mathcal{P}$. Our upcoming results state the impossibility of (blackbox) security reductions, from essentially *any* computational assumption (i.e., property) $\mathcal{P}$. The obvious question is: what if the assumption already *is* an idealized commitment scheme secure under selective openings? The short answer is: "then the security proof will not be blackbox." We give a detailed explanation of what is going on in Appendix D.

Theorem 3.3 (First main result: impossibility of SIM-SO-COM, most general formulation)**.** *Let* $n = n(k)$ *be arbitrary, and let* $\mathcal{I} = (\mathcal{I}_n)_n$ *be arbitrary such that* $\mathcal{I}_n$ *is a set of subsets of* $[n]$ *and* $|\mathcal{I}_n|$ *is superpolynomial in* k.[4] *Let* $\mathcal{X}$ *be an oracle that satisfies property* $\mathcal{P}$*. Then there is a set of oracles relative to which* $\mathcal{X}$ *still satisfies property* $\mathcal{P}$*, but there exists no commitment scheme which is simulatable under selective openings.*

[4] e.g., one could think of $n = 2k$ and $\mathcal{I}_n = \{I \subseteq [n] \mid |I| = n/2\}$ here

Proof. First, let $\mathcal{RO}$ be an oracle for evaluating a random oracle (i.e., a random function $\{0,1\}^* \to \{0,1\}^k$). When writing $\mathcal{RO}(x_1,\dots,x_\ell)$, we assume that $\mathcal{RO}$'s input $x_1,\dots,x_\ell$ is encoded in a prefix-free way, such that all individual x_i can be efficiently reconstructed from $\mathcal{RO}$'s input. Furthermore, let $\mathcal{B}$ be the oracle that proceeds as follows:

1. Upon input $(\mathsf{Com},\mathsf{Ver},com)$, where $com = (com_i)_{i\in[n]}$, return a uniformly chosen $I \in \mathcal{I}$ and record $(\mathsf{Com},\mathsf{Ver},com,I)$.[5]
2. Upon input $(\mathsf{Com},\mathsf{Ver},com,dec_I)$ with $dec_I = (dec_i)_{i\in I}$ for a $(\mathsf{Com},\mathsf{Ver},com,I)$ which was previously recorded, verify using Ver that each dec_i is a valid opening of the respective com_i. If not, reject with output $\perp$. If yes, let M_i denote the message that com_i was opened to, and return the set of all $s \in \{0,1\}^{k/3}$ such that $M_i = \mathcal{RO}(\mathsf{Com},\mathsf{Ver},i,s)$ for all $i \in I$.

Now fix any commitment scheme $(\mathsf{Com}^*,\mathsf{Ver}^*)$ (that may use all the described oracles in its algorithms). Consider the n-message distribution $\mathcal{M}^* = \{(\mathcal{RO}(\mathsf{Com}^*,\mathsf{Ver}^*,i,s^*))_{i\in[n]}\}_{s\in\{0,1\}^{k/3}}$ (i.e., $\mathcal{M}^*$ chooses $s^* \in \{0,1\}^{k/3}$ uniformly and then sets $M_i^* = \mathcal{RO}(\mathsf{Com}^*,\mathsf{Ver}^*,i,s^*)$ for all i).

Lemma 3.4. *There is an adversary A that outputs $out_A = M^*$ with overwhelming probability in the real SIM-SO-COM experiment $\mathsf{Exp}^{\mathsf{sim\text{-}so\text{-}real}}_{\mathsf{Com}^*,\mathsf{Ver}^*,\mathcal{M},A,R}$. Here M^* denotes the full message vector sampled from $\mathcal{M}^*$ by the experiment.*

Proof. Let A be the SIM-SO-COM adversary on $(\mathsf{Com}^*,\mathsf{Ver}^*)$ that relays between its interface to the SIM-SO-COM experiment and $\mathcal{B}$ as follows:

1. Upon receiving $com^* = (com_i^*)_{i\in[n]}$ from the experiment, send $(\mathsf{Com}^*,\mathsf{Ver}^*,com^*)$ to $\mathcal{B}$.
2. Upon receiving $I^* \in \mathcal{I}$ from $\mathcal{B}$, send I^* to the SIM-SO-COM experiment.
3. Upon receiving $dec^*_{I^*} = (dec_i^*)_{i\in I^*}$ from the experiment, send $(\mathsf{Com}^*,\mathsf{Ver}^* com^*, dec^*_{I^*})$ to $\mathcal{B}$.
4. Finally, upon receiving a singleton set $\{s^*\}$ from $\mathcal{B}$, return $out_A = (\mathcal{RO}(\mathsf{Com}^*,\mathsf{Ver}^*,i,s^*))_{i\in[n]}$. If $\mathcal{B}$ returns a set of larger size, return $out_A = \perp$.

(This adversary is straightforwardly split into two PPT parts A_1 and A_2 as required for the SIM-SO-COM experiment.) By construction of $\mathcal{M}^*$ and $\mathcal{B}$, it is clear that $out_A = M^*$ unless $\mathcal{B}$ returns multiple s (which happens only with negligible probability by a counting argument). $\square$

Lemma 3.5. *Any given PPT simulator S will output $out_S = M^*$ in the ideal SIM-SO-COM experiment $\mathsf{Exp}^{\mathsf{sim\text{-}so\text{-}ideal}}_{\mathcal{M},S,R}$ only with negligible probability.*

Proof. Fix a PPT S. We claim that in the ideal SIM-SO-COM experiment, S has a view that is almost statistically independent of s^*, and hence will output $out_S = M^*$ only with negligible probability. To show the claim, denote by I^* the subset that S submits to the SIM-SO-COM experiment, and by $M^*_{I^*}$ the messages that S receives back. Denote by $\mathsf{Com}^j, \mathsf{Ver}^j, I^j, M^j_{I^j}$ the corresponding values used in S's j-th query to $\mathcal{B}$. We first define and bound a number of "bad" events:

- bad_{coll} occurs iff S submits an opened M_i^j to $\mathcal{B}$ for which there are two distinct $s_1, s_2 \in \{0,1\}^{k/3}$ with $\mathcal{RO}(\mathsf{Com}^j,\mathsf{Ver}^j,i,s_1) = M_i^j = \mathcal{RO}(\mathsf{Com}^j,\mathsf{Ver}^j,i,s_2)$.
- bad_{img} occurs iff S submits an opened M_i^j to $\mathcal{B}$ for which an s with $M_i^j = \mathcal{RO}(\mathsf{Com}^j,\mathsf{Ver}^j,i,s)$ exists, but M_i^j has not been obtained through an explicit $\mathcal{RO}$-query (by either S or the SIM-SO-COM experiment).
- bad_{bind} occurs iff $(\mathsf{Com}^j,\mathsf{Ver}^j,I^j,M^j_{I^j}) = (\mathsf{Com}^*,\mathsf{Ver}^*,I^*,M^*_{I^*})$ for some j.
- $\mathsf{bad} := \mathsf{bad}_{coll} \vee \mathsf{bad}_{img} \vee \mathsf{bad}_{bind}$.

These events occur only with negligible probability: informally, bad_{coll} implies a collision among $2^{k/3}$ uniformly distributed k-bit values, which is ruled out by a birthday bound; bad_{img} means that S guessed an element of a very sparse set; bad_{bind} means that S broke $(\mathsf{Com}^*,\mathsf{Ver}^*)$'s binding property. A detailed proof appears in Appendix E.

Now consider the following oracle $\mathcal{B}'$ which is almost identical to $\mathcal{B}$:

[5] Com and Ver denote descriptions of circuits (with access to all oracles) for commitment and verification algorithms. This has the effect that these algorithms will be PPT whenever the entity that uses $\mathcal{B}$ is PPT.

1. Upon input $(\mathsf{Com}, \mathsf{Ver}, com)$, where $com = (com_i)_{i\in[n]}$, return a uniformly chosen $I \in \mathcal{I}$ and record $(\mathsf{Com}, \mathsf{Ver}, com, I)$.
2. Upon input $(\mathsf{Com}, \mathsf{Ver}, com, dec_I)$ with $dec_I = (dec_i)_{i\in I}$ for a $(\mathsf{Com}, \mathsf{Ver}, com, I)$ which was previously recorded, verify using Ver that each dec_i is a valid opening of the respective com_i. If not, reject with output $\perp$. If yes, let M_i denote the message that com_i was opened to. If every M_i is the result of an $\mathcal{RO}(\mathsf{Com}, \mathsf{Ver}, i, s)$-query of S (for the same $s \in \{0,1\}^{k/3}$), then output $\{s\}$. Otherwise, output $\emptyset$.

By construction, the output of $\mathcal{B}$ and $\mathcal{B}'$ can differ only if

- there are multiple s with $M_i = \mathcal{RO}(\mathsf{Com}, \mathsf{Ver}, i, s)$ for some $i \in I$, or
- for some $i \in I$, M_i is not the result of an explicit $\mathcal{RO}$-query of S, but there exists an s with $M_i = \mathcal{RO}(\mathsf{Com}, \mathsf{Ver}, i, s)$ for all $i \in I$.

Assume that event bad does not occur. Then $\neg\mathsf{bad}_{coll}$ ensures that no multiple s with $M_i = \mathcal{RO}(\mathsf{Com}, \mathsf{Ver}, i, s)$ exist, and $\neg\mathsf{bad}_{img}$ ensures that all M_i have been explicitly queried as $M_i = \mathcal{RO}(\mathsf{Com}, \mathsf{Ver}, i, s)$ by either S or the SIM-SO-COM experiment. Now since the SIM-SO-COM experiment makes only queries of the form $M_i^* = \mathcal{RO}(\mathsf{Com}^*, \mathsf{Ver}^*, i, s^*)$, this means that $\mathcal{B}$ and $\mathcal{B}'$ can only differ if $(\mathsf{Com}, \mathsf{Ver}) = (\mathsf{Com}^*, \mathsf{Ver}^*)$, and if M_I contains some M_i from $M^*_{I^*}$. On the other hand, $\neg\mathsf{bad}_{bind}$ implies that then, M_I must also contain some $M_{i'}$ not contained in $M^*_{I^*}$. By $\neg\mathsf{bad}_{img}$, then $M_{i'}$ must have been explicitly queried by S through $M_{i'} = \mathcal{RO}(\mathsf{Com}^j, \mathsf{Ver}^j, i', s^*)$, for the *same* s^* as chosen by the SIM-SO-COM experiment to generate $M_i^* = \mathcal{RO}(\mathsf{Com}^*, \mathsf{Ver}^*, i, s^*)$.

