Possibility and impossibility results for selective decommitments

Dennis Hofheinz

Karlsruhe Institute of Technology
Dennis.Hofheinz@kit.edu

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 provable* for non-interactive or perfectly binding commitment schemes via black-box reductions to standard cryptographic assumptions. *However,* we show how to achieve security in this sense with interaction and a non-black-box reduction to one-way permutations.
2. If only indistinguishability of the unopened commitments from random commitments is desired, then security is *not provable* for (interactive or non-interactive) perfectly binding commitment schemes, via black-box reductions to standard cryptographic assumptions. *However,* any statistically hiding scheme *does* achieve security in this sense.

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.
- When implemented with our secure commitment scheme, the interactive proof for graph 3-coloring due to Goldreich et al. becomes zero-knowledge under parallel composition.

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

Keywords: commitments, zero-knowledge, black-box 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. [22], 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.

Previous 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., [11, 4, 12, 18, 14]). However, these works solved the problem by using (and, in fact, inventing) non-committing encryption, which circumvents the underlying commitment problem. In the zero-knowledge setting, Gennaro and Micali [25] notice a selective decommitment problem and circumvent it by adapting the distribution of the messages committed to. Similarly, a number of works (e.g., Dolev et al. [21], Prabhakaran et al. [40] in the context of zero-knowledge) use "cut-and-choose" techniques on committed values, which is a specific form of selective opening. These works can prove security by using specific properties of the distributions of the committed values (e.g., the fact that the unopened values, conditioned on the opened values, are still uniformly distributed).

Dwork et al. [22] is, to the best of our knowledge, the only previous work that explicitly studies the selective decommitment problem. They prove that a commitment scheme which is secure under selective openings would have interesting applications. In particular, they show that a (non-interactive) commitment scheme that is secure under selective openings gives rise to a 3-round zero-knowledge proof system for NP with negligible soundness error. 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 open the question whether commitment schemes secure under (general) selective openings exist.

From Goldreich and Krawczyk [27], it is known that 3-round black-box zero-knowledge proof systems exist only for languages in BPP. Let us denote a commitment scheme that, when plugged into the construction of [22], gives rise to a black-box zero-knowledge proof system, as *ZK-black-box.* Note that a *non-interactive* ZK-black-box commitment scheme gives rise to a 3-round black-box zero-knowledge proof system. Thus, combining the results of [27] and [22] shows that no non-interactive ZK-black-box commitment schemes exist (or NP$\subseteq$BPP). That is, [27, 22] essentially show that no non-interactive commitment scheme

[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.

exists that is secure under selective openings *and for which the simulator is constructed in a black-box way from the adversary (on the commitment security).* Jumping ahead, one of our results shows that no non-interactive commitment scheme that is secure under selective openings *and uses the computational assumption in a black-box way* exists. These are both negative, but orthogonal statements. Indeed, it is conceivable that a security reduction uses specific, non-black-box properties of the adversary (e.g., it is common in reductions to explicitly make use of the adversary's complexity bounds), but neither scheme nor reduction use specifics (like the code) of the underlying primitive.

Black-box impossibility results from generic assumptions have been derived by Dodis et al. [20]. They show that the security of full-domain hash signatures ([5]) cannot be proved using a black-box reduction to any hardness assumption that is satisfied by a random permutation. Concurrently to and independently from our work, Haitner and Holenstein [29] developed a framework to prove impossibility of black-box reductions from *any* computational assumption. While their formalism is very similar to ours (e.g., their definition of a "cryptographic game" matches our definition of a "security property"), they apply it to an entirely different problem. Namely, [29] prove black-box impossibility of encryption schemes secure in the presence of key-dependent messages.

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 that is secure in this sense *simulatable under selective openings.* We show that no non-interactive or perfectly binding commitment scheme can be proved simulatable under selective openings using black-box reductions to standard assumptions. However, we also show how to construct commitment schemes that *are* simulatable under selective openings, under the assumption that one-way permutations exist. Our construction uses non-black-box techniques (i.e., zero-knowledge proofs) as well as interaction to circumvent our impossibility results. This solves an important open problem from Dwork et al. [22]: our schemes are the first commitment schemes provably secure under selective openings.

We proceed to consider what happens if "security" means that A cannot distinguish the messages inside the unopened commitments from independent messages (where "independent" can of course only mean "independent, conditioned on the already opened messages"). We call a commitment scheme that is secure in this sense *indistinguishable under selective openings.* We show that no perfectly binding commitment scheme (interactive or not) can be proved indistinguishable under selective openings, via black-box reductions from standard assumptions. However, we also show that *all* statistically hiding commitment schemes *are* indistinguishable under selective openings.

Technically, we derive black-box impossibility results in the style of Impagliazzo and Rudich [35], but we can derive stronger claims, similar to Dodis et al. [20]. Concretely, we prove impossibility via $\forall\exists$semi-black-box proofs from *any* computational assumption that can be formalized as an oracle $\mathcal{X}$ and a corresponding security property $\mathcal{P}$ 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, somewhat surprisingly, our impossibility claim holds even if $\mathcal{P}$ models security under selective openings. In that case, however, a reduction will necessarily be non-black-box, see Section A 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 non-standard techniques. Namely, whenever a committing (i.e., ciphertexts commit to messages) encryption scheme is adaptively secure, then it also is, interpreted as a (non-interactive) commitment scheme, simulatable under selective openings. Our impossibility results show that hence, a committing encryption scheme cannot be proved adaptively secure via black-box reductions from standard assumptions.

Second, we apply our results to "commit-choose-open" (CCO) style interactive proof systems such as the graph 3-coloring protocol $\mathsf{G3C}$ from Goldreich et al. [28]. Refining the techniques of Dwork et al. [22], we prove that any CCO protocol becomes zero-knowledge under parallel composition, when implemented with a commitment scheme which is simulatable under selective openings. In particular, our (interactive, but constant-round) commitment scheme enables the parallel composability of $\mathsf{G3C}$. We also show that a CCO protocol becomes witness-indistinguishable, even under parallel composition, when implemented with a commitment scheme which is indistinguishable under selective openings. This shows the usefulness of our indistinguishability-based security definition as a reasonable fallback.

Organization. After fixing some notation in Section 2, we present in Section 3 our possibility and impossibility results for the simulation-based security definition of Dwork et al. [22]. We give an indistinguishability-based security definition, along with possibility and impossibility results in Section 4. In Sections 5 and 6, we consider applications of our results to encryption and interactive proof systems. In Section A, we discuss the role of the computational assumption in our impossibility results.

Publication note and follow-up work. This work constitutes the full version of one part of the Eurocrypt 2009 paper "Possibility and impossibility results for encryption and commitment secure under selective opening" by Bellare et al. [7]. The paper [7] is the merge of two Eurocrypt submissions on the topic of security under selective openings. One submission (by Bellare and Yilek) treated the encryption case, while the other (by Hofheinz) treated the commitment case. The present work is the full version of the latter submission.

Very recently, Hemenway and Ostrovsky [32] have used and improved on the results of [7]. Hemenway and Ostrovsky provide efficient encryption schemes secure under selective openings, along with a generic construction to achieve security even under chosen-ciphertext attacks. They also show that a special type of randomized one-way functions give rise to non-interactive commitment schemes that are secure under selective openings. Their results do not contradict our impossibility results since they consider a trusted set-up of public parameters. (In other words, commitments are performed and checked relative to ideally chosen public parameters.) A more detailed discussion of the ideas of [7, 32], as well as of our impossibility results in the context of trusted set-up information can be found in Section 5.

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 PPT algorithm/machine is a probabilistic algorithm/machine which runs in time polynomial in k. While an algorithm is stateless, a machine maintains a state across activations. 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 $\mathbf{M} = (M_i)_i$ is an indexed set, then we write $M_I := (M_i)_{i \in I}$. We denote the empty (bit-)string by ϵ.

Commitment schemes.

Definition 1 (Commitment scheme). *For a pair of PPT machines* $\mathsf{Com} = (\mathsf{S}, \mathsf{R})$ *and a machine A, consider the following experiments:*

```
Experiment $\mathsf{Exp}^{\mathsf{binding}}_{\mathsf{Com},A}$
  run $\langle \mathsf{R}(\mathtt{recv}), A(\mathtt{com}) \rangle$
  $M'_0 \leftarrow \langle \mathsf{R}(\mathtt{open}), A(\mathtt{open}, 0) \rangle$
  rewind $A$ and $\mathsf{R}$ back to after step 1
  $M'_1 \leftarrow \langle \mathsf{R}(\mathtt{open}), A(\mathtt{open}, 1) \rangle$
  return 1 iff $\bot \neq M'_0 \neq M'_1 \neq \bot$
```

```
Experiment $\mathsf{Exp}^{\mathsf{hiding}\text{-}b}_{\mathsf{Com},A}$
  $(M_0, M_1) \leftarrow A(\mathtt{choose})$
  return $\langle A(\mathtt{recv}), \mathsf{S}(\mathtt{com}, M_b) \rangle$
```

In this, $\langle A, \mathsf{S} \rangle$ denotes the output of A after interacting with S*, and $\langle \mathsf{R}, A \rangle$ denotes the output of* R *after interacting with A. We say that* Com *is a* commitment scheme *iff the following holds:*

Syntax. *For any $M \in \{0,1\}^k$, $\mathsf{S}(\mathtt{com}, M)$ first interacts with $\mathsf{R}(\mathtt{recv})$. We call this the* commit phase*. After that, $\mathsf{S}(\mathtt{open})$ interacts again with $\mathsf{R}(\mathtt{open})$, and R finally outputs a value $M' \in \{0,1\}^k \cup \{\bot\}$. We call this the* opening phase.

Correctness. *We have $M' = M$ always and for all M.*

Hiding. *For a PPT machine A, let*

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

where $\mathsf{Exp}^{\mathsf{hiding}\text{-}b}_{\mathsf{Com},A}$ *is depicted below. For* Com *to be hiding, we demand that* $\mathsf{Adv}^{\mathsf{hiding}}_{\mathsf{Com},A}$ *is negligible for all PPT A that satisfy* $M_0, M_1 \in \{0,1\}^k$ *always.*

Binding. *For a machine A, consider the experiment* $\mathsf{Exp}^{\mathsf{binding}}_{\mathsf{Com},A}$ *below. For* Com *to be binding, we require that* $\mathsf{Adv}^{\mathsf{binding}}_{\mathsf{Com},A} = \Pr\left[\mathsf{Exp}^{\mathsf{binding}}_{\mathsf{Com},A} = 1\right]$ *is negligible for all PPT A.*

Further, we say that Com *is* perfectly binding *iff* $\mathsf{Adv}^{\mathsf{binding}}_{\mathsf{Com},A} = 0$ *for all A. We say that* Com *is* statistically hiding *iff* $\mathsf{Adv}^{\mathsf{hiding}}_{\mathsf{Com},A}$ *is negligible for all (not necessarily PPT) A.*

Definition 2 (Non-interactive commitment scheme). *A* non-interactive commitment scheme *is a commitment scheme* $\mathsf{Com} = (\mathsf{S}, \mathsf{R})$ *in which both commit and opening phase consist of only one message sent from* S *to* R*. We can treat a non-interactive commitment scheme as a pair of algorithms rather than machines. Namely, we write* $(com, dec) \leftarrow \mathsf{S}(M)$ *shorthand for the commit message com and opening message dec sent by* S *on input M. We also denote by* $M' \leftarrow \mathsf{R}(com, dec)$ *the final output of* R *upon receiving com in the commit phase and dec in the opening phase.*

Note that perfectly binding implies that *any* commitment can only be opened to at most one value M. Perfectly binding (non-interactive) commitment schemes can be achieved from any one-way permutation (e.g., Blum [8]). On the other hand, statistically hiding implies that for any $M_0, M_1 \in \{0,1\}^k$, the statistical distance between the respective views of the receiver in the commit phase is negligible. One-way functions suffice to implement statistically hiding (interactive) commitment schemes (Haitner and Reingold [30]), but there are certain lower bounds for the communication complexity of such constructions (Wee [45], Haitner et al. [31]). However, if we assume the existence of (families of) collision-resistant hash functions, then even constant-round statistically hiding commitment schemes exist (Damgård et al. [19], Naor and Yung [38]).

Interactive argument systems and zero-knowledge. We recall some basic definitions concerning interactive argument systems, mostly following Goldreich [26].

Definition 3 (Interactive proof/argument system). *An* interactive proof system *for a language $\mathcal{L}$ with witness relation $\mathcal{R}$ is a pair of PPT machines* $\mathsf{IP} = (\mathsf{P}, \mathsf{V})$ *such that the following holds:*

Completeness. *For every family* $(x_k, w_k)_{k\in\mathbb{N}}$ *such that* $\mathcal{R}(x_k, w_k)$ *for all k and* $|x_k|$ *is polynomial in k, we have that the probability for* $\mathsf{V}(x_k)$ *to output 1 after interacting with* $\mathsf{P}(x_k, w_k)$ *is at least* $2/3$*.*

Soundness. *For every machine* P^* *and every family* $(x_k, z_k)_{k\in\mathbb{N}}$ *such that* $|x_k| = k$ *and* $x_k \notin \mathcal{L}$ *for all k, we have that the probability for* $\mathsf{V}(x_k)$ *to output 1 after interacting with* $P^*(x_k, z_k)$ *is at most* $1/3$*.*

If the soundness condition holds for all PPT machines P^* *(but not necessarily for all unbounded* P^**), then* IP *is an* interactive argument system*. We say that* IP *enjoys* perfect completeness *if* V *always outputs 1 in the completeness condition.*

Furthermore, IP *has* negligible soundness error *if* V *outputs 1 only with negligible probability in the soundness condition.*

Definition 4 (Zero-knowledge). *Let* $\mathsf{IP} = (\mathsf{P}, \mathsf{V})$ *be an interactive proof or argument system for language* $\mathcal{L}$ *with witness relation* $\mathcal{R}$*.* IP *is* zero-knowledge *if for every PPT machine* V^**, there exists a PPT machine* S^* *such that for all sequences* $(x, w) = (x_k, w_k)_{k\in\mathbb{N}}$ *with* $\mathcal{R}(x_k, w_k)$ *for all* k *and* $|x_k|$ *polynomial in* k*, for all PPT machines* D*, and all auxiliary inputs* $z^{V^*} = (z_k^{V^*})_{k\in\mathbb{N}} \in (\{0,1\}^*)^{\mathbb{N}}$ *and* $z^D = (z_k^D)_{k\in\mathbb{N}} \in (\{0,1\}^*)^{\mathbb{N}}$*, we have that*

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

is negligible in k*. Here* $\langle \mathsf{P}(x_k, w_k), V^*(x_k, z_k^{V^*})\rangle$ *denotes the transcript of the interaction between the prover* P *and* V^**.*

Most known interactive proof system achieve perfect completeness. Conversely, most systems do not enjoy a negligible soundness error "by nature"; their soundness has to be amplified via repetition, e.g., via sequential or concurrent composition. Thus, it is important to consider the concurrent composition of an interactive argument system:

Definition 5 (Concurrent zero-knowledge). *Let* $\mathsf{IP} = (\mathsf{P}, \mathsf{V})$ *be an interactive proof or argument system for language* $\mathcal{L}$ *with witness relation* $\mathcal{R}$*.* IP *is* zero-knowledge under concurrent composition *iff for every polynomial* $n = n(k)$ *and PPT machine* V^**, there exists a PPT machine* S^* *such that for all sequences* $(x, w) = (x_{i,k}, w_{i,k})_{k\in\mathbb{N},i\in[n]}$ *with* $\mathcal{R}(x_{i,k}, w_{i,k})$ *for all* i, k *and* $|x_{i,k}|$ *polynomial in* k*, for all PPT machines* D*, and all auxiliary inputs* $z^{V^*} = (z_k^{V^*})_{k\in\mathbb{N}} \in (\{0,1\}^*)^{\mathbb{N}}$ *and* $z^D = (z_k^D)_{k\in\mathbb{N}} \in (\{0,1\}^*)^{\mathbb{N}}$*, we have that*

$$\begin{aligned}&\mathsf{Adv}^{\mathsf{cZK}}_{V^*,S^*,(x,w),D,z^{V^*},z^D} := \\ &\quad \mathsf{Pr}\left[D((x_{i,k})_{i\in[n]}, z_k^D, \langle \mathsf{P}((x_{i,k}, w_{i,k})_{i\in[n]}), V^*((x_{i,k})_{i\in[n]}, z_k^{V^*})\rangle) = 1\right] \\ &\qquad - \mathsf{Pr}\left[D((x_{i,k})_{i\in[n]}, z_k^D, S^*((x_{i,k})_{i\in[n]}, z_k^{V^*})) = 1\right]\end{aligned}$$

is negligible in k*. Here* $\langle \mathsf{P}((x_{i,k}, w_{i,k})_{i\in[n]}), V^*((x_{i,k})_{i\in[n]}, z_k^{V^*})\rangle$ *denotes the transcript of the interaction between* n *copies of the prover* P *(with the respective inputs* $(x_{i,k}, w_{i,k})$ *for* $i = 1, \dots, n$*) on the one hand, and* V^* *on the other hand.*

There exist interactive proof systems (with perfect completeness and negligible soundness error) that achieve Definition 5 for arbitrary NP-languages if one-way permutations exist (e.g., Richardson and Kilian [42]; see also [36, 13, 1, 23, 3] for similar results in related settings). If we assume the existence of (families of) collision-resistant hash functions, then there even exist constant-round interactive proof systems that achieve a bounded version of Definition 5 in which the number of concurrent instances is fixed in advance (Barak [1], Barak and Goldreich [2]).[2]

Black-box reductions. Reingold et al. [41] give an excellent overview and classification of black-box reductions. We recall some of their definitions which are important for our case. A *primitive* $\mathsf{P} = (F_\mathsf{P}, R_\mathsf{P})$ is a set F_P of functions $f : \{0,1\}^* \to \{0,1\}^*$ along with a relation R over pairs (f, A), where $f \in F_\mathsf{P}$, and A is a machine. We say that f is an *implementation* of P iff $f \in F_\mathsf{P}$. Furthermore, f is an *efficient implementation* of P iff $f \in F_\mathsf{P}$ and f can be computed by a PPT machine. A machine A P-*breaks* $f \in F_\mathsf{P}$ iff $R_\mathsf{P}(f, A)$. A primitive P *exists* if there is an efficient implementation $f \in F_\mathsf{P}$ such that no PPT machine P-breaks f. A primitive P *exists relative to an oracle* $\mathcal{B}$ iff there exists an implementation $f \in F_\mathsf{P}$ which is computable by a PPT machine with access to $\mathcal{B}$, such that no PPT machine with access to $\mathcal{B}$ P-breaks f.