In other words, assuming $\neg\mathsf{bad}$, in order to detect a difference between $\mathcal{B}$ and $\mathcal{B}'$, S must already have guessed the hidden s^* used in the SIM-SO-COM experiment. In particular, since up to that point, oracles $\mathcal{B}$ and $\mathcal{B}'$ behave identically, and S can simulate $\mathcal{B}'$ internally, S can either extract the hidden s^* from the SIM-SO-COM experiment with oracles $\mathcal{RO}$ and $\mathcal{X}$ alone, or not at all. However, since we defined $\mathcal{RO}$ independently and after $\mathcal{X}$, these oracles are independent. Hence, using $\mathcal{RO}$ and $\mathcal{X}$ alone, the view of S is independent of s^* unless S explicitly makes a $\mathcal{RO}$-query involving s^*. Since $s^* \in \{0,1\}^{k/3}$ is uniformly chosen from a suitably large domain, and bad occurs with negligible probability, we get that S's view is almost statistically independent of s^*. Hence, S can produce $out_S = M^*$ only with negligible probability. □

Taking things together, this shows that $\mathsf{Adv}^{\mathsf{sim\text{-}so}}_{\mathsf{Com}^*,\mathcal{M}^*,A,S,R}$ is overwhelming for the relation $R(x,y) :\Leftrightarrow x = y$, the described A, and any PPT S. Hence $(\mathsf{Com}^*, \mathsf{Ver}^*)$ is not SIM-SO-COM secure. It remains to argue that in the described computational world, one-way permutations exist.

Lemma 3.6. *$\mathcal{X}$ satisfies $\mathcal{P}$.*

Proof. Assume a PPT adversary A on $\mathcal{X}$'s property $\mathcal{P}$. Since $\mathcal{X}$ and $\mathcal{P}$ do not query $\mathcal{B}$ or $\mathcal{RO}$, these latter two oracles do not help A, in the following sense. Namely, A can break property $\mathcal{P}$ without oracles $\mathcal{RO}$ and $\mathcal{B}$, and use internal simulations of these oracles instead. This achieves the same view for A, $\mathcal{X}$, and $\mathcal{P}$. To see this, note that $\mathcal{RO}$ never queries $\mathcal{X}$. Furthermore, $\mathcal{B}$ queries $\mathcal{X}$ at most a polynomial number of times (for checking the validity of the decommitments dec_I according to Ver). Hence both of these simulations inside A are efficient in terms of $\mathcal{X}$-queries. In fact, using lazy sampling techniques for $\mathcal{RO}$, both simulations can be made PPT. (This includes $\mathcal{B}$'s inversion of $\mathcal{RO}$, since we simulate $\mathcal{B}$ and $\mathcal{RO}$ at the same time.)

So without loss of generality, we can assume that A only uses $\mathcal{X}$-queries when interacting with $\mathcal{P}$. Since we assumed that $\mathcal{P}$ holds in the standard model (i.e., without any auxiliary oracles), $\mathcal{P}$ will hence also hold in presence of $\mathcal{B}$ and $\mathcal{RO}$. □

This concludes the proof of 3.3. □

The following corollary provides an instantiation of 3.3 for a number of standard cryptographic primitives.

Corollary 3.7 (First main result: impossibility of SIM-SO-COM). *Let n and $\mathcal{I}$ as in 3.3. Then no commitment scheme can be proven simulatable under selective openings via a $\forall\exists$semi-blackbox[6] reduction from one or*

[6] A $\forall\exists$semi-blackbox reduction is one where scheme and adversary must use the assumption as a black box, but the security reduction itself is arbitrary. See Reingold et al. [19] for definitions and a classification of blackbox reductions.

more of the following primitives: one-way functions, one-way permutations, trapdoor one-way permutations, IND-CCA secure public key encryption.

The corollary is a special case of 3.3. For instance, to show 3.7 for one-way permutations, one can use the example $\mathcal{X}$ and $\mathcal{P}$ from above: $\mathcal{X}$ is a random permutation of $\{0,1\}^k$, and $\mathcal{P}$ models the one-way experiment with $\mathcal{X}$. Clearly, $\mathcal{X}$ satisfies $\mathcal{P}$, and so we can apply 3.7. This yields impossibility of relativizing proofs for SIM-SO-COM security from one-way permutations. We can get impossibility for ∀∃semi-blackbox reductions using a standard "embedding" technique, cf. Simon [20], Reingold et al. [19]. The other cases are similar. Note that while it is generally not easy to even give a candidate for a cryptographic primitive in the standard model, it is easy to construct an idealized, say, encryption scheme in oracle form.

Generalizations. First, 3.7 constitutes merely an example instantiation of the much more general 3.3. Also, the proof of 3.3 generalizes to interactive commitment schemes in a natural way: in this case, $(\mathsf{Com}, \mathsf{Ver})$ denotes (a description of) interactive machines, and $\mathcal{B}$ performs a commitment/decommitment process as specified by $(\mathsf{Com}, \mathsf{Ver})$. Similarly, A only relays messages used during commitment and decommitment between $\mathcal{B}$ and the SIM-SO-COM experiment. The proof also holds for a relaxation of SIM-SO-COM security considered by Dwork et al. [10], Definition 7.3, where adversary and simulator approximate a function of the message vector.

4 An indistinguishability-based definition

Motivated by the above impossibility result, we relax 3.1 as follows:

Definition 4.1 (Indistinguishable under selective openings/IND-SO-COM)**.** *Let $n = n(k) > 0$, and let $\mathcal{I} = (\mathcal{I}_n)_n$ be a family of sets such that each $\mathcal{I}_n$ is a set of subsets of $[n]$. A commitment scheme* $(\mathsf{Com}, \mathsf{Ver})$ *is* indistinguishable under selective openings *(short* IND-SO-COM*) iff for every PPT n-message distribution $\mathcal{M}(\cdot)$, and every PPT adversary $A = (A_1, A_2)$, we have that $\mathsf{Adv}^{\mathsf{ind\text{-}so}}_{\mathsf{Com},\mathcal{M},A}$ is negligible. Here*

$$\mathsf{Adv}^{\mathsf{ind\text{-}so}}_{\mathsf{Com},\mathcal{M},A} := \Pr\left[\mathsf{Exp}^{\mathsf{ind\text{-}so\text{-}real}}_{\mathsf{Com},\mathcal{M},A} = 1\right] - \Pr\left[\mathsf{Exp}^{\mathsf{ind\text{-}so\text{-}ideal}}_{\mathsf{Com},\mathcal{M},A} = 1\right],$$

where $\mathsf{Exp}^{\mathsf{ind\text{-}so\text{-}real}}_{\mathsf{Com},\mathcal{M},A}$ proceeds as follows:

1. *sample messages $M = (M_i)_{i\in[n]} \leftarrow \mathcal{M}$,*
2. *compute (de-)commitments $(com_i, dec_i) \leftarrow \mathsf{Com}(M_i)$ for $i \in [n]$,*
3. *run $(s, I) \leftarrow A_1(1^k, (com_i)_{i\in[n]})$ to get state information s and a set $I \in \mathcal{I}$,*
4. *run $b \leftarrow A_2(s, (dec_i)_{i\in I}, M)$ to obtain a guess bit b,*
5. *output b.*

On the other hand, $\mathsf{Exp}^{\mathsf{ind\text{-}so\text{-}ideal}}_{\mathsf{Com},\mathcal{M},A}$ proceeds as follows:

1. *sample messages $M = (M_i)_{i\in[n]} \leftarrow \mathcal{M}$,*
2. *compute (de-)commitments $(com_i, dec_i) \leftarrow \mathsf{Com}(M_i)$ for $i \in [n]$,*
3. *run $(s, I) \leftarrow A_1(1^k, (com_i)_{i\in[n]})$ to get state information s and a set $I \in \mathcal{I}$,*
4. *sample $M' \leftarrow \mathcal{M} \mid M_I$, i.e., sample a fresh message M' from $\mathcal{M}$ with $M'_I = M_I$,*
5. *run $b \leftarrow A_2(s, (dec_i)_{i\in I}, M')$ to obtain a guess bit b,*
6. *output b.*

As obvious, for interactive commitments, both experiments perform commitment and decommitment processes with A.

On the conditioned distribution $\mathcal{M} \mid M_I$. We stress that, depending on $\mathcal{M}$, it may be computationally hard to sample $M' \leftarrow \mathcal{M} \mid M_I$, even if (the unconditioned) $\mathcal{M}$ is PPT. This might seem strange at first and inconvenient when *applying* the definition in some larger reduction proof. However, there simply seems to be no other way to capture indistinguishability, since the set of opened commitments depends on the commitments themselves. In particular, in general we cannot predict which commitments the adversary wants opened, and then, say, substitute the not-to-be-opened commitments with random commitments.

What we chose to do instead is to give the adversary either the full message vector, or an independent message vector which "could be" the full message vector, given the opened commitments. We believe that this is the canonical way to capture secrecy of the unopened commitments under selective openings. We should also stress that it is this definition that turns out to be useful in the context of interactive proof system, see Appendix C.