Definition 6 (Relativizing reduction). *There exists a* relativizing reduction *from a primitive* $\mathsf{P} = (F_\mathsf{P}, R_\mathsf{P})$ *to a primitive* $\mathsf{Q} = (F_\mathsf{Q}, R_\mathsf{Q})$ *iff for every oracle* $\mathcal{B}$*, the following holds: if* Q *exists relative to* $\mathcal{B}$*, then so does* P*.*

Definition 7 ($\forall\exists$semi-black-box reduction). *There exists a* $\forall\exists$semi-black-box reduction *from a primitive* $\mathsf{P} = (F_\mathsf{P}, R_\mathsf{P})$ *to a primitive* $\mathsf{Q} = (F_\mathsf{Q}, R_\mathsf{Q})$ *iff for every implementation* $f \in F_\mathsf{Q}$*, there exists a PPT machine* G *such that* $G^f \in F_\mathsf{P}$*, and the following holds: if there exists a PPT machine* A *such that* A^f P*-breaks* G^f*, then there exists a PPT machine* S *such that* S^f Q*-breaks* f*.*

It can be seen that if a relativizing reduction exists, then so does a $\forall\exists$semi-black-box reduction. The converse is true when Q "allows embedding," which essentially means that additional oracles can be embedded into Q without destroying its functionality (see Reingold et al. [41], Definition 3.4 and Theorem 3.5 and Simon [44]). Below we will prove impossibility of relativizing reductions between certain primitives, which also proves impossibility of $\forall\exists$semi-black-box reductions, since the corresponding primitives Q allow embedding.

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 Dwork et al. [22] 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 $R(\mathbf{M}, \mathit{out}_A)$ holds about as often as $R(\mathbf{M}, \mathit{out}_S)$, where $\mathbf{M} = (M_i)_{i\in[n]}$ 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):

[2] It is common to allow the simulator S^* to be *expected polynomial-time*. In fact, the positive results [42, 36] (but not [1]) construct an expected PPT S^*. We will neglect this issue in the following, since our results do not depend the complexity of S^* (as long as S^* is not able to break an underlying computational assumption).

Definition 8 (Simulatable under selective openings). *Assume $n = n(k) > 0$ is polynomially bounded, 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 $\mathsf{Com} = (\mathsf{S}, \mathsf{R})$ is* simulatable under selective openings *(short* SIM-SO-COM secure*) iff for every PPT n-message distribution $\mathcal{M}$, every PPT relation R, and every PPT machine A (the adversary), there is a PPT machine S (the simulator), such that $\mathsf{Adv}^{\mathsf{sim\text{-}so}}_{\mathsf{Com},\mathcal{M},A,S,R}$ is negligible. Here*

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

where the experiments $\mathsf{Exp}^{\mathsf{sim\text{-}so\text{-}real}}_{\mathsf{Com},\mathcal{M},A,R}$ and $\mathsf{Exp}^{\mathsf{sim\text{-}so\text{-}ideal}}_{\mathcal{M},S,R}$ are defined as follows:

```
Experiment $\mathsf{Exp}^{\mathsf{sim\text{-}so\text{-}real}}_{\mathsf{Com},\mathcal{M},A,R}$
  $\mathbf{M} = (M_i)_{i\in[n]} \leftarrow \mathcal{M}$
  $I \leftarrow \langle A(\mathtt{recv}), (\mathsf{S}_i(\mathtt{com}, M_i))_{i\in[n]}\rangle$
  $out_A \leftarrow \langle A(\mathtt{open}), (\mathsf{S}_i(\mathtt{open}))_{i\in I}\rangle$
  return $R(\mathbf{M}, out_A)$
```

```
Experiment $\mathsf{Exp}^{\mathsf{sim\text{-}so\text{-}ideal}}_{\mathcal{M},S,R}$
  $\mathbf{M} = (M_i)_{i\in[n]} \leftarrow \mathcal{M}$
  $I \leftarrow S(\mathtt{choose})$
  $out_S \leftarrow S((M_i)_{i\in I})$
  return $R(\mathbf{M}, out_S)$
```

In this, we require from A that $I \in \mathcal{I}_k$,[3] *and we denote by $\langle A, (\mathsf{S}_i)_i\rangle$ the output of A after interacting concurrently with instances S_i of S.*

Discussion of the definitional choices. While Definition 8 essentially is the selective decommitment definition Dwork et al. [22], Definition 7.1, there are a number of definitional choices we would like to highlight (the following discussion applies equally to the upcoming Definition 10):

- Unlike [22, Definition 7.1], neither adversary A nor relation R get an auxiliary input. Such an auxiliary input is common in cryptographic definitions to ensure some form of composability.
- We do not explicitly hand the chosen set I to the relation R. Handing I to R potentially makes the definition more useful in larger contexts in which I is public.
- One could think of letting R determine the message vector $\mathbf{M}$.[4] (Equivalently, we can view $\mathcal{M}$ as part of R and let $\mathcal{M}$ forward its random coins—or a short seed—to R in a message part M_i which is guaranteed not to be opened, e.g., when $i \notin I$ for all $I \in \mathcal{I}_n$.)
- The order of quantifiers ($\forall \mathcal{M}, R, A \exists S$) is the weakest one possible. In particular, we do not mandate that S is constructed from A in a black-box way.

[3] that is, we actually only quantify over those A for which $I \in \mathcal{I}_k$

[4] This definition is closer to a universally composable definition (cf. Canetti [9]) in the sense that R (almost) takes the role of a UC-environment: R selects all inputs and reads the outputs (in particular the output of A). However, we stress that R may not actively interfere in the commitment protocol. Note that we cannot hope for *fully* UC-secure commitments for reasons not connected to the selective decommitment problem, cf. Canetti and Fischlin [10].

In all of the cases, we chose the weaker definitional variant for simplicity, which makes our negative results only stronger. We stress, however, that our positive results (Theorem 2 and Theorem 4) hold also for all of the stronger definitional variants.

3.1 Impossibility from black-box reductions

Formalization of computational assumptions. Our first result states that SIM-SO-COM security cannot be achieved via black-box reductions from standard assumptions. We want to consider such standard assumptions in a general way that allows to make statements even in the presence of "relativizing" oracles. Thus we make the following definition, which is a special case of the definition of a *primitive* from Reingold et al. [41] (cf. also Section 2).

Definition 9 ((Security) property of an oracle). *Let $\mathcal{X}$ be an oracle. Then a* security property *(or simply* property*) $\mathcal{P}$ of $\mathcal{X}$ is a (not necessarily PPT) machine that, after 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*

$$\mathsf{Adv}^{\mathsf{prop}}_{\mathcal{P},\mathcal{X},A} := \mathsf{Pr}\left[\mathcal{P} \textit{ outputs } b = 1 \textit{ after interacting with } A \textit{ and } \mathcal{X}\right] - 1/2.$$

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

In terms of Reingold et al. [41], the corresponding primitive is $\mathsf{P} = (F_{\mathsf{P}}, R_{\mathsf{P}})$, where $F_{\mathsf{P}} = \{\mathcal{X}\}$, and $R_{\mathsf{P}}(\mathcal{X}, A)$ iff $\mathsf{Adv}^{\mathsf{prop}}_{\mathcal{P},\mathcal{X},A}$ is non-negligible. Our definition is also similar in spirit to "hard games" as used by Dodis et al. [20], but more general.

We emphasize that $\mathcal{P}$ can *only* interact with $\mathcal{X}$ and A, but not with possible additional oracles. (See Section A 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 whether adversary A can "break" $\mathcal{X}$ in the intended way. We give an example, where "breaking" means "breaking $\mathcal{X}$'s one-way property".

Finally, note that a security property as above does not allow to capture statistical properties such as the bijection property of a function, or the (perfect/statistical) correctness property of an encryption scheme. In what follows, we will always consider a fixed oracle $\mathcal{X}$. Its statistical properties will be clear by definition, and in particular they will not change depending on the setting we consider $\mathcal{X}$ in. However, the security properties (such as one-wayness) may change, depending on auxiliary oracles that may help to "break" $\mathcal{X}$.

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 $\Pr[\mathcal{P}$ outputs 1$]$ 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 encryption scheme $\mathcal{X}$) 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 (black-box) security reductions, from essentially *any* computational assumption (i.e., security 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 black-box." We give a detailed explanation of what is going on in Section A.

Stateless breaking oracles. In our impossibility results, we will describe a computational world with a number of oracles. For instance, there will be a "breaking oracle" $\mathcal{B}$, such that $\mathcal{B}$ aids in breaking the SIM-SO-COM security of any given commitment scheme, and in *nothing more.* To this end, $\mathcal{B}$ takes the role of the adversary in the SIM-SO-COM experiment. Namely, $\mathcal{B}$ expects to receive a number of commitments, then chooses a subset of these commitments, and then expects openings of the commitments in this subset. This is an interactive process which would usually require $\mathcal{B}$ to hold a state across invocations. However, stateful oracles are not very useful for establishing black-box separations, so we will have to give a stateless formulation of $\mathcal{B}$. Concretely, suppose that the investigated commitment scheme is non-interactive. Then $\mathcal{B}$ answers deterministically upon queries and expects each query to be prefixed with the history of that query. For instance, $\mathcal{B}$ finally expects to receive openings $\mathit{dec} = (\mathit{dec}_i)_{i \in I}$ *along* with the corresponding previous commitments $\mathit{com} = (\mathit{com}_i)_{i \in [n]}$ and previously selected set I. If I is not the set that $\mathcal{B}$ would have selected when receiving com alone, then $\mathcal{B}$ ignores the query. This way, $\mathcal{B}$ is stateless (but randomized, similarly to a random oracle). Furthermore, for non-interactive commitment schemes, this makes sure that any machine interacting with $\mathcal{B}$ can open commitments to $\mathcal{B}$ only in one way. Hence this formalization preserves the binding property of a commitment scheme, something which we will need in our proofs.

We stress, however, that this method does not necessarily work for interactive commitment schemes. Namely, any machine interacting with such a stateless $\mathcal{B}$ can essentially "rewind" $\mathcal{B}$ during an interactive commitment phase, since $\mathcal{B}$ formalizes a next-message function. Now if the commitment scheme is still binding if the receiver of the commitment can be rewound (e.g., this holds trivially for non-interactive commitment schemes, and also for perfectly binding commitment schemes), then our formalization of $\mathcal{B}$ preserves binding, and our upcoming proof works. If, however, the commitment scheme loses its binding property if the receiver can be rewound, then the following theorem cannot be applied.

We are now ready to state our result.

Theorem 1 (Black-box impossibility of non-interactive or perf. binding SIM-SO-COM, most general formulation). *Let $n = n(k) = 2k$, and let $\mathcal{I} = (\mathcal{I}_n)_n$ with $\mathcal{I}_n = \{I \subseteq [n] \mid |I| = n/2\}$ denote the set of all $n/2$-sized subsets of $[n]$.*[5] *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 non-interactive or perfectly binding commitment scheme which is simulatable under selective openings.*

Proof strategy. We will use a random oracle $\mathcal{RO}$ that, for any given non-interactive commitment scheme Com^*, induces a message distribution $\mathcal{M}^* = \{(\mathcal{RO}(\mathsf{Com}^*, i, X^*))_{i\in[n]}\}_{X^*\in\{0,1\}^{k/3}}$. Here, $\mathcal{RO}(\mathsf{Com}^*)$ denotes the hash of the description of Com^*, and X^* is a short "seed" that ties the values $\mathcal{RO}(\mathsf{Com}^*, i, X^*)$ (with the same X^* but different i) together. Furthermore, we will specify an oracle $\mathcal{B}$ that will help to break Com^* with respect to $\mathcal{M}^*$. Concretely, $\mathcal{B}$ first expects n Com^*-commitments, and then requests openings of a random subset of them. If all openings are valid, $\mathcal{B}$ returns a value X^* consistent (according to $\mathcal{M}^*$) with all opened messages (if such an X^* exists). A suitable SIM-SO-COM adversary A can use $\mathcal{B}$ simply by relaying its challenge to obtain X^* and hence the whole message vector in its SIM-SO-COM experiment.

However, we will prove that $\mathcal{B}$ is useless to any simulator S that gets only a message subset M_I: if S uses $\mathcal{B}$ *before* requesting its own message subset M_I, then $\mathcal{B}$'s answer will not be correlated with the SIM-SO-COM challenge message vector $\mathbf{M}$. (This also holds if S first sends commitments to $\mathcal{B}$ and immediately afterwards requests M_I from the SIM-SO-COM experiment; in that case, S has to break the binding property of Com^* to get an answer from $\mathcal{B}$ which is correlated with $\mathbf{M}$.) But if S uses $\mathcal{B}$ *after* obtaining M_I, then with very high probability, S will have to open at least one commitment to $\mathcal{B}$ whose message is not contained in M_I. By definition of $\mathcal{M}^*$, this opening of S will not be consistent with the other values of M_I (except with small probability), and $\mathcal{B}$'s answer will again not be correlated with $\mathbf{M}$.

Since S cannot efficiently extract the seed X^* from its message subset M_I alone (that would require a brute-force search over exponentially many values), this shows that Com^* is not SIM-SO-COM secure. Consequently, because Com^* was arbitrary (only the message distribution $\mathcal{M}^*$ is specific to Com^*), there exist no SIM-SO-COM secure commitment schemes relative to $\mathcal{RO}$ and $\mathcal{B}$. Finally, it is easy to see that relative to $\mathcal{RO}$ and $\mathcal{B}$, primitive $\mathcal{X}$ still satisfies property $\mathcal{P}$. Concretely, observe that $\mathcal{B}$ does not break any commitment (note that $\mathcal{B}$'s answer depends only on the *opened* commitments), but only inverts a message distribution (or, rather, $\mathcal{RO}$). Hence, any adversary attacking property $\mathcal{P}$ of $\mathcal{X}$ can use efficient internal simulations of $\mathcal{RO}$ and $\mathcal{B}$ instead of the original oracles. Since $\mathcal{X}$ satisfies property $\mathcal{P}$ with respect to adversaries without (additional) oracle access, the claim follows.

We commence with the full proof.

[5] We stress that the proofs of Theorem 1 and Theorem 3 hold literally also for the "cut-and-choose" $\mathcal{I}_n = \{I \subseteq [n] \mid \forall i \in [k] :\ \text{either } 2i-1 \in I \text{ or } 2i \in I\}$.

Proof (of Theorem 1). First, let $\mathcal{RO}$ be a random oracle (i.e., a random function $\{0,1\}^* \to \{0,1\}^k$). When writing $\mathcal{RO}(x_1, \ldots, x_\ell)$, we assume that $\mathcal{RO}$'s input $x_1, \ldots, x_\ell$ is encoded in a prefix-free way, such that all individual x_i can be efficiently reconstructed from $\mathcal{RO}$'s input. Furthermore, to derive our second oracle $\mathcal{B}$, first consider the following machine B:

1. On input Com, interpret Com as the description of two machines (S, R) as in Definition 1. Then, concurrently receive n Com-commitments, indexed by $i \in [n]$.
2. When all commitments are received, output a uniformly chosen $I \in \mathcal{I}$.
3. Engage in $|I|$ concurrent opening phases for the Com-instances with $i \in I$. If all openings are valid (i.e., every receiver instance with $i \in I$ outputs some $M_i \neq \perp$), return the set of all $X \in \{0,1\}^{k/3}$ such that $M_i = \mathcal{RO}(\mathsf{Com}, i, X)$ for all $i \in I$.

Unfortunately, we cannot use B directly in our proof, since B is stateful, and black-box separations require stateless oracles. So let $\mathcal{B}$ be the oracle that evaluates B's *next-message function*. Formally, $\mathcal{B}$ expects queries of the form $h = (h_i)_{i \in [\ell]}$. Upon each such query, $\mathcal{B}$ invokes a fresh copy of B, and feeds it input messages h_1 up to h_ℓ successively, ignoring the respective answers of B. Finally, $\mathcal{B}$ outputs B's answer to the last input h_ℓ. The random coins used for B in a given activation are supplied by $\mathcal{B}$ as a random (but deterministic) function of the previous message history of B. This way, $\mathcal{B}$ itself is randomized but stateless, and can be used to emulate interactions with B. (In fact, $\mathcal{B}$ models a B which can be rewound.)

We now comment on the description of Com that $\mathcal{B}$ receives. Com describes two machines S and R, which may make arbitrary oracle calls (even recursive $\mathcal{B}$-queries[6]). We make no requirement that Com describes a hiding, binding, or correct commitment scheme. However, we do require that S and R are PPT whenever the description Com is generated by a PPT algorithm. We achieve this with a suitable padding: We require that all $\mathcal{B}$-queries h are prefixed with 1^ℓ, where ℓ bounds $\mathcal{B}$'s running time on input h. Here, we count any oracle query with input x as $|x|$ computational steps, and the final computation of all X as one step. This way, not even recursive $\mathcal{B}$-queries consume more than overall ℓ steps (not measuring the time needed to parse ℓ), while any PPT commitment scheme Com can still be encoded efficiently.

For a query $h = (h_i)_{i \in [\ell]}$, let $I^h \in \mathcal{I}$ and ${M^h}_{I^h} = ({M^h}_i)_{i \in I^h}$ denote the variables from the corresponding interaction with B. For a commitment scheme Com and a machine A, we say that A *breaks* Com^* *in* $\mathcal{B}$ iff A manages to output two queries $h = (h_i)_{i \in [\ell]}$ and $h' = (h'_i)_{i \in [\ell']}$ such that the following holds.

[6] Recursive $\mathcal{B}$-queries can be circumvented using the "two-oracle"-technique of Hsiao and Reyzin [34]. Adapted to our setting, we would only have to consider commitment schemes Com which are formulated *independently* of the breaking oracle $\mathcal{B}$, so we can assume that Com itself does not query $\mathcal{B}$. This would directly prove *fully-black-box* reductions impossible. However, at the cost of a little additional care in the mere encoding of our queries (so as to avoid unbounded recursions), we can even show impossibility of relativizing and hence of ∀∃semi-black-box reductions.

- $h_i = h'_i$ for all $i \leq i^I$, where i^I is the (unique) index for which $\mathcal{B}((h_i)_{i\in[i^I]})$ outputs $I^h \in \mathcal{I}$.
- There is an index $j \in [n]$ such that $\bot \neq {M^h}_j \neq {M^{h'}}_j \neq \bot$.

In other words, this holds if A manages to produce interactions with B in which the same commitment is opened in different ways.

From here on, fix a (hiding and binding) commitment scheme $\mathsf{Com}^* = (\mathsf{S}^*, \mathsf{R}^*)$, such that Com^* is non-interactive or perfectly binding (or both). We first show that our modeling of $\mathcal{B}$ preserves the binding property of Com^*.

Lemma 1. *No PPT adversary A breaks Com^* in $\mathcal{B}$ with non-negligible probability.*

Proof. If Com^* is perfectly binding, there never exists a commitment for which two different openings are possible (as long as the receiver acts honestly). Hence there simply are no h and h' as required to break the binding property of Com^* in $\mathcal{B}$. On the other hand, if Com^* is non-interactive, then A must find a non-interactive commitment com along with two non-interactive openings dec_1 and dec_2 in order to break Com^* in $\mathcal{B}$. The (ordinary) binding property of Com^* implies that this is not efficiently possible.

Now consider the distribution $\mathcal{M}^* = \{(\mathcal{RO}(\mathsf{Com}^*, i, X^*))_{i\in[n]}\}_{X^*\in\{0,1\}^{k/3}}$ of message vectors (i.e., $\mathcal{M}^*$ chooses $X^* \in \{0,1\}^{k/3}$ uniformly and then sets ${M^*}_i = \mathcal{RO}(\mathsf{Com}^*, i, X^*)$ for all i).

Lemma 2. *There is an adversary A that outputs $out_A = \mathbf{M}^*$ with overwhelming probability in the real SIM-SO-COM experiment $\mathsf{Exp}^{\mathsf{sim\text{-}so\text{-}real}}_{\mathsf{Com}^*,\mathcal{M},A,R}$. Here $\mathbf{M}^*$ denotes the full message vector sampled from $\mathcal{M}^*$ by the experiment.*

Proof. Let A be the SIM-SO-COM adversary on Com^* that relays between its interface to the SIM-SO-COM experiment and $\mathcal{B}$ as follows. We silently assume that A prefixes queries to $\mathcal{B}$ with the respective message history, and applies a padding as described above.

1. Initially, send Com^* to $\mathcal{B}$.
2. Relay the n commitments from the SIM-SO-COM experiment to $\mathcal{B}$.
3. Upon receiving $I^* \in \mathcal{I}$ from $\mathcal{B}$, send I^* to the SIM-SO-COM experiment.
4. Upon receiving $|I^*|$ openings from the experiment, relay these openings to $\mathcal{B}$.
5. Upon receiving a set $\{X^*\}$ from $\mathcal{B}$, return $out_A = (\mathcal{RO}(\mathsf{Com}^*, i, X^*))_{i\in[n]}$. If $\mathcal{B}$ returns a set of larger size, return $out_A = \bot$.

By construction of $\mathcal{M}^*$ and $\mathcal{B}$, it is clear that $out_A = \mathbf{M}^*$ unless $\mathcal{B}$ returns multiple X (which happens only with negligible probability by a counting argument).

Lemma 3. *Any given PPT simulator S will output $out_S = \mathbf{M}^*$ 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 X^*, and hence will output $out_S = \mathbf{M}^*$ 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 $\mathsf{Com}^j, I^j, {M^j}_{I^j}$ the corresponding values used in S's j-th query $h^j = (h^j_i)_{i\in[\ell^j]}$ to $\mathcal{B}$. Here and in the following, we consider recursive $\mathcal{B}$-queries made by $\mathcal{B}$ during the verification of openings as made by S. We first define and bound a number of "bad" events:

- $\mathsf{bad}_{\mathsf{coll}}$ occurs iff S reveals a message ${M^j}_i$ to $\mathcal{B}$ for which there are two distinct $X^1, X^2 \in \{0,1\}^{k/3}$ with $\mathcal{RO}(\mathsf{Com}^j, i, X^1) = {M^j}_i = \mathcal{RO}(\mathsf{Com}^j, i, X^2)$.
- $\mathsf{bad}_{\mathsf{img}}$ occurs iff S reveals a message ${M^j}_i$ to $\mathcal{B}$ for which an X with ${M^j}_i = \mathcal{RO}(\mathsf{Com}^j, i, X)$ exists, but ${M^j}_i$ has not been obtained through an explicit $\mathcal{RO}$-query (by either S or the SIM-SO-COM experiment).
- $\mathsf{bad}_{\mathsf{bind}}$ occurs iff $(\mathsf{Com}^j, I^j, {M^j}_{I^j}) = (\mathsf{Com}^*, I^*, {M^*}_{I^*})$ for some j.
- $\mathsf{bad} := \mathsf{bad}_{\mathsf{coll}} \vee \mathsf{bad}_{\mathsf{img}} \vee \mathsf{bad}_{\mathsf{bind}}$.

These events occur only with negligible probability: informally, $\mathsf{bad}_{\mathsf{coll}}$ implies a collision among $2^{k/3}$ uniformly distributed k-bit values, which is ruled out by a birthday bound. $\mathsf{bad}_{\mathsf{img}}$ means that S guessed an element of a very sparse set. Finally, $\mathsf{bad}_{\mathsf{bind}}$ means that S broke Com^*'s binding property (or, rather, S broke Com^* in $\mathcal{B}$). A detailed proof can be found in Lemma 4 below.

Now consider the following machine B' which is almost identical to B (the difference to B is *emphasized*):

1. On input Com, interpret Com as the description of two machines (S, R) as in Definition 1. Then, concurrently receive n Com-commitments, indexed by $i \in [n]$.
2. When all commitments are received, output a uniformly chosen $I \in \mathcal{I}$.
3. Engage in $|I|$ concurrent opening phases for the Com-instances with $i \in I$. If all openings are valid (i.e., every receiver instance with $i \in I$ outputs some $M_i \neq \perp$), proceed as follows. *If every M_i is the result of an $\mathcal{RO}(\mathsf{Com}, i, X)$-query of S (for the same and unique $X \in \{0,1\}^{k/3}$), then output $\{X\}$. Otherwise, output $\emptyset$.*

Denote by $\mathcal{B}'$ the oracle that evaluates B''s next-message function. We first remark that $\mathcal{B}'$ can be *efficiently* simulated inside S: $\mathcal{B}'$ running time is (roughly) the same as $\mathcal{B}$'s running time, if we count oracle queries and the final computation of the X as above. Furthermore, by definition, the output of $\mathcal{B}$ and $\mathcal{B}'$ can differ only if

- there are multiple X with $M_i = \mathcal{RO}(\mathsf{Com}, i, X)$ 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 X with $M_i = \mathcal{RO}(\mathsf{Com}, i, X)$ for all $i \in I$.

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

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 X^* 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 X^* 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 X^* unless S explicitly makes a $\mathcal{RO}$-query involving X^*. Since $X^* \in \{0,1\}^{k/3}$ 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 X^*. Consequently, S's view is almost statistically independent of all ${M^*}_i$ with $i \notin I^*$. Hence, S can produce $out_S = \mathbf{M}^*$ only with negligible probability.

It remains to prove that bad occurs only negligibly often.

Lemma 4. *Event* bad *occurs only with negligible probability.*

Proof. We show that any of the events $\mathsf{bad}_{\mathsf{coll}}$, $\mathsf{bad}_{\mathsf{img}}$, $\mathsf{bad}_{\mathsf{bind}}$ 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 $\mathcal{RO}(\mathsf{Com}^j, i, \cdot)$ and $\mathcal{RO}(\mathsf{Com}, i', \cdot)$ are independent as soon as $\mathsf{Com}^j \neq \mathsf{Com}$ or $i \neq i'$. Hence, for all of the events, we can ignore $\mathcal{RO}$- and $\mathcal{B}$-queries with a different Com or i, and assume that $\mathcal{RO}'(\cdot) := \mathcal{RO}(\mathsf{Com}^j, i, \cdot)$ is a fresh random oracle.

$\mathsf{bad}_{\mathsf{coll}}$: Using a birthday bound, we get

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

which implies that there simply exists no ${M^j}_i$ which could raise $\mathsf{bad}_{\mathsf{coll}}$, except with probability at most $2^{-k/3}$,.

$\mathsf{bad}_{\mathsf{img}}$: We show that S's chance to output M_i with $M_i = \mathcal{RO}'(s)$ for some $s \in \{0,1\}^{k/3}$, and such that X 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 Y, outputs the set of all $X \in \{0,1\}^{k/3}$ with $\mathcal{RO}'(X) = Y$. Without loss of generality, we may further assume that S never queries $\mathcal{B}'$ with a Y which has been obtained through an explicit $\mathcal{RO}'(X)$-query. (Namely, unless $\mathsf{bad}_{\mathsf{coll}}$ occurs, which happens only with negligible probability, $\mathcal{B}'$'s answer will then be $\{X\}$.)

Hence, whenever S receives an answer $\neq \emptyset$ from $\mathcal{B}'$, it has already succeeded in producing an M_i with $\mathcal{RO}'(X) = M_i$ for some X, and without querying $\mathcal{RO}'(X)$. So without loss of generality, we can assume that S never queries $\mathcal{B}'$, and hence only produces such an M_i using access to $\mathcal{RO}$ and $\mathcal{X}$ alone. Clearly, $\mathcal{X}$ does not help S, since $\mathcal{X}$ and $\mathcal{RO}$ are independent. But since the set of all Y for which $\mathcal{RO}'(X) = Y$ for some $X \in \{0,1\}^{k/3}$ is sparse in the set of all $Y \in \{0,1\}^k$, and S can only make a polynomial number of $\mathcal{RO}$-queries, S's success in producing such an M_i is negligible.

$\mathsf{bad}_{\mathsf{bind}}$: Let i^I be the (unique) index for which $\mathcal{B}((h_i^j)_{i\in[i^I]})$ outputs I^j. Without loss of generality, assume that S sets I^* after $\mathcal{B}$ first outputs $I^j = \mathcal{B}((h_i^j)_{i\in[i^I]})$. (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 $\mathsf{Com}^j = \mathsf{Com}^*$, since otherwise $\mathsf{bad}_{\mathsf{bind}}$ cannot happen by definition. Hence, S first generates a commit transcript $(h_i^j)_{i\in[i^I]}$, then receives I^j and sends $I^* = I^j$ to the SIM-SO-COM experiment, and only then receives messages ${M^*}_{I^*}$. To achieve $\mathsf{bad}_{\mathsf{bind}}$ in this situation, S must find a full transcript h^j such that ${M^j}_{I^j} = {M^*}_{I^*}$. In particular, there is an $i \in I^j$ such that S opens the i-th commitment in h^j to a value ${M^*}_i$ which S only sees after the transcript of the commit phase is fixed.
Hence, if S achieves $\mathsf{bad}_{\mathsf{bind}}$ with non-negligible probability, we can construct the following PPT machine A. A first simulates S to extract $h = h^j$, and then rewinds S back to the point before it received ${M^*}_{I^*}$. Restarting S with different messages ${M^*}_{I^*}$ then yields a transcript h' that opens the same commitments as in h to different messages. This contradicts Lemma 1.

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 Com^* is not SIM-SO-COM secure. It remains to argue that in the described computational world, $\mathcal{X}$ still satisfies property $\mathcal{P}$.

Lemma 5. *$\mathcal{X}$ 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 $\mathcal{RO}$, A can do without external oracles $\mathcal{RO}$ and $\mathcal{B}$, and use internal simulations of $\mathcal{RO}$ and $\mathcal{B}$ instead. Using lazy sampling for $\mathcal{RO}$, both simulations can even be made PPT. (This includes $\mathcal{B}$'s inversion of $\mathcal{RO}$, since we simulate both $\mathcal{B}$ and $\mathcal{RO}$. We omit the details.)

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), A's advantage $\mathsf{Adv}^{\mathsf{prop}}_{\mathcal{P},\mathcal{X},A}$ must be negligible.

This concludes the proof of Theorem 1.

The following corollary provides an instantiation of Theorem 1 for a number of standard cryptographic primitives.

Corollary 1 (Black-box impossibility of non-interactive or perf. binding SIM-SO-COM). *Let n and $\mathcal{I}$ as in Theorem 1. Then no non-interactive or perfectly binding commitment scheme in the plain model (i.e., without trusted set-up) can be proved simulatable under selective openings via a $\forall\exists$semi-black-box reduction to one or more of the following primitives: one-way functions, one-way permutations, trapdoor one-way permutations, IND-CCA secure public key encryption, homomorphic public key encryption.*

The corollary is a special case of Theorem 1. For instance, to show Corollary 1 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 Corollary 1. This yields impossibility of relativizing proofs for SIM-SO-COM security from one-way permutations. We get impossibility for ∀∃semi-black-box reductions since one-way permutations allow embedding, cf. Simon [44], Reingold et al. [41]. 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.

Of course, it will not be possible to find a real implementation with all the properties of, say, an ideal one-way permutation. For instance, loosely speaking, an ideal one-way permutation contains an exponential amount of entropy. Conversely, the entropy contained in any real one-way permutation is upper bounded by the size of its key. For any efficient implementation of a one-way permutation in the usual sense, this key will be of polynomial size. (A similar argument holds for one-way functions.) Intuitively, this shows the limitations of black-box impossibility results in the style of Corollary 1. In particular, we do not exclude commitment schemes whose security is built on the fact that the used one-way permutation contains only a polynomial amount of entropy. The essence of Corollary 1 hence is: we exclude only constructions that solely build on the, say, one-way property of the employed permutation, but not in any way on how this one-wayness is achieved. We only show that security proofs do not exist which work for any, possibly not efficiently realizable, black box that has the assumed one-way property.

We stress that Corollary 1 makes no assumptions about the nature of the simulation (in the sense of Definition 8). In particular, the simulator may freely use, e.g., the code of the adversary; the only restriction is black-box access to the underlying primitive. As discussed in the introduction, this is quite different from the result one gets upon combining Goldreich and Krawczyk [27] and Dwork et al. [22]: essentially, combining [27, 22] shows impossibility of constructing S in a black-box way from A (i.e., such that S only gets black-box access to A's next-message function).

Finally, we emphasize that our results do not necessarily hold when assuming a model with trusted set-up information. In fact, it is possible to construct non-interactive SIM-SO-COM secure commitment schemes relative to a common reference string. See Section 5 for a detailed explanation.

Generalizations. First, Corollary 1 constitutes merely an example instantiation of the much more general Theorem 1. Second, the proof also holds for a relaxation of SIM-SO-COM security considered by Dwork et al. [22], Definition 7.3, where adversary and simulator approximate a function of the message vector.

3.2 Possibility using non-black-box techniques

Non-black-box techniques vs. interaction. Theorem 1 essentially shows that SIM-SO-COM security cannot be achieved unless one uses non-black-box techniques or interaction. In this section, we will investigate the power of non-black-box techniques to achieve SIM-SO-COM security. As it turns out, for our purposes a concurrently composable zero-knowledge argument system is a suitable non-black-box tool.[7] We stress that the use of this zero-knowledge argument makes our scheme necessarily interactive, and so actually circumvents Theorem 1 in *two* ways: by non-black-box techniques *and* by interaction. However, from a conceptual point of view, our scheme is "non-interactive up to the zero-knowledge argument." In particular, our proof does not use the fact that the zero-knowledge argument is interactive. (That is, if we used a concurrently composable non-interactive zero-knowledge argument in, say, the common reference string model, our proof would still work. See also the discussion in Section 5 on how this relates to our impossibility results.)

The scheme. For our non-black-box scheme, we need an interactive argument system IP with perfect completeness and negligible soundness error, such that IP is zero-knowledge under concurrent composition. We also need a perfectly binding non-interactive commitment scheme $\mathsf{Com}^{\mathsf{b}}$. Both these ingredients can be constructed from one-way permutations. To ease presentation, we only describe a *bit* commitment scheme, which is easily extended (along with the proof) to the multi-bit case. In a nutshell, the sender S^{ZK} commits twice (using $\mathsf{Com}^{\mathsf{b}}$) to the the same bit and proves in zero-knowledge (using IP) that the committed bits are the same.[8] In the opening phase, the sender opens one (randomly selected) commitment. Note that this overall commitment scheme is binding, since IP ensures that both commitments contain the same bits, and the underlying commitment $\mathsf{Com}^{\mathsf{b}}$ is binding. For a SIM-SO-COM simulation, we generate inconsistent overall commitments which can later be opened arbitrarily by choosing which individual $\mathsf{Com}^{\mathsf{b}}$-commitment is opened. We can use the simulator of IP to generate fake consistency proofs for these inconsistent commitments. (Since we consider many concurrent commitment instances in our SIM-SO-COM experiment, we require concurrent composability from IP for that.)