A relaxation. Alternatively, we could let the adversary predict a predicate π of the whole message vector, and consider him successful if $\Pr[b = \pi(M)]$ and $\Pr[b = \pi(M')]$ for the alternative message vector $M' \leftarrow \mathcal{M} \mid M_I$ differ non-negligibly. We stress that our upcoming negative result (as well as the application in Appendix C) also applies to this relaxed notion.

Theorem 4.2 (Second main result: impossibility of perfectly binding IND-SO-COM, most general formulation)**.** *Let $n = n(k) = 2k$, and let $\mathcal{I} = (\mathcal{I}_n)_n$ with $\mathcal{I}_n = \{I \subseteq [n] : |I| = n/2\}$ be the family of all $n/2$-sized subsets of $[n]$. Let $\mathcal{X}$ be an oracle that satisfies property $\mathcal{P}$ even in presence of a PSPACE-oracle. We demand that $\mathcal{X}$ is computable in PSPACE, at least in polynomially bounded contexts.*[7] *Then, there exists a set of oracles relative to which $\mathcal{X}$ still satisfies $\mathcal{P}$, but no perfectly binding commitment scheme is indistinguishable under selective openings.*

Proof. First, let $\varepsilon \in \mathbb{R}$ be a suitably small positive real number that does not depend on n. (We will determine δ later.) Let $\mathbb{F}$ be the finite field of size 2^k. Let $\mathcal{C}$ be the oracle that initially chooses a linear code C over $\mathbb{F}$ with length n, dimension $D \geq (1/2 + \varepsilon)$ and minimum distance $d \geq 5\varepsilon$. That is, $\mathcal{C}$ chooses a full-rank generator matrix $G \in \mathbb{F}^{D \times n}$ such that for any distinct $x, y \in \mathbb{F}^D$, the vectors xG and yG differ in at least d components. We denote with C the induced linear code, i.e., $C = \{xG \mid x \in \{0,1\}^D\} \subseteq \mathbb{F}^n$. For suitably small (but positive) ε and suitably large values of n, such codes exist due to the Gilbert-Varshamov bound (cf., e.g., MacWilliams and Sloane [16], Theorem 12 and Problem 45). We assume such values of n and ε, and we also assume n large enough such that $\varepsilon n \geq 1$. Upon any input, $\mathcal{C}$ replies with C (i.e., with G).

Moreover, let $\mathcal{PSPACE}$ be a PSPACE-oracle, and let $\mathcal{R}$ be the oracle that, upon input $M = (M_i)_{i \in [n]} \in (\mathbb{F} \cup \{\bot\})^n$ and $I \subseteq [n]$, proceeds as follows. If there exists $\tilde{M} = (\tilde{M}_i)_{i \in [n]} \in C$ and $J \supseteq I$ with $|J| \geq (1 - 2\varepsilon)n$ such that $\tilde{M}_J = M_J$, then return $\tilde{M}$. (Since C has minimum distance $d \geq (1 - 5\varepsilon)n$, there is at most one such $\tilde{M}$.) If no such $\tilde{M} \in C$ exists, return $\bot$. Intuitively, $\mathcal{R}$ tries to "error-correct" M and find a vector $\tilde{M} \in C$ which is "close" to M and satisfies $\tilde{M}_I = M_I$. Note that $\mathcal{R}$ can be perfectly emulated using $\mathcal{PSPACE}$ and a description of C alone; we only use an explicit $\mathcal{R}$ to ease presentation.

Finally, let $\mathcal{B}$ be the oracle that proceeds as follows:

1. Upon input $(\mathsf{Com}, \mathsf{Ver}, com)$, where $com = (com_i)_{i \in [n]}$, check that $(\mathsf{Com}, \mathsf{Ver})$ describes a perfectly binding, but not necessarily hiding, commitment scheme.[8] If not, reject with output $\bot$. If yes, return a uniformly chosen $I \in \mathcal{I}$ and record $(\mathsf{Com}, \mathsf{Ver}, com, I)$.
2. Upon input $(\mathsf{Com}, \mathsf{Ver}, com, dec_I)$ with $dec_I = (dec_i)_{i \in I}$ for a $(\mathsf{Com}, \mathsf{Ver}, com, I)$ which was previously recorded, verify using Ver that each dec_i is a valid opening of the respective com_i. If not, reject with output $\bot$. If yes, extract the whole message vector M from com (this is possible uniquely since $(\mathsf{Com}, \mathsf{Ver})$ is perfectly binding), and return $\mathcal{R}(M, I)$.

We should comment on $\mathcal{B}$'s check whether $(\mathsf{Com}, \mathsf{Ver})$ is perfectly binding. We want that, for all possible values of C and states of $\mathcal{X}$, and for all syntactically allowed commitments com_i, there is at most one message M_i to which com_i can be opened in the sense of Ver. Note that by assumption about $\mathcal{X}$, this condition can be checked using PSPACE-oracle $\mathcal{PSPACE}$. (For instance, if $\mathcal{X}$ is a random oracle, then we can let $\mathcal{PSPACE}$ iterate over all possible answers to actually made queries; since there can by only polynomially many such queries in our context, this can be done in PSPACE. More generally, we can iterate over suitable prefixes of $\mathcal{X}$'s random tape.) Note that we completely ignore whether or not $(\mathsf{Com}, \mathsf{Ver})$ is hiding.

Lemma 4.3. *Let $(\mathsf{Com}^*, \mathsf{Ver}^*)$ be a perfectly binding commitment scheme (that may use all of the described oracles in its algorithms). Then $(\mathsf{Com}^*, \mathsf{Ver}^*)$ is not IND-SO-COM secure.*

[7] This is not a contradiction. An example of such an $\mathcal{X}$ is a random oracle or an ideal cipher, using lazy sampling. It will become clearer how we use the PSPACE requirement in the proof.

[8] see the discussion after the description of $\mathcal{B}$

Proof. Consider the n-message distribution $\mathcal{M}^*$ that samples random elements of C. (I.e., $\mathcal{M}^*$ outputs a uniformly sampled $M \in C \subseteq \mathbb{F}^n$.) Consider the following adversary A that relays between the real or ideal IND-SO-COM experiment and oracle $\mathcal{B}$:

1. Upon receiving $com^* = (com_i^*)_{i\in[n]}$ from the experiment, send $(\mathsf{Com}^*, \mathsf{Ver}^*, com^*)$ to $\mathcal{B}$.
2. Upon receiving $I^* \in \mathcal{I}$ from $\mathcal{B}$, send I^* to the IND-SO-COM experiment.
3. Upon receiving openings $dec_{I^*}^* = (dec_i^*)_{i\in I^*}$ and a challenge message M from the experiment, send $(\mathsf{Com}^*, \mathsf{Ver}^*, com^*, dec_{I^*}^*)$ to $\mathcal{B}$.
4. Finally, upon receiving $\tilde{M} \in \mathbb{F}^n$ from $\mathcal{B}$, output $out_A = 1$ iff $M = \tilde{M}$.

(Again, A is straightforwardly split into parts A_1 and A_2.)

Now by construction of the IND-SO-COM experiment and $\mathcal{B}$, we have that the message $\tilde{M}$ that A receives from $\mathcal{B}$ will always be identical to the initially sampled message M^*, both in the real and the ideal IND-SO-COM experiment. Hence, A will always output 1 in the real IND-SO-COM experiment (since then $M = M^*$ by definition). In the ideal experiment, M will be a random codeword with $M_{I^*} = M_{I^*}^*$. However, since code C has dimension $D \geq (1/2 + \varepsilon) \geq |I^*| + 1$, there are at least $|\mathbb{F}| = 2^k$ possible such M, and so $M = M^*$ with probability at most 2^{-k}. Hence A will output 1 with negligible probability in the ideal IND-SO-COM experiment. We get that $\mathsf{Adv}^{\text{ind-so}}_{\mathsf{Com}^*,\mathcal{M}^*,A}$ is overwhelming, which proves the lemma. □

Lemma 4.4. *$\mathcal{X}$ satisfies $\mathcal{P}$.*

Proof. For contradiction, suppose that there is a successful (computationally unbounded, but polynomially bounded in the number of oracle queries) adversary A on $\mathcal{X}$'s property $\mathcal{P}$. We first argue that A can do without $\mathcal{B}$. Formally, we build a refined A' from A such that A' never queries $\mathcal{B}$, but still achieves that $\mathsf{Pr}[1 \leftarrow \mathcal{P}] - 1/2$ is non-negligible.

Now A' simulates A and answers A's $\mathcal{B}$-queries on its own, as follows:

1. Upon input $(\mathsf{Com}, \mathsf{Ver}, com)$ from A, where $com = (com_i)_{i\in[n]}$, check that $(\mathsf{Com}, \mathsf{Ver})$ describes a perfectly binding, but not necessarily hiding, commitment scheme.[9] If not, reject with output $\perp$. If yes, return a uniformly chosen $I \in \mathcal{I}$ and record $(\mathsf{Com}, \mathsf{Ver}, com, I)$.
2. Upon input $(\mathsf{Com}, \mathsf{Ver}, com, dec_I)$ with $dec_I = (dec_i)_{i\in I}$ for a $(\mathsf{Com}, \mathsf{Ver}, com, I)$ which was previously recorded, verify using Ver that each dec_i is a valid opening of the respective com_i. If not, reject with output $\perp$. If yes, apply procedure `Rewind` described below. If `Rewind` fails, return $\perp$. If `Rewind` succeeds, it will return a set $R \subseteq [n]$ of size $|R| \geq (1-\varepsilon)n$ along with messages M'_R such that $M'_R = M_R$ for the unique messages M inside com that would be extracted by $\mathcal{B}$. In this case, return $\mathcal{R}(M', I)$, where we set $M'_i := \perp$ for $i \notin R$.