Scheme 1 (Non-black-box commitment scheme ZKCom). Let $\mathsf{Com}^{\mathsf{b}} = (\mathsf{S}^{\mathsf{b}}, \mathsf{R}^{\mathsf{b}})$ be a perfectly binding non-interactive commitment scheme. Let $\mathsf{IP} = (\mathsf{P}, \mathsf{V})$ be an interactive argument system for NP which enjoys perfect completeness, has negligible soundness error, and which is zero-knowledge under concur-

[7] We require concurrent composability since the SIM-SO-COM definition considers multiple, concurrent sessions of the commitment scheme.

[8] We note that a FOCS referee, reviewing an earlier version of this paper without ZKCom, also suggested to employ zero-knowledge to prove consistency of a given commitment. This suggestion was independent of the eprint version of this paper which at that time already contained our scheme ZKCom. Furthermore, a Eurocrypt referee, reviewing a version of the paper with ZKCom, remarked that alternative constructions of a SIM-SO-COM secure commitment scheme are possible. A more generic construction could be along the lines of "commit using a perfectly binding commitment, then prove consistency of commitment or opening using concurrent zero-knowledge."

rent composition. Define $\mathsf{ZKCom} = (\mathsf{S}^{\mathsf{ZK}}, \mathsf{R}^{\mathsf{ZK}})$ for the following machines S^{ZK} and R^{ZK}:

- Commitment to bit b:
  1. S^{ZK} computes $(com^j, dec^j) \leftarrow \mathsf{S}^{\mathsf{b}}(b)$ for $j \in \{0,1\}$, and then sends (com^0, com^1) to R^{ZK}.
  2. S^{ZK} uses IP to prove to R^{ZK} that com^0 and com^1 commit to the same bit.[9]
- Opening:
  1. S^{ZK} uniformly chooses $j \in \{0,1\}$ and sends (j, dec^j) to R^{ZK}.

The security of ZKCom. It is straightforward to prove that ZKCom is a hiding and binding commitment scheme. (We stress, however, that $\mathsf{Com}^{\mathsf{b}}$'s *perfect* binding property is needed to prove that ZKCom is binding; otherwise, the zero-knowledge argument may become meaningless.) More interestingly, we can also show that ZKCom is SIM-SO-COM secure:

Theorem 2 (Non-black-box possibility of SIM-SO-COM). *Fix n and $\mathcal{I}$ as in Definition 8. Then* ZKCom *is simulatable under selective openings in the sense of Definition 8.*

Proof outline. We start with the real SIM-SO-COM experiment with an arbitrary adversary A. As a first step, we substitute the proofs generated during the commitments by *simulated proofs*. Concretely, we hand to A proofs for the consistency of the commitments that are generated by a suitable simulator S^*. By the concurrent zero-knowledge property of IP, such an S^* exists and yields indistinguishable experiment outputs. Note that S^* does not need witnesses to generate valid-looking proofs, but instead uses (possibly rewinding or even non-black-box) access to A. Hence, we can substitute all ZKCom-commitments with inconsistent commitments of the form (com^0, com^1), where com^0 and com^1 are $\mathsf{Com}^{\mathsf{b}}$-commitments to *different* bits. Such a ZKCom-commitment can later be opened arbitrarily. By the computational hiding property of $\mathsf{Com}^{\mathsf{b}}$ (and since we do not need witnesses to generate consistency proofs anymore), this step does not change the output distribution of the experiment significantly. But note that now, the initial generation of the commitments does not need knowledge of the actual messages. In fact, only the messages $\mathbf{M}_I$ of the actually opened commitments need to be known at opening time. Hence, at this point, the modified experiment is a valid simulator in the sense of the ideal SIM-SO-COM experiment. Since the experiment output has only been changed negligibly by our modifications, we have thus constructed a successful simulator in the sense of Definition 8.

A full proof is given now.

Proof (of Theorem 2). Assume arbitrary n, $\mathcal{I}$, $\mathcal{M}$, R, and A as in Definition 8. We proceed in games.

[9] Formally, the corresponding language $\mathcal{L}$ for IP consists of statements $x = (com^0, com^1)$ and witnesses $w = (dec^0, dec^1)$ such that $\mathcal{R}(x, w)$ iff $\mathsf{R}^{\mathsf{b}}(com^0, dec^0) = \mathsf{R}^{\mathsf{b}}(com^1, dec^1) \in \{0,1\}$.

Game 0 is the real SIM-SO-COM experiment $\mathsf{Exp}^{\textsf{sim-so-real}}_{\mathsf{ZKCom},\mathcal{M},A,R}$ for ZKCom. Define the random variable out_0 as the output of the experiment, so that

$$\mathsf{Pr}\left[\mathsf{Exp}^{\textsf{sim-so-real}}_{\mathsf{ZKCom},\mathcal{M},A,R} = 1\right] = \mathsf{Pr}\left[out_0 = 1\right].$$

In **Game** 1, we interpret the first stage of the experiment as a verifier V^* in the sense of Definition 5. To this end, we constructively define random variables $x_{i,k}, w_{i,k}, z_k^D, z_k^{V^*}$ as follows:

1. sample $\mathbf{M} = (M_i)_{i\in[n]} \in \{0,1\}^n$ from $\mathcal{M}$,
2. uniformly and independently choose n bits $j_1, \ldots, j_n$,
3. for all $i \in [n]$ and $j \in \{0,1\}$, compute $({com^j}_i, {dec^j}_i) \leftarrow \mathsf{S}^{\mathsf{b}}(M_i)$,
4. define $x_{i,k} = ({com^0}_i, {com^1}_i)$, $w_{i,k} = ({dec^0}_i, {dec^1}_i)$, $z_k^{V^*} = \epsilon$ and $z_k^D = (\mathbf{M}, (j_i, {dec^{j_i}}_i)_{i\in[n]})$.

Using this notation, the commitment stage of $\mathsf{Exp}^{\textsf{sim-so-real}}_{\mathsf{ZKCom},\mathcal{M},A,R}$ can be expressed as an interaction of n concurrent instances of prover P with a suitable verifier V^* as in Definition 5.[10] Concretely, we define a verifier V^* that, on input $(x_{i,k})_{i\in[n]} = ({com^0}_i, {com^1}_i)_{i\in[n]}$, internally simulates $\mathsf{Exp}^{\textsf{sim-so-real}}_{\mathsf{ZKCom},\mathcal{M},A,R}$ up to the point where A outputs I. The interactive arguments which show that ${com^0}_i$ and ${com^1}_i$ commit to the same bit are performed concurrently with (n instances of) a prover P that gets $x_{i,k} = ({com^0}_i, {com^1}_i)$ and $w_{i,k} = ({dec^0}_i, {dec^1}_i)$ as input. Finally, V^* outputs $out_{V^*} = I$, so that I will be part of the transcript

$$T_{\mathsf{P},V^*} = \langle \mathsf{P}((x_{i,k}, w_{i,k})_{i\in[n]}), V^*((x_{i,k})_{i\in[n]}, z_k^{V^*})\rangle.$$

We outsource the second stage of the attack into a suitable distinguisher D. Concretely, we define D such that, given $z_k^D = (\mathbf{M}, (j_i, {dec^{j_i}}_i)_{i\in[n]})$ and a transcript T_{P,V^*} (which contains $out_{V^*} = I$), D computes $out_A \leftarrow A((j_i, {dec^{j_i}}_i)_{i\in I})$ and outputs $out_1 = R(\mathbf{M}, out_A)$.

This setting is merely a reformulation of $\mathsf{Exp}^{\textsf{sim-so-real}}_{\mathsf{ZKCom},\mathcal{M},A,R}$ as a concurrent zero-knowledge argument, so we have that

$$\mathsf{Pr}\left[out_1 = 1\right] = \mathsf{Pr}\left[out_0 = 1\right].$$

In **Game** 2, we use IP's concurrent zero-knowledge property. That is, Game 1 already specifies a PPT verifier V^* and a PPT distinguisher D, as well as random variables (x, w), z^{V^*}, and z^D, as in Definition 5. Hence our assumption on IP guarantees the existence of a PPT simulator S^* such that $\mathsf{Adv}^{\mathsf{cZK}}_{V^*,S^*,(x,w),D,z^{V^*},z^D}$ is negligible. We substitute V^* (along with all instances of P) from Game 1 with that simulator S^* in Game 2. Note that now, the execution of Game 2 does not require $w_{i,k} = ({dec^0}_i, {dec^1}_i)$ anymore, but instead only *one* opening ${dec^{j_i}}_i$ for each argument session. If we let out_2 denote D's output (on input z_k^D and out_{S^*}) in this setting, we get that

$$\mathsf{Pr}\left[out_1 = 1\right] - \mathsf{Pr}\left[out_2 = 1\right] = \mathsf{Adv}^{\mathsf{cZK}}_{V^*,S^*,(x,w),D,z^{V^*},z^D}$$

[10] Note that Definition 5 trivially implies security for all *distributions* on (x, w), z^{V^*} and z^D. Also recall that Definition 5 models two different auxiliary inputs z^{V^*} (for V^* and S^*) and z^D (for D). We emphasize again that this is without loss of generality, cf. the discussion after Definition 4.

is negligible.

In **Game** 3, we use $\mathsf{Com}^{\mathsf{b}}$'s hiding property. Namely, we now change the generation of the $x_{i,k} = ({com^0}_i, {com^1}_i)$. While we still generate ${com^{j_i}}_i$ as a commitment to M_i, we now define ${com^{1-j_i}}_i$ as a commitment to $1 - M_i$, so that ${com^0}_i$ and ${com^1}_i$ are commitments to *different* bits. Since ${dec^{1-j_i}}_i$ is never used in Game 2, this does not result in a detectable change in D's output. Concretely, we have that

$$\mathsf{Pr}\left[out_3 = 1\right] - \mathsf{Pr}\left[out_2 = 1\right] = \mathsf{Adv}^{\mathsf{hiding}}_{\mathsf{Com}^{\mathsf{b}},A'}$$

for a suitable adversary A' on $\mathsf{Com}^{\mathsf{b}}$'s hiding property, so that $\mathsf{Pr}\left[out_3 = 1\right] - \mathsf{Pr}\left[out_2 = 1\right]$ is negligible.

To construct **Game** 4, observe that in Game 3, distinguisher D only needs the openings ${dec^{j_i}}_i$ for $i \in I$ from its input $z^D_k = (\mathbf{M}, ({dec^{j_i}}_i)_{i\in[n]})$. We can exploit this fact as follows. We now generate the commitments $x_{i,k} = ({com^0}_i, {com^1}_i)$ and openings ${dec^{j_i}}_i$, as well as the $j_i \in \{0, 1\}$ slightly differently. Concretely, for each message bit M_i, we first choose a random bit b_i and compute $({com^0}_i, {dec^0}_i) \leftarrow \mathsf{S}^{\mathsf{b}}(b_i)$ and $({com^1}_i, {dec^1}_i) \leftarrow \mathsf{S}^{\mathsf{b}}(1 - b_i)$. This modification does not change S^*'s view. When D requires an opening ${dec^{j_i}}_i$ (for $i \in I$), we define $j_i = b_i \oplus M_i$, so that ${dec^{j_i}}_i$ opens the "right" message M_i. This does not change the view of S^* or D, so that we have

$$\mathsf{Pr}\left[out_4 = 1\right] = \mathsf{Pr}\left[out_3 = 1\right].$$

The crucial conceptual difference to Game 3 is that now the execution of D requires only knowledge about the message parts $(M_i)_{i\in I}$ selected by S^* and not the full message vector $\mathbf{M}$.

We can now reformulate Game 4 as an ideal SIM-SO-COM experiment. First, we define a simulator S as follows: first, S prepares bits b_i and commitments $({com^0}_i, {com^1}_i)$ as in Game 4 and then runs an internal simulation of S^* on these commitments. Upon obtaining I from S^*, S outputs I. Then, upon input $(M_i)_{i\in I}$, S runs an internal simulation of A on input $(j_i, {dec^{j_i}}_i)_{i\in I}$ for $j_i = b_i \oplus M_i$ as in Game 4. Finally, S outputs $out_S = out_A$. By construction, the ideal SIM-SO-COM experiment $\mathsf{Exp}^{\mathsf{sim\text{-}so\text{-}ideal}}_{\mathcal{M},S,R}$ with this S is only a reformulation of Game 4, so that

$$\mathsf{Pr}\left[\mathsf{Exp}^{\mathsf{sim\text{-}so\text{-}ideal}}_{\mathcal{M},S,R} = 1\right] = \mathsf{Pr}\left[out_4 = 1\right].$$

Putting things together, we get that

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

is negligible, which proves the theorem.

Where is the non-black-box component? Interestingly, the used argument system IP itself can well be black-box zero-knowledge (where black-box zero-knowledge means that the simulator S^* from Definition 5 has only black-box access to the next-message function of V^*). The essential fact that allows us to circumvent our negative result Theorem 1 is the way we employ IP. Namely, ZKCom uses IP to prove a statement about two given commitments (com^0, com^1). This proof (or, rather, argument) uses an explicit and non-black-box description of the employed commitment scheme $\mathsf{Com}^{\mathsf{b}}$. It is this argument that cannot even be expressed when $\mathsf{Com}^{\mathsf{b}}$ makes use of, say, a one-way function given in oracle form.

The role of the commitment randomness. Observe that the opening of a ZKCom-commitment does not release all randomness used for constructing the commitment. In fact, it is easy to see that our proof would not hold if S^{ZK} opened *both* commitments com^0 and com^1 in the opening phase. Hence, ZKCom is not suitable for settings in which an opening corresponds to a corruption of a party (e.g., in a multi-party computation setting), and when one cannot assume no trusted erasures.

Generalizations. First, ZKCom can be straightforwardly extended to a multi-bit commitment scheme, e.g., by running several sessions of ZKCom in parallel. Second, ZKCom is SIM-SO-COM secure also against adversaries with auxiliary input z: our proof holds literally, where of course we also require security of $\mathsf{Com}^{\mathsf{b}}$ against non-uniform adversaries.

4 An indistinguishability-based definition

Motivated by the impossibility result from the previous section, we now relax Definition 8 as follows:

Definition 10 (Indistinguishable under selective openings). *Let $n = n(k) > 0$ be polynomially bounded, 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 $\mathsf{Com} = (\mathsf{S}, \mathsf{R})$ is* indistinguishable under selective openings *(short* IND-SO-COM secure*) iff for every PPT n-message distribution $\mathcal{M}$, and every PPT adversary A, we have that $\mathsf{Adv}^{\mathsf{ind\text{-}so}}_{\mathsf{Com},\mathcal{M},A}$ is negligible. Here*

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

where the experiments $\mathsf{Exp}^{\mathsf{ind\text{-}so\text{-}real}}_{\mathsf{Com},\mathcal{M},A}$ and $\mathsf{Exp}^{\mathsf{ind\text{-}so\text{-}ideal}}_{\mathsf{Com},\mathcal{M},A}$ are defined as follows:

```
Experiment $\mathsf{Exp}^{\mathsf{ind\text{-}so\text{-}real}}_{\mathsf{Com},\mathcal{M},A}$
  $\mathbf{M} = (M_i)_{i\in[n]} \leftarrow \mathcal{M}$
  $I \leftarrow \langle A(\mathtt{recv}), (\mathsf{S}_i(\mathtt{com}, M_i))_{i\in[n]}\rangle$
  $out_A \leftarrow \langle A(\mathtt{open}), (\mathsf{S}_i(\mathtt{open}))_{i\in I}\rangle$

  return $A(\mathsf{guess}, \mathbf{M})$
```

```
Experiment $\mathsf{Exp}^{\mathsf{ind\text{-}so\text{-}ideal}}_{\mathsf{Com},\mathcal{M},A}$
  $\mathbf{M} = (M_i)_{i\in[n]} \leftarrow \mathcal{M}$
  $I \leftarrow \langle A(\mathtt{recv}), (\mathsf{S}_i(\mathtt{com}, M_i))_{i\in[n]}\rangle$
  $out_A \leftarrow \langle A(\mathtt{open}), (\mathsf{S}_i(\mathtt{open}))_{i\in I}\rangle$
  $\mathbf{M}' \leftarrow \mathcal{M} \mid M_I$
  return $A(\mathsf{guess}, \mathbf{M}')$
```

Again, we require from A that $I \in \mathcal{I}_k$, and we denote by $\langle A, (\mathsf{S}_i)_i\rangle$ the output of A after interacting concurrently with instances S_i of S. Furthermore, $\mathcal{M} \mid M_I$ denotes the message distribution $\mathcal{M}$ conditioned on the values of M_I.

On the conditioned distribution $\mathcal{M} \mid M_I$. We stress that, depending on $\mathcal{M}$, it may be computationally hard to sample $\mathbf{M}' \leftarrow \mathcal{M} \mid M_I$, 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 argument systems, see Section 6.

The relation between SIM-SO-COM and IND-SO-COM security. Unfortunately, we (currently) cannot prove that SIM-SO-COM security implies IND-SO-COM security (although this seems plausible, since usually simulation-based definitions imply their indistinguishability-based counterparts). Technically, the reason why we are unable to prove an implication is the conditioned distribution $\mathcal{M} \mid M_I$ in the ideal IND-SO-COM experiment, which cannot be sampled from during an (efficient) reduction.

A relaxation. Alternatively, we could let the adversary predict a predicate π of the whole message vector, and consider him successful if $\mathsf{Pr}\left[b = \pi(\mathbf{M})\right]$ and $\mathsf{Pr}\left[b = \pi(\mathbf{M}')\right]$ for the alternative message vector $\mathbf{M}' \leftarrow \mathcal{M} \mid M_I$ differ non-negligibly. We stress that our upcoming negative result (as well as the application in Section 6) also applies to this relaxed notion.

4.1 Impossibility from black-box reductions

Theorem 3 (Black-box impossibility of perf. binding IND-SO-COM, most general formulation). *Let $n = n(k) = 2k$, and let $\mathcal{I} = (\mathcal{I}_n)_n$ with $\mathcal{I}_n = \{I \subseteq [n] \mid |I| = n/2\}$ denote the set of all $n/2$-sized subsets of $[n]$. Let $\mathcal{X}$ be an oracle that satisfies a property $\mathcal{P}$ even in presence of an EXPSPACE-oracle. We also assume that $\mathcal{X}$ is computable in EXPSPACE.*[11] *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 outline. Similarly to Theorem 1, we specify an oracle $\mathcal{RO}$ which induces a message distribution $\mathcal{M}^*$. This time, however, $\mathcal{RO}$ maps $\mathbb{E}^{n/2+1}$-elements to message vectors in $\mathbb{E}^n$, where $\mathbb{E} = \{0,1\}^k$ is the domain of each individual message. Hence, $n/2$ messages usually do not fix the whole message vector, but more messages do. Now fix any perfectly binding commitment scheme Com^*. We

[11] Examples of such $\mathcal{X}$ are random oracles or ideal ciphers. It will become clearer how we use the EXPSPACE requirement in the proof.

define a breaking oracle $\mathcal{B}$ that, like the $\mathcal{B}$ from Theorem 1, asks for n Com^*-commitments and subsequent openings of a random subset $I \in \mathcal{I}_n$ of these commitments. If all openings are valid, $\mathcal{B}$ extracts the *whole* message vector in the commitments (note that this is possible since Com^* is perfectly binding), and returns a "close" (with respect to Hamming distance) element in the message distribution $\mathcal{M}^*$ if there is a sufficiently close one.

It is easy to see that an adversary can use $\mathcal{B}$ to obtain the whole message vector $\mathbf{M}$ in the real IND-SO-COM experiment. But a message vector freshly sampled from $\mathcal{M}^*$, conditioned on the opened messages M_I, will most likely be different from $\mathbf{M}$. Hence, our adversary easily distinguishes the real from the ideal IND-SO-COM experiment.

The main part of the proof shows that oracle $\mathcal{B}$ is useless to an adversary attacking $\mathcal{X}$'s property $\mathcal{P}$. Assume first that the commitment scheme Com with respect to which an adversary A on $\mathcal{X}$ queries $\mathcal{B}$ is perfectly binding. In that case, a somewhat technical but straightforward combinatorial argument shows that A's successfully opened messages M_I, *together with A's queries to $\mathcal{RO}$,* determine $\mathcal{B}$'s answer (except with small probability). Hence A can use internal simulations of $\mathcal{B}$ and $\mathcal{RO}$ instead of the original oracles, and hence property $\mathcal{P}$ of $\mathcal{X}$ is not damaged by the presence of $\mathcal{B}$. To ensure that $\mathcal{B}$ is only useful for perfectly binding commitment schemes Com, we let $\mathcal{B}$ *test* whether Com is perfectly binding. Since we demand that Com is *perfectly* binding, this test is independent of the random coins used by $\mathcal{X}$. Indeed, $\mathcal{B}$ needs to check that for all syntactically possible commitments and decommitments, and *all* possible random coins used by $\mathcal{X}$, the opened message is unique. Hence, by assumption about $\mathcal{X}$, this test can also be performed by A using an EXPSPACE-oracle, and the above proof idea applies.

A full proof follows.

Proof (of Theorem 3). Let $\mathbb{E} = \{0,1\}^k$ and $\varepsilon := .01$. Let $\mathcal{EXPSPACE}$ be an EXPSPACE-oracle. We stress that $\mathcal{EXPSPACE}$ can be used to perform inefficient computations, but $\mathcal{EXPSPACE}$ itself never makes oracle queries (e.g., to $\mathcal{X}$ or the oracles $\mathcal{RO}$ and $\mathcal{B}$ presented below). Let $\mathcal{RO}$ be a random function from $\mathbb{E}^{n/2+1}$ to $\mathbb{E}^n$. We write $\mathbf{M} \in \mathcal{RO}$ when $\mathbf{M} \in \mathbb{E}^n$ lies in the range of $\mathcal{RO}$. For $\mathbf{M}, \mathbf{M}' \in \mathbb{E}^n$ and $\epsilon > 0$, we write $\mathbf{M} \equiv_\varepsilon \mathbf{M}'$ iff $\mathbf{M}$ and $\mathbf{M}'$ coincide in at least $\lceil (1-\varepsilon)n \rceil$ components (i.e., iff there exists $R \subseteq [n]$, $|R| \geq \lceil (1-\varepsilon)n \rceil$, with $M_R = M'_R$). To construct our last oracle $\mathcal{B}$, let B be the machine that proceeds as follows.

1. Upon receiving Com as input, check that Com describes a perfectly binding (but not necessarily hiding) commitment scheme (see the discussion after the description of $\mathcal{B}$). If not, reject with output $\perp$. If yes, concurrently receive n Com-commitments, indexed by $i \in [n]$.
2. When all commitments are received, output a uniformly chosen $I \in \mathcal{I}$.
3. Engage in $|I|$ concurrent opening phases for the Com-instances with $i \in I$. If all openings are valid (i.e., every Com-receiver instance with $i \in I$ outputs some $M_i \neq \perp$), then extract the whole message vector $\mathbf{M} = (M_i)_{i \in [n]} \in$

$\mathbb{E}^n$ from the commitments (this is possible uniquely since Com is perfectly binding). Output the set of all $\mathbf{M}' \in \mathcal{RO}$ with ${M'}_I = M_I$ and $\mathbf{M}' \equiv_\varepsilon \mathbf{M}$.

We should comment on B's check whether Com is perfectly binding. We want that, for all possible values of $\mathcal{RO}$ and states of $\mathcal{X}$, and for all syntactically allowed commitments, there is at most one message M_i to which a commitment can be opened in the sense of Com. Note that by assumption about $\mathcal{X}$, this condition can be checked using $\mathcal{EXPSPACE}$. Concretely, we let $\mathcal{EXPSPACE}$ iterate internally over all possible internal states of $\mathcal{X}$ and $\mathcal{B}$, and over all possible random tapes of an honest verifier. $\mathcal{EXPSPACE}$ then checks whether a syntactically possible commitment along with two openings to different messages exists. At this point, our requirement on $\mathcal{X}$ should become clear: We require that $\mathcal{X}$ uses only an exponentially large random tape; furthermore, $\mathcal{X}$'s computation should be simulatable by $\mathcal{EXPSPACE}$ for any fixed random tape. We completely ignore whether or not Com is hiding.

Again, we cannot use B directly, since B is stateful, and black-box separations require stateless oracles. So let $\mathcal{B}$ be the oracle that evaluates B's next-message function, suitably padded as in the proof of Theorem 1. We note that, similarly to Lemma 1, we can derive that the perfect binding property of a perfectly binding commitment scheme is preserved by the rewindable formalization in $\mathcal{B}$. In particular, (the transcript of) a commitment phase uniquely determines the only possible opening message.

Lemma 6. *Let* Com^* *be a perfectly binding commitment scheme (that may use all of the described oracles in its algorithms). Then,* Com^* *is not indistinguishable under selective openings.*

Proof. Consider the n-message distribution $\mathcal{M}^*$ that samples random elements in the range of $\mathcal{RO}$. (I.e., $\mathcal{M}^*$ outputs $\mathcal{RO}(X)$ for a uniformly sampled $X \in \mathbb{E}^{n/2+1}$.) Consider the following adversary A that relays between the real or ideal IND-SO-COM experiment and oracle $\mathcal{B}$. (Again, we silently assume that A prefixes queries to $\mathcal{B}$ with the respective message history.)

1. Initially, send Com^* to $\mathcal{B}$.
2. Relay the n commitments from the IND-SO-COM experiment to $\mathcal{B}$.
3. Upon receiving $I^* \in \mathcal{I}$ from $\mathcal{B}$, send I^* to the IND-SO-COM experiment.
4. Upon receiving $|I^*|$ openings from the experiment, relay these openings to $\mathcal{B}$.
5. Upon receiving a challenge message $\mathbf{M}$ from the experiment, and a set $S \subseteq \mathbb{E}^n$ from $\mathcal{B}$, output $out_A = 1$ iff $S = \{\mathbf{M}\}$.

First, we claim that the probability for $S = \{\mathbf{M}^*\}$ is overwhelming, where $\mathbf{M}^*$ denotes the message vector sampled by the IND-SO-COM experiment. By construction of $\mathcal{B}$, we have $\mathbf{M}^* \in S$. Furthermore, for any $\mathbf{M}' \in S$, it must hold that $\mathbf{M}' \equiv_\varepsilon \mathbf{M}^*$. But for any distinct $X^1, X^2 \in \mathbb{E}^{n/2+1}$, we have that $\mathcal{RO}(X^1) \equiv_\varepsilon \mathcal{RO}(X^2)$ with probability $\binom{n}{\lceil(1-\varepsilon)n\rceil}/|\mathbb{E}|^{\lceil(1-\varepsilon)n\rceil}$. A union bound over all $\mathbf{M}' \in \mathcal{RO}$ shows that the probability that there exists an $\mathbf{M}' \in S$, $\mathbf{M}' \neq \mathbf{M}^*$ is negligible. Hence $S = \{\mathbf{M}^*\}$ with overwhelming probability.

Thus, A outputs 1 in the real IND-SO-COM experiment with overwhelming probability, since then $\mathbf{M} = \mathbf{M}^*$. However, in the ideal IND-SO-COM experiment, $\mathbf{M} \neq \mathbf{M}^*$ with overwhelming probability (since for uniformly chosen

$\mathbf{M}^* \in \mathcal{RO}$, the expected number of $\mathbf{M} \in \mathcal{RO}$ with $M_I = M^*{}_I$ is about $|\mathbb{E}| = 2^k$). Consequently, A outputs 1 in the ideal IND-SO-COM experiment only with negligible probability. We get that $\mathsf{Adv}^{\mathsf{ind\text{-}so}}_{\mathsf{Com}^*,\mathcal{M}^*,A}$ is overwhelming, which proves the lemma.

Lemma 7. *$\mathcal{X}$ satisfies $\mathcal{P}$.*

Proof. Consider a PPT adversary A on $\mathcal{X}$'s property $\mathcal{P}$. Note that A may use $\mathcal{RO}$, $\mathcal{B}$, and $\mathcal{EXPSPACE}$ freely. We proceed in games to show that $\mathsf{Adv}^{\mathsf{prop}}_{\mathcal{P},\mathcal{X},A}$ is negligible.

Let **Game** 0 by the original security experiment in which A attacks $\mathcal{X}$'s property $\mathcal{P}$. We say that a $\mathcal{B}$-query is a *commit query* (resp. *open query*) if it finishes the commitment (resp. opening) phase in the corresponding interaction with B, such that $\mathcal{B}$ responds with an $I \in \mathcal{I}$ (resp. a set of $\mathbf{M}' \in \mathcal{RO}$). In the following, we count a recursive $\mathcal{B}$-query made by $\mathcal{B}$ in the role of R as made by A. Without loss of generality, we may assume that A always makes precisely $p(k)$ open queries for a fixed polynomial p, and never makes a query twice. We also assume that for any of A's open queries, A made a corresponding commit query first.[12] Let out_0 denote $\mathcal{P}$'s output in Game 0. By definition, we have

$$\mathsf{Pr}\left[out_0 = 1\right] - 1/2 = \mathsf{Adv}^{\mathsf{prop}}_{\mathcal{P},\mathcal{X},A}.$$

In **Game i (for $0 < i \leq p(k)$)**, we use an oracle $\mathcal{B}_i$ instead of oracle $\mathcal{B}$. Here, $\mathcal{B}_i$ behaves like $\mathcal{B}$, except that $\mathcal{B}_i$ answers each of A's first i opening queries as follows. Here, M_I denotes the opened messages, as before.

- If all openings are valid, then return the set of all $\mathbf{M}' \in \mathcal{RO}$ which have been explicitly obtained through $\mathcal{RO}$-queries by A (or $\mathcal{B}_i$, in the role of a receiver), and for which $M'{}_I = M_I$.

We stress that oracle $\mathcal{B}_i$ does not break a commitment or use internal access to $\mathcal{RO}$ until the $(i+1)$-th open query. Let out_i denote $\mathcal{P}$'s output in Game i. To show that out_i is not significantly affected by our changes, fix an i. Let h denote A's i-th open query in Game i. Let $S = \mathcal{B}_i(h)$ denote the answer A gets in Game i, and let $S' = \mathcal{B}_{i-1}(h)$ denote the answer that A would have received in Game $i-1$. We show in Lemma 8 below that $S = S'$ except with probability asymptotically smaller than $2^{-3\varepsilon k}$, so that

$$\mathsf{Pr}\left[out_i = 1\right] - \mathsf{Pr}\left[out_{i-1} = 1\right] \leq 2^{-(\varepsilon/2)k}$$

for sufficiently large k and all $i \in [p(k)]$.

Observe that in Game $p(k)$, $\mathcal{B}_{p(k)}$ and $\mathcal{RO}$ can both be simulated efficiently inside A. Indeed, $\mathcal{B}_{p(k)}$ only needs knowledge about A's $\mathcal{RO}$-queries, as well as access to $\mathcal{EXPSPACE}$ to check whether a given commitment scheme is perfectly binding. Hence,

$$\mathsf{Adv}^{\mathsf{prop}}_{\mathcal{P},\mathcal{X},A'} = \mathsf{Pr}\left[out_{p(k)} = 1\right] - 1/2$$

[12] In order to violate this assumption, A would have to guess an $I \in \mathcal{I}$ as chosen by $\mathcal{B}$ upon the corresponding commit query. Since $|\mathcal{I}|$ is large, we ignore this possibility.

for a suitable PPT adversary A' that internally simulates A, $\mathcal{RO}$, and $\mathcal{B}_{p(k)}$, and only needs access to $\mathcal{EXPSPACE}$. By assumption about $\mathcal{X}$, $\mathsf{Adv}^{\mathsf{prop}}_{\mathcal{P},\mathcal{X},A'}$ is negligible, and hence so must be $\mathsf{Adv}^{\mathsf{prop}}_{\mathcal{P},\mathcal{X},A}$.

It remains to prove that, in the situation of Lemma 7, $S = S'$ with high probability.

Lemma 8. *In the situation of Lemma 7,* $\Pr[S \neq S'] \leq 2^{-(\varepsilon/2)k}$ *for sufficiently large* k*.*

Combining Lemmas 9, 10, 11, and 12 below shows Lemma 8.

Lemma 9. *In the situation of Lemma 7,* $|S| \leq 1$ *except with probability at most* $q(k)2^{-k}$ *for some polynomial* q*.*

Proof. We interpret the whole Game i (including A, $\mathcal{P}$, $\mathcal{X}$, $\mathcal{B}_i$, and $\mathcal{EXPSPACE}$) as a machine A' interacting with $\mathcal{RO}$. Note that A' may be computationally unbounded, but only makes a polynomial number of $\mathcal{RO}$-queries, at least until A's i-th open query. Let $Q_{\mathcal{RO}}$ denote the set of $\mathcal{RO}$-queries of A'. Now $|S| > 1$ implies that there are $X^1, X^2 \in Q_{\mathcal{RO}}$ with $X^1 \neq X^2$, such that $\mathcal{RO}(X^1), \mathcal{RO}(X^2) \in S$, and so $\mathcal{RO}(X^1)_I = \mathcal{RO}(X^2)_I$. However, the statistical properties of $\mathcal{RO}$ imply that for any $X^1, X^2 \in Q_{\mathcal{RO}}$, $\mathcal{RO}(X^1)$ and $\mathcal{RO}(X^2)$ match in at least one component with probability at most $n2^{-k}$. A union bound over all such pairs shows the claim.

Lemma 10. *In the situation of Lemma 7,* $|S'| \leq 1$ *except with probability at most* $q(k)2^{-k}$ *for some polynomial* q*.*

Proof. As in Lemma 9, we interpret Game i as a machine A' interacting with $\mathcal{RO}$. Again, let $Q_{\mathcal{RO}}$ denote the set of $\mathcal{RO}$-queries of A'. Now let $\mathbb{X}$ be the set of all $X \in \mathbb{E}^{n/2+1} \setminus Q_{\mathcal{RO}}$ with $\mathcal{RO}(X)_I = M_I$. Using, e.g., Chebyshev's inequality, we get $|\mathbb{X}| < 2|\mathbb{E}|$, except with probability at most 2^{-k}. Furthermore, $Q_{\mathcal{RO}}$ contains at most one query X with $\mathcal{RO}(X)_I = M_I$ except with probability at most $q_1(k)2^{-k}$ for some polynomial q_1 (with similar reasoning as in Lemma 9). Let $\mathbb{X}' := \mathbb{X} \cup \{X\}$ for that $X \in Q_{\mathcal{RO}}$, or $\mathbb{X}' := \mathbb{X}$ if no such X exists. By the preceding discussion, $|\mathbb{X}'| \leq 2\mathbb{E}$ except with probability $(q_1(k)+1)2^{-k}$.