We denote this simulation of $\mathcal{B}$ by $\mathcal{B}'$. Before we analyze $\mathcal{B}'$ further, we sketch the `Rewind` procedure. `Rewind` rewinds A back to the state just before receiving $I \in \mathcal{I}$ from $\mathcal{B}'$, and replaces I with a freshly sampled $I' \in \mathcal{I}$, in the hope that A opens $M_{I'}$ later. We argue in Appendix E that rewinding a sufficient (polynomial[10]) number of times will with high probability allow to extract $(M_i)_{i\in R}$ for $R \supseteq I$ with $|R| \geq (1-\varepsilon)n$. In particular, we will prove that the probability for `Rewind` to fail in *any* $\mathcal{B}'$-query is significantly smaller than A's success probability. For that reason, we will henceforth silently assume that `Rewind` did not fail. A detailed description and analysis of `Rewind` appears in Appendix E.

First we remark that, given a fixed vector com (that also fixes M), there is at most one $M^{\mathcal{B}} \in C$ that $\mathcal{B}$ could possibly output, and this $M^{\mathcal{B}}$ does not depend on I. Indeed, whenever $\mathcal{B}$ outputs some $M^{\mathcal{B}} \in C$, then by definition of $\mathcal{R}$, there must be a $J \subset [n]$, $|J| \geq (1-2\varepsilon)n$, such that $M_J^{\mathcal{B}} = M_J$. For any two possible $\mathcal{B}$-outputs $M^{\mathcal{B},1}, M^{\mathcal{B},2} \in C$ and corresponding subsets J_1, J_2, we have $M^{\mathcal{B},1}_{J_1\cap J_2} = M_{J_1\cap J_2} = M^{\mathcal{B},1}_{J_1\cap J_2}$. Hence $M^{\mathcal{B},1}$ and $M^{\mathcal{B},2}$ match in at least $|J_1 \cap J_2| \geq (1-4\varepsilon)n \geq (1-5\varepsilon)n + 1$ components, which implies $M^{\mathcal{B},1} = M^{\mathcal{B},2}$ by definition of C. The same argument shows that, given com, also $\mathcal{B}'$ can only output either $\perp$ or $M^{\mathcal{B}'} = M^{\mathcal{B}}$.

[9] By assumption, this can be efficiently done by A' using $\mathcal{PSPACE}$.

[10] Although A' will not necessarily be polynomial-time, we need to keep the number of A''s oracle queries polynomial, hence the need to bound the number of rewindings.

We claim that $\mathcal{B}'$ will output what $\mathcal{B}$ would have output, except with negligible probility. Indeed, if $\mathcal{B}'$ outputs $M^{\mathcal{B}'} \in C$, then already M_R can be error-corrected (in the sense of $\mathcal{R}$) to a unique vector $\tilde{M} = M^{\mathcal{B}'} \in C$. Hence also M can be error-corrected to the same $\tilde{M} \in C$, and so $\mathcal{B}$ would have output $M^{\mathcal{B}} = M^{\mathcal{B}'}$.

Conversely, assume that $\mathcal{B}$ would have output $M^{\mathcal{B}} \in C$ (and not $\perp$). For contradiction, assume that $\mathcal{B}'$ outputs $\perp$. Since $\mathcal{B}$ would have output $M^{\mathcal{B}}$, there exists a subset $J \supseteq I$ of size $|J| \geq (1-2\varepsilon)n$ with $M_J = M^{\mathcal{B}}_J$. Denote by $R \supseteq I$ the indices for which `Rewind` extracts M_R. Call an index $i \in [n]$ *bad* iff $M_i \neq M^{\mathcal{B}}_i$. Because we assumed that $\mathcal{B}'$ outputs $\perp$, every $(1-2\varepsilon)n$-sized subset of M_R must contain at least one bad index. Since $|R| \geq (1-\varepsilon)n$, we get that R contains at least εn bad indices. Now n grows linearly in k, and so a uniformly chosen subset $I \subseteq [n]$ contains hence a bad index with overwhelming probability over I. For any such choice of I, $\mathcal{B}$ would have output $\perp$ and not $M^{\mathcal{B}}$ since $M^{\mathcal{B}}_I = M_I$ by definition of $\mathcal{R}$. So $\mathcal{B}$'s probability to output $M^{\mathcal{B}}$ must be negligible in the first place. This shows that the claim "whenever $\mathcal{B}$ would have output $M^{\mathcal{B}} \in C$, then $\mathcal{B}'$ outputs $M^{\mathcal{B}'} = M^{\mathcal{B}}$" holds with overwhelming probability.

We conclude that the internal simulation $\mathcal{B}'$ behaves like $\mathcal{B}$, except with sufficiently small probability. Hence A' breaks property $\mathcal{P}$, but without querying $\mathcal{B}$. Without loss of generality, we can also assume that A' never queries $\mathcal{R}$, since $\mathcal{R}$-queries can be efficiently emulated using $\mathcal{PSPACE}$ (that itself cannot access $\mathcal{X}$) and a description of code C alone. Hence A' breaks property $\mathcal{P}$ with only a polynomial number of queries to $\mathcal{X}$ and $\mathcal{PSPACE}$. This contradicts our assumption on $\mathcal{X}$ and creates the desired contradiction. □

Taking the two lemmas together proves 4.2. □

Similarly to 3.7, we get for concrete choices of $\mathcal{X}$ and $\mathcal{P}$:

Corollary 4.5 (Second main result: impossibility of perfectly binding IND-SO-COM)**.** *Let n and $\mathcal{I}$ as in 4.2. Then no perfectly binding commitment scheme can be proven simulatable under selective openings via a $\forall\exists$semi-blackbox reduction from one or more of the following primitives: one-way functions, one-way permutations, trapdoor one-way permutations, an IND-CCA secure encryption scheme.*

Fortunately, things look different for statistically hiding commitment schemes (since the proof of the following theorem is rather straightforward, we postpone it to Appendix E):

Theorem 4.6 (Possibility of statistically hiding IND-SO-COM)**.** *Let* $(\mathsf{Com}, \mathsf{Ver})$ *be a statistically hiding commitment scheme. Then* $(\mathsf{Com}, \mathsf{Ver})$ *is indistinguishable under selective openings.*

Now statistically hiding (and hence IND-SO-COM secure) commitment schemes can be constructed using a blackbox reduction from one-way functions (Haitner and Reingold [14]), but 3.7 implies that this is not possible for SIM-SO-COM security. This immediately implies that IND-SO-COM security does not imply SIM-SO-COM security via a blackbox reduction.

Generalizations. Again, 4.5 constitutes merely an example instantiation of the much more general 4.2. Also, when considering interactive commitment schemes, the proof 4.2 generalizes with the same changes as for the proof of 3.3. 4.6 generalizes as well: here, it is only necessary to note that any statistically hiding commitment can be opened (also interactively) as a commitment to any other message. We stress that the proof of 4.6 also holds (literally) in case A and/or $\mathcal{M}$ gets an additional auxiliary input $z_k \in \{0,1\}^*$. However, we stress that the proof for 4.2 does *not* apply to "almost-perfectly binding" commitment schemes such as the one by Naor [17]. (For such schemes, oracle $\mathcal{B}$'s check that the supplied commitment scheme is binding might tell something about $\mathcal{X}$.)

Acknowledgements. I am indebted to Enav Weinreb, Marc Stevens, and Serge Fehr for many insightful discussions. In particular, Enav suggested interactive proof systems as an application, and Marc contributed to the first negative result.

References

[1] Donald Beaver. Plug and play encryption. In Joan Feigenbaum, editor, *Advances in Cryptology, Proceedings of CRYPTO '91*, number 576 in Lecture Notes in Computer Science, pages 75–89. Springer-Verlag, 1992.

[2] Mihir Bellare and Phillip Rogaway. Optimal asymmetric encryption—how to encrypt with RSA. In Alfredo de Santis, editor, *Advances in Cryptology, Proceedings of EUROCRYPT '94*, number 950 in Lecture Notes in Computer Science, pages 92–111. Springer-Verlag, 1995.

[3] Manuel Blum. Coin flipping by telephone. In Allen Gersho, editor, *Advances in Cryptology, A report on CRYPTO 81*, number 82-04 in ECE Report, pages 11–15. University of California, Electrical and Computer Engineering, 1982.

[4] Ran Canetti, Uri Feige, Oded Goldreich, and Moni Naor. Adaptively secure multi-party computation. In *Twenty-Eighth Annual ACM Symposium on Theory of Computing, Proceedings of STOC 1995*, pages 639–648. ACM Press, 1996.

[5] Ran Canetti, Cynthia Dwork, Moni Naor, and Rafail Ostrovsky. Deniable encryption. In Burton S. Kaliski Jr., editor, *Advances in Cryptology, Proceedings of CRYPTO '97*, number 1294 in Lecture Notes in Computer Science, pages 90–104. Springer-Verlag, 1997.

[6] Ran Canetti, Shai Halevi, and Jonathan Katz. Adaptively-secure, non-interactive public-key encryption. In Joe Kilian, editor, *Theory of Cryptography, Proceedings of TCC 2005*, number 3378 in Lecture Notes in Computer Science, pages 150–168. Springer-Verlag, 2005.

[7] Ivan Damgård. A "proof-reading" of some issues in cryptography. In Lars Arge, Christian Cachin, Tomasz Jurdzinski, and Andrzej Tarlecki, editors, *Automata, Languages and Programming, 34th International Colloquium, Proceedings of ICALP 2007*, number 4596 in Lecture Notes in Computer Science, pages 2–11. Springer-Verlag, 2007.

[8] Ivan Damgård and Jesper Buus Nielsen. Improved non-committing encryption schemes based on general complexity assumptions. In Mihir Bellare, editor, *Advances in Cryptology, Proceedings of CRYPTO 2000*, number 1880 in Lecture Notes in Computer Science, pages 432–450. Springer-Verlag, 2000.

[9] Yevgeniy Dodis, Roberto Oliveira, and Krzysztof Pietrzak. On the generic insecurity of the full domain hash. In Victor Shoup, editor, *Advances in Cryptology, Proceedings of CRYPTO 2005*, number 3621 in Lecture Notes in Computer Science, pages 449–466. Springer-Verlag, 2005.