Now $|S'| > 1$ implies that $X^1, X^2 \in \mathbb{X}'$ exist, such that $X^1 \neq X^2$ but $\mathcal{RO}(X^1) \equiv_\varepsilon \mathbf{M} \equiv_\varepsilon \mathcal{RO}(X^2)$, so $\mathcal{RO}(X^1) \equiv_{2\varepsilon} \mathcal{RO}(X^2)$. Observe that the values $\mathcal{RO}(X)$ for $X \in \mathbb{X}'$ are independent, conditioned only on $\mathcal{RO}(X)_I = M_I$. For any fixed $X^1, X^2 \in \mathbb{X}'$ with $X^1 \neq X^2$, the probability that $\mathcal{RO}(X^1) \equiv_{2\varepsilon} \mathcal{RO}(X^2)$ is $\binom{n/2}{\lceil (1/2-2\varepsilon)n \rceil}/|\mathbb{E}|^{\lceil (1/2-2\varepsilon)n \rceil}$, which is less than 2^{-3k-2} for sufficiently large k. Assuming that $|\mathbb{X}'| \leq 2|\mathbb{E}| = 2^{k+1}$, a union bound yields that no such X^1, X^2 exist, and hence $|S'| \leq 1$, except with probability 2^{-k}. Summing up shows the claim.

Lemma 11. *In the situation of Lemma 7, we have* $S = \emptyset$ *and* $|S'| = 1$ *simultaneously with probability at most* $q(k)2^{-k/2}$ *for some polynomial* q*.*

Proof. Let bad denote the event that $S = \emptyset$ but $S' = \{\mathbf{M}'\}$ for some $\mathbf{M}'$, and let bad_j denote the event that bad occurs and A's i-th open query refers to A's j-th commit query. Since A makes only polynomially many $\mathcal{B}_i$-queries, there is a polynomial $q_1 = q_1(k)$ and a function $j = j(k)$ such that $\mathsf{Pr}\left[\mathsf{bad}_j\right] \geq \mathsf{Pr}\left[\mathsf{bad}\right]/q_1(k)$.

Consider the machine A' that simulates Game i and interacts externally only with oracle $\mathcal{RO}$. Call $I^1 \in \mathcal{I}$ the answer of $\mathcal{B}_i$ to A's j-th commit query. After A submits its i-th open query, A' rewinds the simulation back to A's j-th commit query, and then restarts with a freshly sampled $I^2 \in \mathcal{I}$ as $\mathcal{B}_i$'s answer to A's j-th commit query. By $\mathsf{bad}_{j,1}$, resp. $\mathsf{bad}_{j,2}$, we denote the events that bad_j occurs before, resp. after the rewinding. It is clear that $\mathsf{Pr}\left[\mathsf{bad}_{j,1}\right] = \mathsf{Pr}\left[\mathsf{bad}_{j,2}\right] = \mathsf{Pr}\left[\mathsf{bad}_j\right]$, but unfortunately, the events $\mathsf{bad}_{j,1}$ and $\mathsf{bad}_{j,2}$ may be dependent. We have to work to establish that $\mathsf{bad}_{j,1}$ and $\mathsf{bad}_{j,2}$ occur simultaneously with sufficiently large probability. Consider a prefix E_j of A''s execution until A's j-th commit query. Given any such E_j and a fixed oracle $\mathcal{RO}$, the events $\mathsf{bad}_{j,1}$ and $\mathsf{bad}_{j,2}$ are independent and occur with the same probability, so that

$$\begin{aligned}\mathsf{Pr}\left[\mathsf{bad}_{j,1} \wedge \mathsf{bad}_{j,2}\right] &= \sum_{E_j,\mathcal{RO}} \mathsf{Pr}\left[\mathsf{bad}_{j,1} \wedge \mathsf{bad}_{j,2} \mid E_j, \mathcal{RO}\right] \cdot \mathsf{Pr}\left[E_j, \mathcal{RO}\right] \\ &= \sum_{E_j,\mathcal{RO}} \mathsf{Pr}\left[\mathsf{bad}_{j,1} \mid E_j, \mathcal{RO}\right]^2 \cdot \mathsf{Pr}\left[E_j, \mathcal{RO}\right] \\ &\overset{(*)}{\geq} \left(\sum_{E_j,\mathcal{RO}} \mathsf{Pr}\left[\mathsf{bad}_{j,1} \mid E_j, \mathcal{RO}\right] \cdot \mathsf{Pr}\left[E_j, \mathcal{RO}\right]\right)^2 \\ &= \mathsf{Pr}\left[\mathsf{bad}_{j,1}\right]^2 = \mathsf{Pr}\left[\mathsf{bad}_j\right]^2 \geq \mathsf{Pr}\left[\mathsf{bad}\right]^2/q_1(k)^2,\end{aligned}$$

where $(*)$ uses that $\sum_i c_i x_i^2 \geq (\sum_i c_i x_i)^2$ for $c_i, x_i \geq 0$ with $\sum_i c_i = 1$ by Jensen's inequality.

Let $Q_{\mathcal{RO},1}$ denote the set of A''s $\mathcal{RO}$-queries before the rewinding, and let $Q_{\mathcal{RO},2}$ denote the set of A''s $\mathcal{RO}$-queries after the rewinding and before A's j-th commit query. The rationale here is that $Q_{\mathcal{RO},1}$ are A's queries in the run related to I^1, and $Q_{\mathcal{RO},2}$ are A's queries in the run related to I^2. Note that $Q_{\mathcal{RO},1}$ and $Q_{\mathcal{RO},2}$ share A's queries before the j-th commitment. We write $\mathcal{RO}(Q_{\mathcal{RO},i})$ for the set of all $\mathcal{RO}(X)$ for $X \in Q_{\mathcal{RO},i}$.

Now $\mathsf{bad}_{j,1} \wedge \mathsf{bad}_{j,2}$ implies that A opens two subsets M_{I^1} and M_{I^2} message vector $\mathbf{M}$ inside the j-th commit query, such that there exist $\mathbf{M}^1, \mathbf{M}^2 \in \mathcal{RO}$ with the following properties:

- ${M^1}_{I^1} = M_{I^1}$ and ${M^2}_{I^2} = M_{I^2}$,
- $\mathbf{M}^1 \equiv_\varepsilon \mathbf{M} \equiv_\varepsilon \mathbf{M}^2$ and hence $\mathbf{M}^1 \equiv_{2\varepsilon} \mathbf{M}^2$,
- $\mathbf{M}^1 \notin \mathcal{RO}(Q_{\mathcal{RO},1})$ and $\mathbf{M}^2 \notin \mathcal{RO}(Q_{\mathcal{RO},2})$.

We claim that $\mathbf{M}^1 = \mathbf{M}^2$ with high probability. To see this, let $\mathbb{M}$ be set of all $\mathbf{M}' \in \mathcal{RO} \setminus \mathcal{RO}(Q_{\mathcal{RO},1})$ which satisfy ${M'}_{I^1 \cap I^2} = M_{I^1 \cap I^2}$. A simple calculation shows that $m := |I^1 \cap I^2| \geq n/10$ except with probability at most 2^{-k} for sufficiently large k. Now $|\mathbb{M}|$'s expected value is, depending on $|Q_{\mathcal{RO},1}|$, at most $|\mathbb{E}|^{n/2+1-m}$. A Chebyshev bound as in Lemma 10 yields that $|\mathbb{M}| \leq |\mathbb{E}|^{n/2-m+2}$

except with probability at most $q_2(k)2^{-k}$ for some polynomial q_2. So assume $|I^1 \cap I^2| \geq n/10$ and $|\mathbb{M}| \leq |\mathbb{E}|^{n/2-m+2}$. Then, for any two $\mathbf{M}^1, \mathbf{M}^2 \in \mathbb{M}$ with $\mathbf{M}^1 \neq \mathbf{M}^2$, we have $\mathbf{M}^1 \equiv_{2\varepsilon} \mathbf{M}^2$ with probability at most $\binom{n-m}{\lfloor 2\varepsilon n \rfloor}/|\mathbb{E}|^{n-m-\lfloor 2\varepsilon n \rfloor}$. A simple calculation and a union bound over all $\mathbf{M}^1, \mathbf{M}^2 \in \mathbb{M}$ yield that there do not exist $\mathbf{M}^1, \mathbf{M}^2 \in \mathbb{M}$ with $\mathbf{M}^1 \equiv_{2\varepsilon} \mathbf{M}^2$ yet $\mathbf{M}^1 \neq \mathbf{M}^2$, except with probability at most $q_3(k)2^{-k}$ for some polynomial q_3. So for the $\mathbf{M}^1, \mathbf{M}^2$ guaranteed by $\mathsf{bad}_{j,1} \wedge \mathsf{bad}_{j,2}$, either $\mathbf{M}^1 = \mathbf{M}^2$, or $\mathbf{M}^2 \notin \mathbb{M}$ with high probability.

Now $\mathbf{M}^2 \notin \mathbb{M}$ implies $\mathbf{M}^2 = \mathcal{RO}(X)$ for some $X \in Q_{\mathcal{RO},1}$, and $\mathsf{bad}_{j,2}$ even dictates $X \in Q_{\mathcal{RO},1} \setminus Q_{\mathcal{RO},2}$. Put differently, $\mathbf{M}^2 \notin \mathbb{M}$ implies that in the execution after the rewinding, $M_{I^2} = {M^2}_{I^2}$ contains a component of an $\mathcal{RO}$-image $\mathbf{M}^2$ obtained (independently, since $\mathbf{M}^2 \notin Q_{\mathcal{RO},2}$) before the rewinding. By symmetry, the probability that this happens equals the probability that M_{I^1} contains a component of an $\mathcal{RO}$-image $\mathbf{M}^1$ queried after the rewinding. However, this essentially means that A' has guessed a component of the result of an *upcoming* $\mathcal{RO}$-query, which can happen with probability at most $q_4(k)2^{-k}$ for some polynomial q_4 by the statistical properties of $\mathcal{RO}$. We conclude that hence, $\mathbf{M}^2 \in \mathbb{M}$ and so $\mathbf{M}^1 = \mathbf{M}^2$ except with probability at most $q_5(k)2^{-k}$ for a polynomial q_5.

Finally, a counting argument shows that $|I^1 \cup I^2| < n/2 + 2$ happens with probability less than 2^{-k} for large enough k. Summarizing, $\mathsf{bad}_{\mathsf{glue}} := \mathsf{bad}_{j,1} \wedge \mathsf{bad}_{j,2} \wedge (\mathbf{M}^1 = \mathbf{M}^2) \wedge (|I^1 \cup I^2| \geq n/2 + 2)$ happens with probability at least $\mathsf{Pr}\left[\mathsf{bad}\right]^2 - q_6(k)2^{-k}$ for some polynomial q_6. But $\mathsf{bad}_{\mathsf{glue}}$ implies that A' has found $J := I^1 \cup I^2$ with $|J| \geq n/2+2$, such that there exists an $\mathbf{M}' := \mathbf{M}^1 = \mathbf{M}^2 \in \mathcal{RO}$ with ${M'}_J = M_J$, and A' has not obtained $\mathbf{M}'$ through an explicit $\mathcal{RO}$-query. Another Chebyshev bound shows that no such $\mathbf{M}'$ *exists*, except with probability (over the images $\mathcal{RO} \setminus \mathcal{RO}(Q_{\mathcal{RO},1} \cup Q_{\mathcal{RO},2})$ not queried by A') at most 2^{-k}. Hence, $\mathsf{Pr}\left[\mathsf{bad}_{\mathsf{glue}}\right] \leq 2^{-k}$, so that we finally have $\mathsf{Pr}\left[\mathsf{bad}\right] \leq q(k)2^{k/2}$ for some polynomial q.

Lemma 12. *In the situation of Lemma 7, we have $|S| = 1$ and $S' = \emptyset$ simultaneously with probability at most $2^{-(\varepsilon/2)k}$ for large enough k.*

Proof. Again, we interpret the whole Game i (except for $\mathcal{RO}$) as a machine A' interacting with $\mathcal{RO}$. As in Lemma 11, A' waits for A's i-th open query M_I, and then rewinds the whole game back to A's j-th commit query. Again, A' resamples an $I \leftarrow \mathcal{I}$ as a fresh answer to A's j-th commit query, in the hope that A opens M_I in the i-th open query. However, this time A' repeats this process $p(k)$ times for a suitable number $p(k)$ to be determined later. Let S^ℓ and I^ℓ denote the values of I and S from the ℓ-th rewinding.

Now fix random tapes for all machines simulated inside A', and fix an $\mathcal{RO}$. This means that the only randomness during the execution of A' comes from the choice of the I^ℓ. Let bad denote the event that $|S| = 1$ but $S' = \emptyset$, and let bad_j denote the event that bad occurs and A's i-th open query refers to A's j-th commit query. Since A makes only polynomially many $\mathcal{B}_i$-queries, there is a polynomial $q = q(k)$ and a function $j = j(k)$ such that $\mathsf{Pr}\left[\mathsf{bad}_j\right] \geq \mathsf{Pr}\left[\mathsf{bad}\right]/q(k)$, where the probability is only over $I \in \mathcal{I}$.

Suppose that $\mathsf{Pr}[\mathsf{bad}] > 2^{-(\varepsilon/2)k}$ for contradiction, so that $\mathsf{Pr}[\mathsf{bad}_j] > 2^{-\varepsilon k}$ for large enough k. Let $\mathcal{I}' \subseteq \mathcal{I}$ be the set of all I such that bad_j occurs when A receives I upon the j-th commit query. Note that $\mathcal{I}'$ is well-defined, since we fixed all randomness except for I. Assume first that there exists a subset $B \subseteq [n]$ of size $|B| > \lfloor \varepsilon n \rfloor$ with $\mathsf{Pr}[I \in \mathcal{I}' \wedge i \in I] < 2^{-2\varepsilon k}$ for all $i \in B$, where the probability is over $I \in \mathcal{I}$. We have $\mathsf{Pr}[I \cap B = \emptyset] = \binom{\lceil (1-\varepsilon)n \rceil}{n/2} / \binom{n}{n/2} \leq 2^{-\varepsilon n} = 2^{-2\varepsilon k}$, so

$$\begin{aligned} 2^{-\varepsilon k} - 2^{-2\varepsilon k} \leq \mathsf{Pr}[I \in \mathcal{I}'] - \mathsf{Pr}[I \cap B = \emptyset] &\leq \mathsf{Pr}[I \in \mathcal{I}' \wedge I \cap B \neq \emptyset] \\ &\leq \sum_{i \in B} \mathsf{Pr}[I \in \mathcal{I}' \wedge i \in I] < n \cdot 2^{-2\varepsilon k} \end{aligned}$$

creates a contradiction for sufficiently large k. Hence, no such B exists, so there must be a subset $R \subseteq [n]$ of size $|R| \geq \lceil (1-\varepsilon)n \rceil$ such that $\mathsf{Pr}[I \in \mathcal{I}' \wedge i \in I] \geq 2^{-2\varepsilon k}$ for all $i \in R$.

Our goal is now to use A' to extract M_R with high probability. To this end, we first finish our description of A'. Let L denote the set of all $\ell \in [p(k)]$ for which bad_j occurs in the ℓ-th rewinding. After $p(k) := 2^{8\varepsilon k}$ rewindings, A' outputs M_J, where $J = \bigcup_{\ell \in L} I^\ell$ is the union of all successfully extracted partial message subsets. For $\ell \in L$, we have $|S^\ell| = 1$ by definition of bad_j, so say $S^\ell = \{\mathbf{M}^\ell\}$. By definition, $\mathbf{M}^\ell$ has been obtained by A' through an explicit $\mathcal{RO}$-query, and we have ${M^\ell}_{I^\ell} = M_{I^\ell}$ for the message vector $\mathbf{M}$ inside A's j-th commit query. Similar to Lemma 9, all components of all $\mathcal{RO}$-images obtained by A' are pairwise distinct, except with probability at most $2^{-k/2}$ for large enough k. As in Lemma 11, we can show that all the $\mathcal{RO}$-images $\mathbf{M}^\ell$ are identical, except with probability $2^{-k/2}$ for sufficiently large k. Thus, there exists one single $\mathbf{M}' \in \mathcal{RO}$ with ${M'}_J = M_J$. Now note that the I^ℓ are independent. Hence, a Chebyshev bound shows that for each fixed $i \in R$, there is an $I^\ell \in L \subseteq \mathcal{I}'$ with $i \in I^\ell$, except with probability at most $2^{-6\varepsilon k}$. A union bound over all $i \in R$ yields $R \subseteq J$ except with probability at most $2^{-5\varepsilon k}$ for large enough k. So, except with probability $2^{-6\varepsilon k} + 2^{k/2} < \mathsf{Pr}[\mathsf{bad}]$, A' shows the existence of an $\mathbf{M}' \in \mathcal{RO}$ with ${M'}_J = M_J$ for $|J| \geq \lceil (1-\varepsilon)n \rceil$, such that $\mathbf{M}' \equiv_\varepsilon \mathbf{M}$. Since ${M'}_{I^\ell} = M_{I^\ell}$ for any $I^\ell \in L$, this contradicts bad_j and thus bad. Hence, our assumption on $\mathsf{Pr}[\mathsf{bad}]$ must have been incorrect, and we have proved the lemma.

Combining Lemma 6 and Lemma 7 shows Theorem 3.

On the requirement on $\mathcal{X}$. We require that $\mathcal{X}$ is secure even in the presence of an EXPSPACE-oracle, but can be simulated by an EXPSPACE-oracle as well. This means that in any polynomial interaction with black-box access to $\mathcal{X}$, it is computationally infeasible to break $\mathcal{X}$, even with exponential computing powers. At the same time, $\mathcal{X}$ itself should be implementable in exponential time. We stress that this requirement on $\mathcal{X}$ is a rather mild one. For instance, random oracles are one-way even against computationally unbounded adversaries, as long as the adversary makes only a polynomial number of oracle queries. Hence, an EXPSPACE-oracle (which itself does not perform oracle queries) is not helpful in breaking a random oracle. This even holds when the random oracle uses lazy sampling (and hence is polynomial-time in any polynomial context), or when the random oracle is actually only a t-wise independent hash function for slightly superpolynomial t (say, $t(k) = k \log k$). It is also noteworthy that the choice of EXPSPACE was rather arbitrary; any oracle that is sufficient to simulate $\mathcal{X}$, and at the same time not sufficient to break $\mathcal{X}$ will do in place of $\mathcal{EXPSPACE}$.