[10] Cynthia Dwork, Moni Naor, Omer Reingold, and Larry Stockmeyer. Magic functions. *Journal of the ACM*, 50(6):852–921, 2003.

[11] Oded Goldreich. *Foundations of Cryptography – Volume 1 (Basic Tools).* Cambridge University Press, August 2001.

[12] Oded Goldreich and Hugo Krawczyk. On the composition of zero-knowledge proof systems. In Mike Paterson, editor, *Automata, Languages and Programming, 17th International Colloquium, Proceedings of ICALP 90*, number 443 in Lecture Notes in Computer Science, pages 268–282. Springer-Verlag, 1990.

[13] Oded Goldreich, Silvio Micali, and Avi Wigderson. Proofs that yield nothing but their validity or all languages in NP have zero-knowledge proof systems. *Journal of the ACM*, 38(1):691–729, 1991.

[14] Iftach Haitner and Omer Reingold. Statistically-hiding commitment from any one-way function. In *39th Annual ACM Symposium on Theory of Computing, Proceedings of STOC 2007*, pages 1–10. ACM Press, 2007.

[15] Russell Impagliazzo and Steven Rudich. Limits on the provable consequences of one-way permutations. In *Twenty-First Annual ACM Symposium on Theory of Computing, Proceedings of STOC 1989*, pages 44–61. ACM Press, 1989. Extended abstract.

[16] F. Jessie MacWilliams and Neil J.A. Sloane. *The Theory of Error-Correcting Codes*. Elsevier Science Publishers, 1988.

[17] Moni Naor. Bit commitment using pseudo-randomness. *Journal of Cryptology*, 4(2):151–158, 1991.

[18] Jesper B. Nielsen. Separating random oracle proofs from complexity theoretic proofs: The non-committing encryption case. In Moti Yung, editor, *Advances in Cryptology, Proceedings of CRYPTO 2002*, number 2442 in Lecture Notes in Computer Science, pages 111–126. Springer-Verlag, 2002.

[19] Omer Reingold, Luca Trevisan, and Salil P. Vadhan. Notions of reducibility between cryptographic primitives. In Moni Naor, editor, *Theory of Cryptography, Proceedings of TCC 2004*, number 2951 in Lecture Notes in Computer Science, pages 1–20. Springer-Verlag, 2004.

[20] Daniel R. Simon. Finding collisions on a one-way street: Can secure hash functions be based on general assumptions? In Kaisa Nyberg, editor, *Advances in Cryptology, Proceedings of EUROCRYPT '98*, number 1403 in Lecture Notes in Computer Science, pages 334–345. Springer-Verlag, 1998.

A Commitment schemes

For clarity, we give a more detailed definition of commitment schemes.

Definition A.1 (Commitment scheme). *A (noninteractive)* commitment scheme $(\mathsf{Com}, \mathsf{Ver})$ *is a pair of PPT algorithms, such that the following holds:*

Syntax. *For any $M \in \{0,1\}^k$, algorithm $\mathsf{Com}(M)$ outputs a pair (com, dec), and $\mathsf{Ver}(com, dec)$ deterministically outputs either a message $M \in \{0,1\}^k$ or rejects with output $\perp$.*

Correctness. *For all $M \in \{0,1\}^k$ and $(com, dec) \leftarrow \mathsf{Com}(M)$, we have $\mathsf{Ver}(com, dec) = M$.*

Binding. *For an algorithm A, let $\mathsf{Adv}^{\mathsf{binding}}_{\mathsf{Com},\mathsf{Ver},A}$ be the probability that A outputs (com, dec_1, dec_2) with*

$$\mathsf{Ver}(com, dec_1) = M_1 \neq M_2 = \mathsf{Ver}(com, dec_2).$$

We demand that for any PPT A, $\mathsf{Adv}^{\mathsf{binding}}_{\mathsf{Com},\mathsf{Ver},A}$ is negligible in the security parameter.

Hiding. *For a pair $A = (A_1, A_2)$ of algorithms, let*

$$\mathsf{Adv}^{\mathsf{hiding}}_{\mathsf{Com},A} := \Pr\left[\mathsf{Exp}^{\mathsf{hiding}\text{-}0}_{\mathsf{Com},A} = 1\right] - \Pr\left[\mathsf{Exp}^{\mathsf{hiding}\text{-}1}_{\mathsf{Com},A} = 1\right].$$

Here, $\mathsf{Exp}^{\mathsf{hiding}\text{-}b}_{\mathsf{Com},A}$ proceeds as follows:

1. *run $(s, M_0, M_1) \leftarrow A_1(1^k)$ to obtain two messages $M_0, M_1 \in \{0,1\}^k$ and a state s,*
2. *compute $(com, dec) \leftarrow \mathsf{Com}(M_b)$,*
3. *run $b' \leftarrow A_2(s, com)$ to obtain a guess bit b'*
4. *output b'.*

We demand that $\mathsf{Adv}^{\mathsf{hiding}}_{\mathsf{Com},A}$ is negligible for any PPT A.

Furthermore, if $\mathsf{Adv}^{\mathsf{hiding}}_{\mathsf{Com},A}$ is negligible for all (not necessarily PPT) A, then $(\mathsf{Com}, \mathsf{Ver})$ is statistically hiding. *If $\mathsf{Adv}^{\mathsf{binding}}_{\mathsf{Com},\mathsf{Ver},A} = 0$ for all A, then $(\mathsf{Com}, \mathsf{Ver})$ is* perfectly binding.

B Application to adaptively secure encryption

Motivation and setting. Taking up the motivation of Damgård [7], we consider the setting of an adversary A that may corrupt, in an adaptive manner, a subset of a set of parties $P_1, \dots, P_n$. Assume that for all i, the public encryption key pk_i with which party P_i encrypts outgoing messages, is publicly known. Suppose further that A may corrupt parties based on all public keys and all so far received ciphertexts. When A corrupts P_i, A learns P_i's internal state and history, in particular A learns the randomness used for all of that party's encryptions, and its secret key sk_i. We assume the following:

1. The number of parties is $n = 2k$ for the security parameter k,
2. It is allowed for A to choose at some point a subset $I \subseteq [n]$ of size $n/2$ and to corrupt all these P_i $(i \in I)$.
3. We can interpret the used encryption scheme as a (hiding and binding) commitment scheme $(\mathsf{Com}, \mathsf{Ver})$ in the following sense: $\mathsf{Com}(M)$ generates a fresh public key pk and outputs a commitment $com = (pk, \mathsf{Enc}(pk, M; r))$ and a decommitment $dec = (M, r)$. Here Enc denotes the encryption algorithm of the encryption scheme, and r denotes the randomness used while encrypting M. Verification of $(com, dec) = (pk, C, M, r)$ checks that $\mathsf{Enc}(pk, M; r) = C$.

Note that the third assumption does not follow from the scheme's correctness. Indeed, correctness implies that honestly generated (pk, M) are committing. However, there are schemes for which it is easy to come up with fake public keys and ciphertexts (i.e., fake commitments) which are computationally indistinguishable from honestly generated commitments, but can be opened in arbitrary ways. Prominent examples of such schemes are non-committing encryption schemes [4, 1, 5, 8, 6], which however generally contain an interactive set-up phase and are comparatively inefficient.

Application of our impossibility results. Attacks in this setting cannot be easily simulated in the sense of, e.g., Canetti et al. [4]: such a simulator would in particular be able to simulate openings (in the sense of Ver, i.e., openings of ciphertexts). Hence, this would imply a simulator for $(\mathsf{Com}, \mathsf{Ver})$ in the sense of 3.1. Now from 3.7 we know that the construction and security analysis of such a simulator requires either a very strong computational assumption, or fundamentally non-blackbox techniques. Even worse: if $(\mathsf{Com}, \mathsf{Ver})$ is perfectly binding[11], then 4.5 shows that not even secrecy in the sense of 4.1[12] can be proven (blackbox) secure.

We stress that these negative results only apply if encryption really constitutes a (binding) commitment scheme in the above sense. In fact, e.g., [4] construct a sophisticated *non-committing* (i.e., non-binding) encryption scheme and prove simulatability for their scheme. Our results show that such a non-committing property is necessary.

C Application to zero-knowledge proof systems

Application of our impossibility results. Dwork et al. [10] considered the applications of SIM-SO-COM secure commitment schemes to zero-knowledge protocols. In their Theorem 7.6, they show that SIM-SO-COM secure commitments essentially make the well-known graph 3-coloring protocol of [13] composable *in parallel*. This would be a very exciting result, given the negative results for the concurrent composition of zero-knowledge protocols of Goldreich and Krawczyk [12]. Unfortunately, 3.7 shatters any hopes to construct such a scheme from standard assumptions and using standard proof techniques.

Application of our positive result. A natural question is whether IND-SO-COM security, our achievable relaxation of SIM-SO-COM security, provides a reasonable fallback in this situation. Now first, our results show that even when using IND-SO-COM secure schemes, we cannot rely on perfectly binding commitment schemes because of 4.2. For many interesting interactive proofs (and in particular the graph 3-coloring proof

[11] in the presence of non-uniform adversaries, this is already implied by the fact that the scheme is noninteractive and computationally binding

[12] in the context of encryption, 4.1 would translate to a variant of indistinguishability of ciphertexts

from [13]), this unfortunately means that the proof system degrades to an argument system. But, assuming we are willing to pay this price, what do we get from IND-SO-COM security?

The answer is "essentially witness indistinguishability," as we will argue in a minute. Intuitively, any commitment scheme which satisfies (a slight variation of) IND-SO-COM security can be used to implement "commit-choose-open" style interactive argument systems, such that

- the argument system is witness indistinguishable,
- the security reduction is tight (and in particular does not lose a factor of $|I|$, where $|I|$ is the number of possible choices in the second stage), and
- we get composability essentially "for free."

(More details follow.) Now witness indistinguishable argument systems already enjoy a composition theorem (see, e.g., Goldreich [11], Lemma 4.6.6), so at least the second of these claims is not surprising. However, our point here is that the security notion of IND-SO-COM secure commitments itself is a "good" and useful notion.

Formal setting. Consider an interactive argument system (P, V) for an NP-language $\mathcal{L}$ with witness relation R. (We refer to Goldreich [11], Chapter 4 for an introduction to zero-knowledge interactive proof/argument systems.) We assume that (P, V) is of the following "commit-choose-open" form, where the prover P gets as input a statement $x \in \mathcal{L}$ along with a witness w such that $R(x, w)$, and the verifier only gets x.

1. P generates n commitments $com = (com_i)_{i \in [n']}$ and sends them to V,
2. V chooses a subset $I \subseteq [n']$,
3. P opens the commitments com_i for $i \in I$ by sending $dec_I = (dec_i)_{i \in I}$ to V,
4. V accepts if the openings are valid and if the opened values satisfy some fixed relation specified by the protocol.

We suppose further that the value of the actually opened messages M_I is always statistically independent of the used witness w. These are strong assumptions, but at least one of the most important zero-knowledge interactive proof systems (namely, the graph 3-coloring protocol from Goldreich et al. [13]) is of this form.

Connection to IND-SO-COM security. To state our claim, let us recall the definition of witness indistinguishability from Goldreich [11] (we chose a slightly different but equivalent formulation):

Definition C.1 (Witness indistinguishability)**.** *Let (P, V) be an interactive argument system for language $\mathcal{L}$ with witness relation R. (P, V) is* witness indistinguishable for R *iff for every PPT machines V^* and D, all sequences $x = (x_k)_{k \in \mathbb{N}}$, $w^0 = (w^0_k)_{k \in \mathbb{N}}$, and $w^1 = (w^1_k)_{k \in \mathbb{N}}$ with $|x_k| = k$ and $R(x_k, w^0_k)$ and $R(x_k, w^1_k)$, and all auxiliary inputs $z = (z_k)_{k \in \mathbb{N}} \in (\{0,1\}^*)^{\mathbb{N}}$, we have that*

$$\begin{aligned}\mathsf{Adv}^{\mathsf{WI}}_{x,w^0,w^1,V^*,D,z} := \mathsf{Pr}\left[D(x_k, z_k, \langle P(x_k, w^0_k), V^*(x_k, z_k)\rangle) = 1\right] \\ - \mathsf{Pr}\left[D(x_k, z_k, \langle P(x_k, w^1_k), V^*(x_k, z_k)\rangle) = 1\right]\end{aligned}$$

is negligible in k. Here, $\langle P(x, w), V^*(x)\rangle$ denotes a transcript of the interaction between P and V^*.

Since this definition involves auxiliary input z given to the verifier/adversary V^*, we also consider a variation of 4.1 that involves auxiliary input. Namely,

Definition C.2 (Auxiliary-input-IND-SO-COM)**.** *In the situation of 4.1, we call* $(\mathsf{Com}, \mathsf{Ver})$ auxiliary-input-IND-SO-COM *iff* $\mathsf{Adv}^{\mathsf{ind\text{-}so}}_{\mathsf{Com},\mathcal{M},A,z}$ *is negligible for all PPT $\mathcal{M}$ and A and all auxiliary inputs $z = (z_k)_{k \in \mathbb{N}} \in (\{0,1\}^*)^{\mathbb{N}}$, where both $\mathcal{M}$ and A are invoked with additional auxiliary input z_k.*

Now we are ready to prove the following connection between witness indistinguishability and auxiliary-input-IND-SO-COM:

Theorem C.3 (Auxiliary-input-IND-SO-COM implies witness indistinguishability)**.** *Assume an interactive argument system (P, V) as above. Then, if the commitment scheme in (P, V) is auxiliary-input-IND-SO-COM for parameters $n = n' + 1$ and all subsets $\mathcal{I}$ of $[n']$ as possible in (P, V), then (P, V) is witness indistinguishable. The security reduction loses only a factor of 2.*

Proof. For contradiction, assume x, w^0, w^1, V^*, D, z such that $\mathsf{Adv}^{\mathsf{WI}}_{x,w^0,w^1,V^*,D,z}$ is non-negligible. We construct a message distribution $\mathcal{M}$, an adversary A, and a z' such that

$$\mathsf{Adv}^{\mathsf{ind\text{-}so}}_{\mathsf{Com},\mathcal{M},A,z} = \frac{1}{2}\mathsf{Adv}^{\mathsf{WI}}_{x,w^0,w^1,V^*,D,z}. \tag{1}$$

First, define $z'_k = (x_k, w^0_k, w^1_k, z_k)$, so that $\mathcal{M}$ and A are both invoked with *both* witnesses and z_k. Then, let $\mathcal{M}$ be the following PPT algorithm:

1. upon input $z'_k = (x_k, w^0_k, w^1_k, z_k)$, toss a coin $b \in \{0, 1\}$,
2. sample messages $(M_i)_{i\in[n']}$ by running P on input (x_k, w^b_k),
3. define $M_{n'+1} := b$,
4. return the $(n'+1)$-message vector $(M_i)_{i\in[n'+1]}$.

Now adversary A proceeds as follows:

1. upon input $z'_k = (x_k, w^0_k, w^1_k, z_k)$ and commitments $com = (com_i)_{i\in[n'+1]}$, run V^* on input (x_k, z_k) and commitments $(com_i)_{i\in[n']}$,
2. when V^* chooses a set $I \subseteq [n']$, relay this set (interpreted as a subset of $[n] = [n'+1]$) to the IND-SO-COM experiment,
3. upon receiving openings $(dec_i)_{i\in I}$ and a message vector $M^* = (M^*_i)_{i\in[n]}$ from the experiment, run D on input $(x_k, z_k, (com, I, dec_I))$ to receive a guess b' from D,
4. output $b' \oplus M^*_{n'+1}$.

(As usual, A is straightforwardly split up into (A_1, A_2) as required by the IND-SO-COM experiment.)

Now in the real IND-SO-COM experiment $\mathsf{Exp}^{\mathsf{ind\text{-}so\text{-}real}}_{\mathsf{Com},\mathcal{M},A,z}$, the following happens: if $\mathcal{M}$ chose $b = 0$, then an interaction of $P(x_k, w^0_k)$ and $V^*(x_k, z_k)$ is perfectly simulated, so that A (and hence, since $M^*_{n'+1} = b = 0$, also $\mathsf{Exp}^{\mathsf{ind\text{-}so\text{-}real}}_{\mathsf{Com},\mathcal{M},A,z}$) outputs $D(x_k, z_k, \langle P(x_k, w^0_k), V^*(x_k, z_k)\rangle)$. Conversely, if $b = 1$, then $\mathsf{Exp}^{\mathsf{ind\text{-}so\text{-}real}}_{\mathsf{Com},\mathcal{M},A,z}$ outputs $1 - D(x_k, z_k, \langle P(x_k, w^1_k), V^*(x_k, z_k)\rangle)$ because $M^*_{n'+1} = b = 1$ then. We get that

$$\begin{aligned}\Pr\left[\mathsf{Exp}^{\mathsf{ind\text{-}so\text{-}real}}_{\mathsf{Com},\mathcal{M},A,z} = 1\right] = \frac{1}{2}\Big(&\Pr\left[D(x_k, z_k, \langle P(x_k, w^0_k), V^*(x_k, z_k)\rangle) = 1\right] \\ &+ 1 - \Pr\left[D(x_k, z_k, \langle P(x_k, w^0_k), V^*(x_k, z_k)\rangle) = 1\right]\Big) = \frac{1}{2}\mathsf{Adv}^{\mathsf{WI}}_{x,w^0,w^1,V^*,D,z} + \frac{1}{2}.\end{aligned}$$

On the other hand, in the ideal IND-SO-COM experiment, the message $M^*_{n'+1}$ that A receives from the experiment results from a resampling of $\mathcal{M}$, conditioned on $M^*_I = M_I$. Since we assumed about (P, V) that M_I is independent of the used witness, M_I is also independent of b, and hence $M^*_{n'+1}$ will be a freshly tossed coin. We get

$$\Pr\left[\mathsf{Exp}^{\mathsf{ind\text{-}so\text{-}ideal}}_{\mathsf{Com},\mathcal{M},A,z} = 1\right] = \frac{1}{2}.$$

Putting things together proves Equation 1. □

Tightness in the reduction and composition. We stress that we only lose a factor of 2 in our security reduction, which contrasts the loss of a factor of about n'^2 in the proof of Goldreich et al. [13]. Admittedly, their proof works also for perfectly binding commitment schemes (thus achieving an interactive *proof* system), which we (almost) cannot hope to satisfy IND-SO-COM security, according to 4.2. However, since we can give IND-SO-COM secure schemes for arbitrary parameters, n and $\mathcal{I}$, we can hope to apply C.3 even to protocols where $|\mathcal{I}_n|$ is superpolynomial.[13] In particular, our proof shows that we can even map several parallel executions of a protocol (P, V) to the IND-SO-COM security experiment. This derives a parallel composition theorem (for this particular class of protocols and witness indistinguishability) at virtually no extra cost.

[13]Of course, it is possibly to directly prove, say, witness indistinguishability for the case of superpolynomial $|\mathcal{I}_n|$ from statistically hiding commitment schemes. However, our point here is to illustrate the usefulness of our definition.