So similarly to Corollary 1, we get for concrete choices of $\mathcal{X}$ and $\mathcal{P}$:

Corollary 2 (Black-box impossibility of perf. binding IND-SO-COM). *Let n and $\mathcal{I}$ as in Theorem 3. Then no perfectly binding commitment scheme in the plain model (i.e., without trusted set-up) can be proved indistinguishable under selective openings via a $\forall\exists$semi-black-box reduction to one or more of the following primitives: one-way functions, one-way permutations, trapdoor one-way permutations, IND-CCA secure public key encryption, homomorphic public key encryption.*

Generalizations. Again, Corollary 2 constitutes merely an example instantiation of the much more general Theorem 3. We stress, however, that the proof for Theorem 3 does *not* apply to "almost-perfectly binding" commitment schemes such as the one from Naor [37]. (For instance, for such schemes, $\mathcal{B}$'s check that the supplied commitment scheme is binding might tell something about $\mathcal{X}$.)

4.2 Statistically hiding schemes are secure

Fortunately, things look different for statistically hiding commitment schemes:

Theorem 4 (Statistically hiding schemes are IND-SO-COM secure). *Fix arbitrary n and $\mathcal{I}$ as in Definition 10, and let* $\mathsf{Com} = (\mathsf{S}, \mathsf{R})$ *be a statistically hiding commitment scheme. Then* Com *is indistinguishable under selective openings in the sense of Definition 10.*

Proof outline. Intuitively, the claim holds since an adversary A's views in the real, resp. ideal IND-SO-COM experiment are statistically close (and hence so must be A's outputs). However, the fact that A's views are indeed statistically close is less obvious than it may seem at first glance. Our proof proceeds in games and starts with the real IND-SO-COM experiment with A. As a first modification, we change the opening phase of the experiment, so that the opening of each selected commitment is produced solely from the commitment itself and the "target message" M_i to which it should be opened (but not from opening information previously generated alongside the commitment). Note that this change is merely conceptual and does not alter A's view at all. This makes the opening phase inefficient, but since we are dealing with statistically hiding commitment schemes, we need not worry about that. Indeed, by the statistical hiding property, we can now substitute all commitments (in a hybrid argument) with commitments to a fixed value (say, 0^k) without affecting the experiment output. We can reduce this step to the hiding property of the commitment scheme since the experiment only needs commitments as input, and produces all openings on its own. At this point, all commitments that A gets are independent of $\mathbf{M}$, and so the whole view of A is independent of the unopened values $M_{[n]\setminus I}$. Hence A's output is (almost) independent of $M_{[n]\setminus I}$ in the real IND-SO-COM experiment and, with similar reasoning, also in the ideal IND-SO-COM experiment. This shows the claim.

We proceed to the full proof.

Proof (of Theorem 4). Fix an n-message distribution $\mathcal{M}$ and a PPT adversary A on the SIM-SO-COM security of Com. We proceed in games.

Game -1 is the real IND-SO-COM experiment $\mathsf{Exp}^{\mathsf{ind\text{-}so\text{-}real}}_{\mathsf{Com},\mathcal{M},A}$. Let out_{-1} denote the output of the experiment, so that we have

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

Game 0 constitutes our first modification of $\mathsf{Exp}^{\mathsf{ind\text{-}so\text{-}real}}_{\mathsf{Com},\mathcal{M},A}$, and proceeds as follows (*emphasized* steps are different from $\mathsf{Exp}^{\mathsf{ind\text{-}so\text{-}real}}_{\mathsf{Com},\mathcal{M},A}$):

1. sample messages $\mathbf{M} = (M_i)_{i\in[n]} \leftarrow \mathcal{M}$,
2. let $A(\mathtt{recv})$ interact with n concurrent instances $(\mathsf{S}_i(\mathtt{com}, M_i))_{i\in[n]}$ of S,
3. let $I \in \mathcal{I}$ be A's output after interacting with the S_i,
4. *for $i \in I$, set S_i's state to the output of procedure $\mathtt{AltDec}(H_i, M_i)$ (described below), where H_i denotes the exchanged messages during the commit phase of the i-th* Com *instance*,
5. let $A(\mathtt{open})$ interact concurrently with the $|I|$ instances $(\mathsf{S}_i(\mathtt{open}))_{i\in I}$ of S,
6. send the full message vector $\mathbf{M}$ to A,
7. output A's final output b.

The (in general inefficient) procedure $\mathtt{AltDec}$ takes as input a history H_i of exchanged messages in the commit phase and a message M_i. We call a random tape t for S *consistent with H_i and M_i* iff $\mathsf{S}(\mathtt{com}, M_i)$ (with random tape t) produces the sender's messages in H_i when receiving the respective receiver's replies in H_i. Let T_{H_i,M_i} denote the set of all random tapes t for S which are consistent with H_i and M_i. Now $\mathtt{AltDec}(H_i, M_i)$ samples uniformly a random tape t from T_{H_i,M_i} and returns the state of S with random tape t and after an interaction according to H_i. If $T_{H_i,M_i} = \emptyset$, then $\mathtt{AltDec}$ returns $\perp$ (and Game 0 aborts with output 0). In other words, $\mathtt{AltDec}$ returns the state of a sender S with initial input M_i, conditioned on the transcript T_i of the commit phase.

In Game 0, $\mathtt{AltDec}$ will never return $\perp$ (since $\mathtt{AltDec}$ is invoked with a transcript H_i that has actually been produced as a commit phase to M_i). Moreover, the view of the adversary is not altered by re-sampling the internal state of the sender, conditioned on all previous actions, as $\mathtt{AltDec}$ does. Hence, we have

$$\mathsf{Pr}\left[out_0 = 1\right] = \mathsf{Pr}\left[out_{-1} = 1\right]$$

for the output out_0 of the experiment in Game 0.

We describe Game j (for $j \in [n]$). Game j is identical to Game 0, except for step 2:

2*. let $A(\mathtt{recv})$ interact with n concurrent instances $(\mathsf{S}_i(\mathtt{com}, {M^*}_i))_{i\in[n]}$ of S, where we set ${M^*}_i = 0^k$ for $i \leq j$ and ${M^*}_i = M_i$ for $j > i$, Obviously, for $j = 0$ we would get Game 0. Note that only difference between Game $j-1$ and Game j is the commitment to M_j. In fact, we can now construct an adversary A' on Com's statistical hiding property. A' first uniformly chooses $j \in [n]$, then simulates Game $j-1$, but picks M_j and 0^k as challenge messages for its own experiment $\mathsf{Exp}^{\mathsf{hiding}\text{-}b}_{\mathsf{Com},A'}$. The j-th commitment (to either M_j or 0^k) is performed through the experiment. $\mathsf{Exp}^{\mathsf{hiding}\text{-}0}_{\mathsf{Com},A'}$ is then a perfect simulation of Game $j-1$, and $\mathsf{Exp}^{\mathsf{hiding}\text{-}1}_{\mathsf{Com},A'}$ perfectly simulates Game j. (However, we stress that A' is inherently unbounded: A' must run procedure $\mathtt{AltDec}$.) We get that

$$\mathsf{Pr}\left[out_n = 1\right] - \mathsf{Pr}\left[out_0 = 1\right] = n \cdot \mathsf{Adv}^{\mathsf{hiding}}_{\mathsf{Com},A'}$$

must be negligible, which proves that

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

is negligible.

We can apply the same reasoning for the ideal IND-SO-COM experiment $\mathsf{Exp}^{\mathsf{ind\text{-}so\text{-}real}}_{\mathsf{Com},\mathcal{M},A}$: we first construct the openings using the commit transcripts H_i and the target messages M_i alone as in Game 0 above. Then we change the actual commitments to commitments to 0^k, as in Game 1 up to Game n above. At this point, the modified ideal experiment first samples $\mathbf{M} \leftarrow \mathcal{M}$ and then $\mathbf{M'} \leftarrow \mathcal{M} \mid M_I$, but *never uses* $\mathbf{M}$. Hence we can sample $\mathbf{M'} \leftarrow \mathcal{M}$ in the first place without changing A's view. But this is then exactly Game n from above, so that we get that

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

is negligible. Hence $\mathsf{Adv}^{\mathsf{ind\text{-}so}}_{\mathsf{Com},\mathcal{M},A}$ is negligible as well, which shows the theorem.

We stress that the proof of Theorem 4 also holds (literally) in case A and/or $\mathcal{M}$ gets an additional auxiliary input z.

5 Application to adaptively secure encryption

Motivation and setting. Taking up the motivation of Damgård [17], we consider the setting of an adversary A that may corrupt, in an adaptive manner, a subset of a set of parties $P_1, \ldots, P_n$. 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 (non-interactive, hiding and binding) commitment scheme $\mathsf{Com} = (\mathsf{S}, \mathsf{R})$ in the following sense: $\mathsf{S}(\mathbf{M})$ generates a fresh public key pk and outputs a commitment $com = (pk, \mathsf{Enc}(pk, \mathbf{M}; r))$ and an opening $dec = (\mathbf{M}, r)$. Here Enc denotes the encryption algorithm of the encryption scheme, and r denotes the randomness used while encrypting $\mathbf{M}$. Verification of $(com, dec) = (pk, C, \mathbf{M}, r)$ checks that $\mathsf{Enc}(pk, \mathbf{M}; r) = C$.

Note that the third assumption does not follow from the scheme's correctness. Indeed, correctness implies only that *honestly* generated $(pk, \mathbf{M})$ 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 [11, 4, 12, 18, 14], which however generally involve periodical interaction and are comparatively inefficient.

Application of our impossibility results. Attacks in this setting cannot be simulated in a black-box way in the sense of, e.g., Canetti et al. [11]: such a simulator would in particular be able to simulate openings (in the sense of Com, i.e., openings of ciphertexts). Hence, this would imply a simulator for Com in the sense of SIM-SO-COM security (Definition 8). Now from Corollary 1 we know that the construction and security analysis of such a simulator requires either a very strong computational assumption, or fundamentally non-black-box techniques. Even worse: if Com is perfectly binding[13], then Corollary 2 shows that not even secrecy in the sense of Definition 10[14] can be proved in a black-box way. On top of that, we cannot hope to use our (non-black-box) SIM-SO-COM secure scheme ZKCom to construct an encryption scheme in a non-black-box way, since ZKCom's commitment phase is inherently interactive.

We stress that these negative results only apply if encryption really constitutes a (binding) commitment scheme in the above sense. In fact, e.g., [11] 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 to a certain extent necessary.[15]

Relation to the works of Bellare et al. and Hemenway and Ostrovsky. Bellare et al. [7] and Hemenway and Ostrovsky [32] construct encryption schemes that are secure under selective openings. Their schemes achieve a security notion that is comparable to our SIM-SO-COM and IND-SO-COM definitions, only

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

[14] in the context of encryption, Definition 10 would translate to a variant of indistinguishability of ciphertexts

[15] "To a certain extent necessary" means that our results imply (black-box) impossibility of adaptively secure *committing* encryption. Of course, there is a gap between "not committing" (in the sense that public keys and encryptions do not commit to plaintexts) and "non-committing" (as defined in [11]), and our arguments do not apply to schemes in that gap.

for encryption schemes instead of commitments. There is no contradiction to our impossibility results because their encryption schemes do not constitute commitment schemes. In fact, they employ "lossy encryption," which allows for the following idea to prove (encryption) security under selective openings. Start with the real attack game. Then, substitute all public keys with "lossy public keys" (i.e., public keys that yield ciphertexts that contain no information about the message). Openings for the adversary are prepared "on the fly," where we assume that a lossy ciphertext can be efficiently opened to any given message. This modified game essentially is a simulated attack game in which the adversary does not get any (real) encryptions, but only lossy encryptions, which can be prepared independently of the true message vector.

In view of our impossibility results about commitment schemes, it is noteworthy that [7, 32] crucially use that honestly generated tuples $(pk, \mathsf{Enc}(pk, \mathbf{M}; r))$ are computationally indistinguishable from tuples (pk', C') for lossy public keys pk' and lossy ciphertexts C'. The proof of selective opening security of their encryption schemes shows furthermore that such "lossy tuples" do not at all commit to a particular message. Hence, their schemes do not constitute commitment schemes in the above sense.

Sender corruptions vs. receiver corruptions. Note that the schemes of [7, 32] build on the fact that only "sender corruptions" are considered. (With "sender corruptions," an opening of a ciphertext does not reveal the secret key, but only the encryption randomness.) In case of "receiver corruptions" (in which an opening reveals the secret key), Nielsen [39] argues that every encryption scheme is "eventually committing." Here, "eventually committing" means that, after encrypting suitably many messages, there is only one secret key that decrypts all of these ciphertexts back to their original messages.[16] In that sense, suitably many encryptions of known messages commit to (the functionality of) the secret key. Hence, using our results, no encryption scheme can be proved secure under selective openings against receiver corruptions with a black-box reduction to standard cryptographic assumptions.

Commitment schemes with trusted set-up. Our impossibility results are formulated in the plain model, i.e., in a model without trusted set-up (such as a common reference string). We remark that a trusted set-up actually does allow to circumvent our impossibility result. For instance, when assuming a trusted set-up of (honestly generated) public keys, *any* public key encryption scheme constitutes a (non-interactive) commitment scheme in the above sense. Namely, the trusted set-up of the encryption public key and correctness of the scheme (for honestly generated public keys) guarantee that an encryption can only be opened one way. Moreover, if an encryption scheme is secure under selective openings, then so it is when viewed as a commitment scheme. In particular, the

[16] Technically, this is not entirely true, since there are schemes like that of Cramer and Shoup [16], in which the secret key is never completely determined by (honest) encryptions. However, our point here is that the functionality of the secret key (i.e., its action on *honest* encryptions) is uniquely determined.

schemes of [7, 32] constitute non-interactive SIM-SO-COM secure commitment schemes, *assuming a trusted set-up.*

In a similar vein, it should be noted that [32] also build non-interactive IND-SO-COM secure commitment schemes from a special type of randomized one-way functions, tightly related to homomorphic encryption. The security (both secrecy and binding) of this commitment scheme relies on the ideal choice of a set of public parameters. Since our impossibility results do not take into account an ideal parameter set-up, there is no contradiction.

As another example, recall our SIM-SO-COM secure commitment scheme ZKCom from Section 3.2. This scheme actually becomes non-interactive when using a *non-interactive* zero-knowledge (NIZK) proof system IP (see, e.g., Goldreich [26], Section 4.10). There is no contradiction to our impossibility results because the resulting scheme still uses non-black-box techniques, and because NIZK proof systems do not exist in the plain model (i.e., without set-up or other additional assumptions). However, when assuming a trusted set-up, one can at least implement NIZK proof systems, and thus build non-interactive SIM-SO-COM secure commitment schemes.

(Non-)programmable trusted set-up. So we can summarize that relative to a trusted set-up, our impossibility results do not necessarily hold. Now it may be interesting to see where exactly things go wrong in the proof of, e.g., Theorem 1, with a trusted set-up. To this end, we can distinguish two kinds of security proofs in trusted set-up settings:
(a) proofs during which the trusted set-up is "re-programmed," and
(b) proofs that take the trusted set-up as a given.
Let us explain: type-(a) proofs actively modify the generation and/or distribution of the trusted set-up. For instance, assume an honestly generated public key for the schemes of [7, 32] as a trusted set-up. As outlined above, we can then view encryptions as commitments, and the resulting commitment scheme is secure under selective openings. However, the proof of, e.g., SIM-SO-COM security constructs a simulator S that tweaks the distribution of the set-up. (Namely, S expects to work with lossy keys instead of honestly generated public keys.) Similarly, during the security proof of NIZK proof systems in the common reference string (CRS) model, the NIZK simulator S^* expects to actively choose the CRS, so that it knows a certain CRS trapdoor.

So abstractly, type-(a) proofs "re-program" the set-up information and consider settings in which the set-up may lose its guarantees. All discussed proofs for selective opening secure commitment schemes in the trusted set-up model are type-(a) proofs. Similarly, proofs in the *programmable* random oracle model (such as the proof of the non-committing encryption scheme of Nielsen [39]) can be seen as type-(a) proofs.

On the other hand, a type-(b) simulator uses an externally given trusted set-up. Many protocols that use a public key infrastructure, or protocols in the generalized universal composability model have type-(b) proofs (e.g., [33, 15]). Also, proofs in the *non-programmable* random oracle model (such as the proof of OAEP [6, 43]) can be seen as type-(b) proofs.

Applicability of our impossibility proofs. Roughly, Theorem 1 says that no *relativizing* (and thus no black-box) reductions for the SIM-SO-COM security of certain commitment schemes exist. In fact, going through the proof shows that this covers type-(b) reductions as above. Technically, a type-(b) reduction simply assumes another oracle that models the trusted set-up information available to all parties. This oracle can be viewed as part of the computational assumption, which is already modeled as an oracle in the proof of Theorem 1.

However, things lie differently for type-(a) reductions. In a type-(a) reduction, the simulator S expects to actively modify the set-up information. As a consequence, S may violate the binding property of the considered commitment scheme. (In fact, this is precisely what the simulators for the SIM-SO-COM secure commitments with trusted set-up sketched above would do in order to put up a successful simulation.) On the other hand, our proof of Theorem 1 uses that the commitment scheme is binding even to S. (More concretely, this is used in the proof of Lemma 3, when event $\mathsf{bad}_{\mathsf{coll}}$ is shown to occur only with negligible probability.) Hence the proof of Theorem 1 fails relative to a trusted set-up and type-(a) simulators.