D On the role of property $\mathcal{P}$

The intuitive contradiction. The formulations of 3.3 and 4.2 seem intuitively much too general: essentially they claim impossibility of blackbox proofs from *any* computational assumption which is formulated as a property $\mathcal{P}$ of an oracle $\mathcal{X}$. Why can't we choose $\mathcal{X}$ to be an ideally secure commitment scheme, and $\mathcal{P}$ a property that models precisely what we want to achieve, e.g., 4.1 (i.e., IND-SO-COM security)? After all, 4.1 can be rephrased as a property $\mathcal{P}$ by letting A choose a message distribution $\mathcal{M}$ and send this distribution (as a description of a PPT algorithm $\mathcal{M}$) to $\mathcal{P}$. Then, $\mathcal{P}$ could perform the $\mathsf{Exp}^{\mathsf{ind\text{-}so\text{-}real}}_{\mathsf{Com},\mathcal{M},A}$ or the $\mathsf{Exp}^{\mathsf{ind\text{-}so\text{-}ideal}}_{\mathsf{Com},\mathcal{M},A}$ experiment with A, depending on an internal coin toss (the output of $\mathcal{P}$ will then depend on A's output and on that coin toss). This $\mathcal{P}$ models 4.1, in the sense that

$$\mathsf{Adv}^{\mathsf{ind\text{-}so}}_{\mathsf{Com},\mathcal{M},A} = 2\mathsf{Adv}^{\mathcal{P}}_{A}.$$

Also, using a truly random permutation as a basis, it is natural to assume that we can construct an *ideal* (i.e., as an oracle) perfectly binding commitment scheme $\mathcal{X}$ that satisfies $\mathcal{P}$. (Note that although $\mathcal{X}$ is perfectly binding, A's view may still be almost statistically independent of the unopened messages, since the scheme $\mathcal{X}$ is given in oracle form.)

Hence, if the assumption essentially *is* already IND-SO-COM security, we can certainly achieve IND-SO-COM security (using a trivial reduction), and this seems to contradict 4.2. So where is the problem?

Resolving the situation. The problem in the above argument is that $\mathcal{P}$-security (our assumption) implies IND-SO-COM security (our goal) in a fundamentally non-blackbox way. Namely, the proof converts an IND-SO-COM adversary A and a message distribution $\mathcal{M}$ into a $\mathcal{P}$-adversary A' that sends a description of $\mathcal{M}$ to $\mathcal{P}$. This very step makes use of an *explicit representation* of the message distribution $\mathcal{M}$, and this is what makes the whole proof non-blackbox. In other words, this way of achieving IND-SO-COM security cannot be blackbox, and there is no contradiction to our results.

Viewed from a different angle, the essence of our impossibility proofs is: build a very specific message distribution, based on oracles ($\mathcal{RO}$, resp. $\mathcal{C}$), such that another "breaking oracle" $\mathcal{B}$ "breaks" this message distribution if and only if the adversary can prove that it can open commitments. This step relies on the fact that we can specify message distributions which depend on oracles. Relative to such oracles, property $\mathcal{P}$ still holds (as we prove), but may not reflect IND-SO-COM security anymore. Namely, since $\mathcal{P}$ itself cannot access additional oracles[14], $\mathcal{P}$ is also not able to sample a message space that depends on additional (i.e., on top of $\mathcal{X}$) oracles. So in our reduction, although A itself can, both in the IND-SO-COM experiment and when interacting with $\mathcal{P}$, access all oracles, it will not be able to communicate a message distribution $\mathcal{M}$ that depends on additional oracles (on top of $\mathcal{X}$) to $\mathcal{P}$. On the other hand, any PPT algorithm $\mathcal{M}$, as formalized in 4.1, *can* access all available oracles.

So for the above modeling of IND-SO-COM security as a property $\mathcal{P}$ in the sense of 3.2, our impossibility results still hold, but become meaningless (since basically using property $\mathcal{P}$ makes the proof non-blackbox). In a certain sense, this comes from the fact that the modeling of $\mathcal{P}$ is inherently non-blackbox.

What computational assumptions can be formalized as properties in a "blackbox" way? Fortunately, most standard computational assumptions can be modeled in a blackbox way as a property $\mathcal{P}$. Besides the mentioned one-way property (and its variants), in particular, e.g., the IND-CCA security game for encryption schemes can be modeled. Observe that in this game, we can let the IND-CCA adversary himself sample challenge messages M_0, M_1 for the IND-CCA experiment from his favorite distribution; no PPT algorithm has to be transported to the security game. In fact, the only properties which do not allow for blackbox proofs are those that involve an explicit transmission of code (i.e., a description of a circuit or a Turing machine). In that sense, the formulation of 3.3 and 4.2 is very general and useful.

(Non-)programmable random oracles. We stress that the blackbox requirement for random oracles (when used in the role of $\mathcal{X}$) corresponds to "non-programmable random oracles" (as used by, e.g., Bellare

[14] simply because the definition demands that $\mathcal{P}$ must be specified independently of additional oracles, cf. 3.2; if we did allow $\mathcal{P}$ to access additional oracles, this would break our impossibility proofs

and Rogaway [2]) as opposed to "programmable random oracles" (as used by, e.g., Nielsen [18]). Roughly, a proof in the programmable random oracle model translates an attack on a cryptographic scheme into an attack on a *simulated* random oracle (that is, an oracle completely under control of simulator). Naturally, such a reduction is not blackbox. And indeed, with programmable random oracles, SIM-SO-COM secure commitment schemes can be built relatively painless. As an example, [18] proves a simple encryption scheme (which can be interpreted as a commitment scheme) secure under selective openings.

E Postponed proofs

Lemma E.1. *In the situation of the proof of 3.3, event* bad *occurs only with negligible probability.*

Proof. We show that any of the events bad_{coll}, bad_{img}, bad_{bind} occurs only with negligible probability for any fixed i, j. The full claim then can be derived by a union bound over i, j, and the individual events. So first fix i, j, and note that the functions $\mathcal{RO}(\mathsf{Com}^j, \mathsf{Ver}^j, i, \cdot)$ and $\mathcal{RO}(\mathsf{Com}', \mathsf{Ver}', i', \cdot)$ are independent as soon as $\mathsf{Com}^j \neq \mathsf{Com}'$ or $\mathsf{Ver}^j \neq \mathsf{Ver}'$ or $i \neq i'$. Hence, for all of the events, we can ignore $\mathcal{RO}$- and $\mathcal{B}$-queries with different Com, Ver, or i, and assume that $\mathcal{RO}'(\cdot) := \mathcal{RO}(\mathsf{Com}^j, \mathsf{Ver}^j, i, \cdot)$ is a fresh random oracle.

bad_{coll}: Using a birthday bound, we get

$$\mathsf{Pr}\left[\exists s_1, s_2 \in \{0,1\}^{k/3}, s_1 \neq s_2 : \ \mathcal{RO}'(s_1) = \mathcal{RO}'(s_2)\right] \leq \frac{(2^{k/3})^2}{2^k} = 2^{-k/3},$$

which implies that with large probability, there simply exists no M_i^j which could raise bad_{coll}.

bad_{img}: We show that S's chance to output M with $M = \mathcal{RO}'(s)$ for some $s \in \{0,1\}^{k/3}$, and such that s has not been queried to $\mathcal{RO}'$-query, is negligible. Now S's access to the $\mathcal{B}$-oracle can be emulated using an oracle $\mathcal{B}'$ that, upon input M, outputs the set of all $s \in \{0,1\}^{k/3}$ with $\mathcal{RO}'(s) = M$. Without loss of generality, we may further assume that S never queries $\mathcal{B}'$ with an M which has been obtained through an explicit $\mathcal{RO}'(s)$-query. (Namely, unless bad_{coll} occurs, which happens only with negligible probability, $\mathcal{B}'$'s answer will then be $\{s\}$.) Hence, whenever S receives an answer $\neq \emptyset$ from $\mathcal{B}'$, it has already succeeded in producing an M with $\mathcal{RO}'(s) = M$ for some s, and without querying $\mathcal{RO}'(s)$. So without loss of generality, we can assume that S never queries $\mathcal{B}'$, and hence only produces such an M using access to $\mathcal{RO}$ and $\mathcal{RP}$ alone. Clearly, $\mathcal{RP}$ does not help S, since $\mathcal{RP}$ and $\mathcal{RO}$ are independent. But since the set of all M for which $\mathcal{RO}'(s) = M$ for some $s \in \{0,1\}^{k/3}$ is sparse in the set of all $M \in \{0,1\}^k$, and S can only make a polynomial number of $\mathcal{RO}$-queries, S's success in producing such an M is negligible.

bad_{bind}: Without loss of generality, assume that S sets I^* after $\mathcal{B}$ chooses I^j. (Otherwise, $I^j = I^*$ occurs only with probability $1/|\mathcal{I}|$, since I^j is chosen uniformly and then independent of I^*.) We can also assume that $(\mathsf{Com}^j, \mathsf{Ver}^j) = (\mathsf{Com}^*, \mathsf{Ver}^*)$, since otherwise bad_{bind} cannot happen by definition. This means that S first commits to $\mathcal{B}$ via sending $(\mathsf{Com}^j, \mathsf{Ver}^j, com)$, then receives I^j, and then sends $I^* = I^j$ to its own experiment to receive $M^*_{I^j}$. Finally, to achieve bad_{bind}, S must open com_{I^j} to $M^j_{I^j}$. In particular, there is an i such that S opens com_i to a value M_i^* which S only sees after defining com_i. This directly breaks the binding property of $(\mathsf{Com}^j, \mathsf{Ver}^j) = (\mathsf{Com}^*, \mathsf{Ver}^*)$.

□

Lemma E.2. *In the situation of the proof of 4.2, procedure* `Rewind` *extracts a subvector* M_R *of the message vector* M *with* $R \supseteq I$ *and* $|R| \geq (1-\varepsilon)n$ *upon success. The probability that* `Rewind` *fails in at least one of* A*'s* $\mathcal{B}$*-queries is at most half of the advantage of* A *against* $\mathcal{P}$.

Proof. First, we detail how `Rewind` works. Generally, since A makes only polynomially many oracle queries, and we assumed A to be successfully attacking $\mathcal{X}$'s property $\mathcal{P}$, we can assume that there is a polynomial p such that (a) for infinitely many values of the security parameter k, A's advantage against $\mathcal{P}$ is at least

$1/p(k)$, and (b) A always makes at most $p(k)$ queries. Assume concretely that A previously submitted $(\mathsf{Com}, \mathsf{Ver}, com)$ to $\mathcal{B}'$, such that $(\mathsf{Com}, \mathsf{Ver})$ is perfectly binding. Assume further that A successfully opened com_I to M_I for a subset $I \in \mathcal{I}$ uniformly chosen by $\mathcal{B}'$. Let P denote A's probability (over a uniform choice of $I \in \mathcal{I}$) to correctly open com_I.

In this situation, procedure `Rewind` records M_I and then rewinds A's state to the point where A received $I \in \mathcal{I}$ from $\mathcal{B}'$ (without altering A's random tape). `Rewind` then substitutes I with a fresh I' uniformly sampled from $\mathcal{I}$. If A later successfully opens $M_{I'}$, then `Rewind` records $M_{I'}$. (Note that there can be no contradiction among the different M_I, since any com_i can only be opened to at most one message M_i.) This process is repeated until either

- at least $(1-\varepsilon)n$ individual messages M_i have been gathered, or
- it turns out that with overwhelming probability, $P \leq 1/2p(k)^2$.

In other words, we rewind until either enough messages have been extracted, or it becomes clear that the event that A opened the first M_I successfully (which triggered `Rewind`) was very unlikely in the first place (in which case we can safely abort). We have to show that this process only takes up a polynomial number of rewindings, to show that `Rewind` is efficient.

Now, a Chernoff bound shows that using a polynomial number of rewindings, we can approximate P sufficiently well. In particular, if, say, $P \leq 1/3p(k)^2$, then we will detect that $P \leq 1/2p(k)^2$ with overwhelming probability, and we can abort. Note that `Rewind` is only triggered when A opens the first M_I. Using a union bound, we can hence conclude that the probability that in any of A's $\mathcal{B}$-queries, A opens the first M_I but `Rewind` then aborts, is at most $1/2p(k)$, i.e., at most half of A's overall success probability. That means that aborting does not significantly alter A's success.

It remains to show that, once $P > 1/3p^2(k)$, we can, using a polynomial number of rewindings, indeed extract at least $(1-\varepsilon)n$ messages M_i from A. To this end, let $\mathcal{I}' \subseteq \mathcal{I}$ be the set of I for which A opens M_I. (Note that this definition is meaningful, since we fixed A's random tape.) First, for contradiction, suppose that there is a set $B \subseteq [n]$ with $|B| \geq \varepsilon n$ such that for all $i \in B$, we have $\mathsf{Pr}\left[I \in \mathcal{I}' \wedge i \in I\}\right] < P/2n$, where the probability is over $I \in \mathcal{I}$. Then

$$
\begin{aligned}
P - \mathsf{Pr}\left[I \cap B = \emptyset\right] &= \mathsf{Pr}\left[I \in \mathcal{I}'\right] - \mathsf{Pr}\left[I \cap B = \emptyset\right] \\
&\leq \mathsf{Pr}\left[I \in \mathcal{I}' \wedge I \cap B \neq \emptyset\right] \leq \sum_{i \in B} \mathsf{Pr}\left[I \in \mathcal{I}' \wedge i \in I\right] \leq n \cdot \frac{P}{2n} = \frac{P}{2},
\end{aligned}
$$

so that $\mathsf{Pr}\left[I \cap B = \emptyset\right] \geq P/2 \geq 1/6p(k)^2$. So if $|B|$ was a constant of fraction of n as we assumed, then this probability would be negligible in $n = n(k) = 2k$. So we have a contradiction to our assumption on $|B|$, and hence there can be no such B. Thus there is an $(1-\varepsilon)n$-sized subset $R \subseteq [n]$ and all $i \in R$, we have $\mathsf{Pr}\left[I \in \mathcal{I}' \wedge i \in I\right] \geq P/2n \geq 1/12kp(k)^2$. Again using a Chernoff bound shows that a sufficient (polynomial) number of rewinding retrieves M_i for all $i \in R$, except with negligible probability. Since we started with collecting all M_I, we have $R \supseteq I$. □

Theorem E.3 (Possibility of IND-SO-COM, restated)**.** *Let* $(\mathsf{Com}, \mathsf{Ver})$ *be a statistically hiding commitment scheme. Then* $(\mathsf{Com}, \mathsf{Ver})$ *is indistinguishable under selective openings.*

Proof. Fix an n-message distribution $\mathcal{M}$ and a PPT adversary A on the SIM-SO-COM security of $(\mathsf{Com}, \mathsf{Ver})$. We start by considering the $\mathsf{Exp}^{\mathsf{ind\text{-}so\text{-}real}}_{\mathsf{Com},\mathcal{M},A}$ experiment. We refine this experiment stepwise, in each step preserving A's output distribution.

Our first modification of $\mathsf{Exp}^{\mathsf{ind\text{-}so\text{-}real}}_{\mathsf{Com},\mathcal{M},A}$ is experiment H_0, which proceeds as follows (*emphasized* steps are different from $\mathsf{Exp}^{\mathsf{ind\text{-}so\text{-}real}}_{\mathsf{Com},\mathcal{M},A}$):

1. sample messages $M = (M_i)_{i \in [n]} \leftarrow \mathcal{M}$,
2. compute (de-)commitments $(com_i, dec_i) \leftarrow \mathsf{Com}(M_i)$ for $i \in [n]$,
3. run $(s, I) \leftarrow A_1(1^k, (com_i)_{i \in [n]})$ to get state information s and a set $I \in \mathcal{I}$,

4. for every $i \in I$, compute an alternative decommitment $dec'_i \leftarrow \texttt{AltDec}(com_i, M_i)$ (procedure `AltDec` *is described below),*

5. run $b \leftarrow A_2(s, (dec'_i)_{i \in I}, M)$ to obtain a guess bit b,

6. output b.

To describe the (in general inefficient) procedure `AltDec`, consider $\mathsf{Com}(M)$'s output distribution $C_M = (C_{M,1}, C_{M,2})$. Now $\texttt{AltDec}(com_i, M_i)$ samples from C_{M_i}, conditioned on the event that $C_{M_i,1} = com_i$. `AltDec` returns the sampled $C_{M_i,2}$. In other words, `AltDec` looks, given M_i, com_i, for a corresponding decommitment dec_i, as could have been output by $\mathsf{Com}(M_i)$. If no such dec_i exists (i.e., if the probability that $\mathsf{Com}(M_i)$ returns com_i is 0), then `AltDec` returns $\perp$.

Note that the distributions of dec_i and dec'_i are identical (even given com_i, M_i), and hence

$$\mathsf{Pr}\left[\mathsf{Exp}^{\mathsf{ind\text{-}so\text{-}real}}_{\mathsf{Com},\mathcal{M},A} = 1\right] = \mathsf{Pr}\left[H_0 = 1\right].$$

We now define a generalization H_j: H_j runs like H_0, except that H_j runs this alternative step 2' instead of step 2:

2'. for every $i \le j$, compute $(com_i, dec_i) \leftarrow \mathsf{Com}(0^k)$; for $i > j$, compute $(com_i, dec_i) \leftarrow \mathsf{Com}(M_i)$.

Obviously, for $j = 0$ we get H_0. Note that H_j computes commitments com_i which, for $i \le j$, do no longer depend on M_i. From H_j, we can now construct an adversary A' on $(\mathsf{Com}, \mathsf{Ver})$'s statistical hiding property. A' first uniformly picks $j \in [n]$, then simulates H_{j-1}, but constructs com_j using its own experiment $\mathsf{Exp}^{\mathsf{hiding}\text{-}b}_{\mathsf{Com},A'}$. Namely, A' asks for a commitment to either M_j or 0^k, and uses the obtained com_j for further simulation in H_{j-1}. In $\mathsf{Exp}^{\mathsf{hiding\text{-}0}}_{\mathsf{Com},A'}$, this means that com_j is constructed as a commitment to M_j, and we obtain experiment H_{j-1}. On the other hand, in $\mathsf{Exp}^{\mathsf{hiding\text{-}1}}_{\mathsf{Com},A'}$, com_j is a commitment to 0^k, and we obtain experiment H_j. This way, we get that

$$\mathsf{Adv}^{\mathsf{hiding}}_{\mathsf{Com},A'} = \frac{1}{n}\left(\sum_{j=1}^{n} \mathsf{Pr}\left[H_j = 1\right] - \mathsf{Pr}\left[H_{j-1} = 1\right]\right) = \frac{1}{n}\left(\mathsf{Pr}\left[H_n = 1\right] - \mathsf{Pr}\left[H_0 = 1\right]\right)$$

is negligible, and hence so must be $\mathsf{Pr}\left[H_n = 1\right] - \mathsf{Pr}\left[H_0 = 1\right]$. Note that in H_n, the view of the adversary now only depends on M_I; all commitments are produced as commitments to 0^k.

With the same reasoning, we can show that the output of experiment $\mathsf{Exp}^{\mathsf{ind\text{-}so\text{-}ideal}}_{\mathsf{Com},\mathcal{M},A}$ is negligibly close to that of the analogously modified experiment H'_n, where all commitments are generated as commitments to 0^k. Since $H'_n = H_n$, we hence obtain that

$$\mathsf{Adv}^{\mathsf{ind\text{-}so}}_{\mathsf{Com},\mathcal{M},A} = \mathsf{Exp}^{\mathsf{ind\text{-}so\text{-}real}}_{\mathsf{Com},\mathcal{M},A} - \mathsf{Exp}^{\mathsf{ind\text{-}so\text{-}ideal}}_{\mathsf{Com},\mathcal{M},A}$$

must be negligible, which proves the theorem. $\square$