6 Application to zero-knowledge proof systems

6.1 Graph 3-coloring is composable in parallel

Overview. Dwork et al. [22] have considered the applications of SIM-SO-COM secure commitment schemes to zero-knowledge protocols, in particular to the graph 3-coloring interactive proof system $\mathsf{G3C}$ of Goldreich et al. [28]. Concretely, [22, Theorem 7.6] states that $\mathsf{G3C}$, when instantiated with a SIM-SO-COM secure commitment scheme, retains a relaxed zero-knowledge property called "$S(V, T, D)$ zero-knowledge" under parallel composition. $S(V, T, D)$ zero-knowledge is a variant of zero-knowledge in which the simulator S may depend on the verifier V, on the distinguisher T between real and simulated transcript, and on the considered message distribution D. Unfortunately, [22] could not give a SIM-SO-COM secure commitment scheme to implement their theorem.

Using our scheme ZKCom, we can instantiate and in fact generalize [22, Theorem 7.6]. Concretely, using a refined analysis and the specific structure of ZKCom, we show that $\mathsf{G3C}$, when implemented with ZKCom, is zero-knowledge under parallel composition. This result does not contradict the negative composability results of Goldreich and Krawczyk [27], Canetti et al. [13]. Namely, on the one hand, we use non-black-box techniques, similar to Barak [1]. On the other hand, our construction of ZKCom already assumes a concurrently composable zero-knowledge proof system. Hence, while our result does not immediately yield an efficient instantiation of $\mathsf{G3C}$, it shows that there is hope for realizing $\mathsf{G3C}$ in a composable way.

A detailed technical treatment follows.

Commit-choose-open protocols. We can actually prove parallel composability of a larger class of "commit-choose-open" style interactive argument systems:

Definition 11 (Commit-choose-open (CCO) protocol). *Let* $\mathsf{IP} = (\mathsf{P}, \mathsf{V})$ *be an interactive argument system for an NP-language* $\mathcal{L}$ *with witness relation* $\mathcal{R}$*. Let* $n = n(k) > 0$ *be polynomially bounded, 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* $[n]$*. We say that* IP *is a* commit-choose-open (CCO) protocol (that uses commitment scheme Com) *if the following holds. First, we require that* IP *is of the following form:*

1. P*, upon input* (x, w) *with* $x \in \mathcal{L}$ *and* $\mathcal{R}(x, w)$*, selects* n *messages* $(M_i)_{i\in[n]}$*,*
2. P *engages in* n *instances of* Com *to commit to the* M_i *at* R*,*
3. V*, upon input* x*, chooses a subset* $I \in \mathcal{I}_n$ *and sends* I *to* P*,*
4. P *opens all* Com*-commitments to* M_i *with* $i \in I$*,*
5. V *accepts if the openings are valid and if the opened values satisfy some fixed relation specified by the protocol.*

Second, we require that the messages M_I *opened by* P *in the third step are uniform and independent values over their respective domain. (In particular,* M_I *can be efficiently sampled without knowing a witness* w*.)*

It is easy to verify that the mentioned graph 3-coloring protocol $\mathsf{G3C}$ [28] is a CCO protocol. Also, trivially, the parallel composition of many instances of a CCO protocol is again a CCO protocol. In particular, in the following, we will for simplicity only talk about a single CCO protocol, while one should actually have the parallel composition of, e.g., $\mathsf{G3C}$ in mind.

Auxiliary-input SIM-SO-COM security. We will prove that any CCO protocol, when using a commitment scheme which is simulatable under selective openings, is black-box zero-knowledge. To this end, we need a refinement of SIM-SO-COM security, which captures auxiliary input and an order of quantifiers as in the zero-knowledge definition.

Definition 12 (AI-SIM-SO-COM). *In the situation of Definition 8, we say that* Com *is AI-SIM-SO-COM secure, iff for every PPT adversary* A*, there exists a PPT simulator* S*, such that for every PPT relation* R*, every PPT* n*-message distribution* $\mathcal{M}$*, and all auxiliary inputs* $z^{\mathcal{M}} = (z^{\mathcal{M}}_k)_{k\in\mathbb{N}} \in (\{0,1\}^*)^{\mathbb{N}}$*,* $z^A = (z^A_k)_{k\in\mathbb{N}} \in (\{0,1\}^*)^{\mathbb{N}}$*, and* $z^R = (z^R_k)_{k\in\mathbb{N}} \in (\{0,1\}^*)^{\mathbb{N}}$*, we have that the advantage* $\mathsf{Adv}^{\mathsf{sim\text{-}so}}_{\mathsf{Com},\mathcal{M},A,S,R,z^{\mathcal{M}},z^A,z^R}$ *is negligible. Here,* $\mathsf{Adv}^{\mathsf{sim\text{-}so}}_{\mathsf{Com},\mathcal{M},A,S,R,z^{\mathcal{M}},z^A,z^R}$ *is defined as* $\mathsf{Adv}^{\mathsf{sim\text{-}so}}_{\mathsf{Com},\mathcal{M},A,S,R}$*, with the following differences:*

- $\mathcal{M}$ *gets additional input* $z^{\mathcal{M}}$*,*
- A *and* S *get additional input* z^A*, and*
- R *gets additional input* z^R*.*

We claim that our scheme ZKCom from Section 3.2 satisfies Definition 12. To see this, recall that the simulator S constructed in the proof of Theorem 2 works also in the presence of auxiliary input. Furthermore, S does not depend on $\mathcal{M}$ and R. However, since $\mathcal{M}$, S, A, and R all receive an auxiliary input in the

AI-SIM-SO-COM experiment, we must demand that the commitment schemes $\mathsf{Com}^{\mathsf{b}}$ and $\mathsf{Com}^{\mathsf{h}}$ against non-uniform adversaries. We get:

Theorem 5 (ZKCom is AI-SIM-SO-COM). *Suppose that there exist one-way permutations secure against non-uniform adversaries. Then the commitment scheme* ZKCom *from Section 3.2 can be instantiated such that* ZKCom *achieves AI-SIM-SO-COM security for arbitrary n, $\mathcal{I}$.*

The following theorem is a generalization of Dwork et al. [22], Theorem 7.6:

Theorem 6 (AI-SIM-SO-COM implies zero-knowledge). *Let* $\mathsf{IP} = (\mathsf{P}, \mathsf{V})$ *be a CCO protocol that uses a commitment scheme* Com *that is AI-SIM-SO-COM secure for n and $\mathcal{I}$ as used in* IP*. Then* IP *is zero-knowledge in the sense of Definition 4.*

Proof. Assume V^*, (x, w), D, z^{V^*}, and z^D as in Definition 4. We will construct a suitable PPT simulator S^*. Since IP is a CCO protocol, we can immediately use the AI-SIM-SO-COM security of Com. To this end, we define an adversary A, a message distribution $\mathcal{M}$, a relation R, and auxiliary inputs z^A and z^R as in Definition 12.

Concretely, define $z^{\mathcal{M}} = (x, w)$ and let $\mathcal{M}$ be the PPT n-message distribution that is induced by P on input (x, w). Furthermore, let $z^A = (x_k, z^{V^*})$ and let $A = V^*$, except that A finally outputs a transcript of its conversation. We hence have $out_A = \langle \mathsf{P}(x_k, w_k), V^*(x_k, z_k^{V^*})\rangle$. Finally, set $z^R = z^D$ and $R(\mathbf{M}, out, z^R) = D(out, z^R)$, such that R outputs exactly what D outputs on real transcripts as in Definition 4. Now Definition 12 guarantees that there exists a PPT machine S such that

$$\begin{aligned} &\Pr\left[R(\mathbf{M}, out_A, z^R) = 1\right] - \Pr\left[R(\mathbf{M}, out_S, z^R) = 1\right] \\ &\qquad = \Pr\left[D(\langle \mathsf{P}(x_k, w_k), V^*(x_k, z_k^{V^*})\rangle, z^D) = 1\right] - \Pr\left[D(out_S, z^D) = 1\right] \end{aligned}$$

is negligible, where out_S is the final output of S in the ideal AI-SIM-SO-COM experiment. Note that out_S is still obtained through an interactive experiment that in particular requires knowledge about $\mathbf{M}$ and hence the witness w. However, the only information S actually receives about the message vector $\mathbf{M}$ is the subset M_I. Since IP is a CCO protocol in the sense of Definition 11, M_I is statistically independent of (x, w). Hence we can construct the following machine S^* which has oracle access to $A = V^*$. Namely, S^* internally simulates S (and relays to S^* its own oracle access to A). As soon as S outputs a set I, S^* answers with a uniformly and independently sampled set M_I. Note that S^* no longer takes part in a AI-SIM-SO-COM experiment, but instead works with input $z^A = (x_k, z^{V^*})$ and oracle access to V^* alone. By the CCO property of IP, we obtain

$$\Pr\left[D(out_S, z^D) = 1\right] = \Pr\left[D(S^*(x_k, z^{V^*}, z^D) = 1\right],$$

and hence, putting things together shows that $\mathsf{Adv}^{\mathsf{ZK}}_{V^*, S^*, (x,w), D, z^{V^*}, z^D}$ is indeed negligible.

Observing that the mentioned graph 3-coloring protocol $\mathsf{G3C}$ from Goldreich et al. [28] is a CCO protocol, and that the set of CCO protocols are closed under parallel composition we get:

Corollary 3 ($\mathsf{G3C}$ is composable in parallel). *The graph 3-coloring protocol* $\mathsf{G3C}$*, when implemented with our commitment scheme* ZKCom*, is zero-knowledge, even under parallel composition.*

6.2 IND-SO-COM security and witness indistinguishability

Overview. A natural question is whether IND-SO-COM security, our relaxation of SIM-SO-COM security, provides a reasonable fallback for SIM-SO-COM security. Now first, our results show that even when using IND-SO-COM secure schemes, we cannot rely on perfectly binding commitment schemes because of Theorem 3. For many interesting interactive proofs (and in particular the mentioned graph 3-coloring protocol $\mathsf{G3C}$), 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." Namely, 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. The resulting argument system will be witness-indistinguishable, and the security reduction is tight. (In particular, the security reduction does not lose a factor of $|\mathcal{I}|$, where $|\mathcal{I}|$ is the number of possible challenges sent by the verifier.)

We stress that, since the set of commit-choose-open protocols is closed under parallel composition, we get composability "for free." Now witness indistinguishable argument systems already enjoy a composition theorem (see Feige and Shamir [24] or Goldreich [26], Lemma 4.6.6), so the compositionality claim is not surprising. However, we believe that our results demonstrate that the security notion of IND-SO-COM secure commitments itself is a reasonable fallback to SIM-SO-COM security.

We proceed to provide details.

Witness indistinguishability. We first recall the definition of witness indistinguishability (a relaxation of zero-knowledge) from Feige and Shamir [24], where we chose a slightly different but equivalent formulation:

Definition 13 (Witness indistinguishability). *Let* $\mathsf{IP} = (\mathsf{P}, \mathsf{V})$ *be an interactive proof or argument system for language* $\mathcal{L}$ *with witness relation* $\mathcal{R}$. IP *is* witness indistinguishable *iff for every PPT machines* V^* *and* D*, all sequences* $x = (x_k)_{k\in\mathbb{N}}$*,* $w^0 = (w_k^0)_{k\in\mathbb{N}}$*, and* $w^1 = (w_k^1)_{k\in\mathbb{N}}$ *with* $\mathcal{R}(x_k, w_k^0)$ *and* $\mathcal{R}(x_k, w_k^1)$ *for all* k *and* $|x_k|$ *polynomial in* k*, and all auxiliary inputs* $z = (z_k)_{k\in\mathbb{N}} \in (\{0,1\}^*)^{\mathbb{N}}$*, we have that*

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

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

Auxiliary-input IND-SO-COM security. Since the standard definition of witness indistinguishability (see Definition 13) involves an auxiliary input z given to the verifier/adversary V^*, we also consider a variation of Definition 10 that involves auxiliary input. Namely,

Definition 14 (AI-IND-SO-COM). *In the situation of Definition 10, we say that* Com *is* AI-IND-SO-COM secure *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.*

We stress that the proof of Theorem 4 shows AI-IND-SO-COM security, once the investigated commitment scheme is statistically hiding against non-uniform adversaries.

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

Theorem 7 (AI-IND-SO-COM implies witness indistinguishability). *Assume a CCO protocol* IP *with parameters n' and $\mathcal{I}'$ that uses commitment scheme* Com *as in Definition 11. If* Com *is AI-IND-SO-COM for parameters $n = n' + 1$ and $\mathcal{I} = \mathcal{I}'$, then* IP *is witness indistinguishable. The security reduction loses only a factor of 2.*

Proof. Assume arbitrary x, w^0, w^1, V^*, D, z as in Definition 13. We construct a message distribution $\mathcal{M}$, an adversary A, and a z' such that

$$\mathsf{Adv}^{\mathsf{ind\text{-}so}}_{\mathsf{Com},\mathcal{M},A,z} = \Pr\left[\mathsf{Exp}^{\mathsf{ind\text{-}so\text{-}real}}_{\mathsf{Com},\mathcal{M},A,z} = 1\right] - \Pr\left[\mathsf{Exp}^{\mathsf{ind\text{-}so\text{-}ideal}}_{\mathsf{Com},\mathcal{M},A,z} = 1\right]$$
$$= \frac{1}{2}\mathsf{Adv}^{\mathsf{WI}}_{x,w^0,w^1,V^*,D,z}.$$

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, running in the IND-SO-COM experiment, proceeds as follows:

1. upon input $z'_k = (x_k, w^0_k, w^1_k, z_k)$, start an internal simulation of V^* on input (x_k, z_k),
2. upon receiving $n = n' + 1$ Com-commitments from the experiment, relay the first n' of these commitments to V^*, and receive the $(n' + 1)$-th commitment,
3. when V^* chooses a set $I \subseteq [n']$, relay this set (interpreted as a subset of $[n] = [n' + 1]$) to the experiment,
4. upon receiving openings (for $i \in I$) from the experiment, relay these openings to V^*,

5. when the interaction between experiment and machine V^* finishes, run $b' \leftarrow D(x_k, z_k, T)$ to obtain a bit b', where T denotes the transcript of the interaction between the experiment and V^*,
6. upon receiving a message vector $\mathbf{M}^* = ({M^*}_i)_{i \in [n]}$ from the experiment, output $b' \oplus {M^*}_{n'+1}$.

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 $\mathsf{P}(x_k, w_k^0)$ and $V^*(x_k, z_k)$ is perfectly simulated. Since ${M^*}_{n'+1} = b = 0$, consequently A and also $\mathsf{Exp}^{\mathsf{ind\text{-}so\text{-}real}}_{\mathsf{Com},\mathcal{M},A,z}$ output $D(x_k, z_k, \langle \mathsf{P}(x_k, w_k^0), V^*(x_k, z_k)\rangle)$. Conversely, if $b = 1$, then $\mathsf{Exp}^{\mathsf{ind\text{-}so\text{-}real}}_{\mathsf{Com},\mathcal{M},A,z}$ outputs $1 - D(x_k, z_k, \langle \mathsf{P}(x_k, w_k^1), V^*(x_k, z_k)\rangle)$ because ${M^*}_{n'+1} = b = 1$ then. We get that

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

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 IP is a CCO protocol, M_I is independent of the used witness. Hence M_I is also independent of b, and so ${M^*}_{n'+1}$ 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 the theorem.

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. [28]. Their proof works also for perfectly binding commitment schemes (thus achieving an interactive *proof* system), which we (almost) cannot hope to satisfy AI-IND-SO-COM security, according to Theorem 3. However, since we can instantiate AI-IND-SO-COM secure schemes for arbitrary parameters n and $\mathcal{I}$, we can hope to apply Theorem 7 even to protocols where $|\mathcal{I}_n|$ is super-polynomial.[17] In particular, we can apply our theorem to a parallel composition of a CCO protocol (which is again a CCO protocol). This gives a composition theorem for the witness indistinguishability of CCO protocols (implemented with AI-IND-SO-COM secure commitments) at virtually no extra cost.

What our positive results do not imply (and what our negative results do imply). We emphasize as well that our results do *not* imply that there are no, in the terminology of [22], "magic functions." In order to prove nonexistence of magic functions with [22, Theorem 5.1], one would have to find a

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

non-interactive SIM-SO-COM secure commitment scheme. Our negative result Theorem 1 states that this will not be possible with black-box reductions to standard assumptions.

Acknowledgments

I am indebted to Enav Weinreb, Marc Stevens, Serge Fehr, Krzysztof Pietrzak, and Ivan Damgård for many insightful discussions. In particular, Enav suggested interactive proof systems as an application, Marc contributed to the first negative result, and Ivan suggested some of the definitional variants. I am grateful to the anonymous referees whose useful comments helped to significantly improve the paper.

References

[1] B. Barak. How to go beyond the black-box simulation barrier. In *42th Annual Symposium on Foundations of Computer Science, Proceedings of FOCS 2001*, pages 106–115. IEEE Computer Society, 2001.

[2] B. Barak and O. Goldreich. Universal arguments and their applications. In *17th Annual IEEE Conference on Computational Complexity, Proceedings of CoCo 2002*, pages 194–203. IEEE Computer Society, 2002.

[3] B. Barak, M. Prabhakaran, and A. Sahai. Concurrent non-malleable zero-knowledge. In *47th Annual Symposium on Foundations of Computer Science, Proceedings of FOCS 2006*, pages 345–354. IEEE Computer Society, 2006.

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

[5] M. Bellare and P. Rogaway. Random oracles are practical: A paradigm for designing efficient protocols. In *1st ACM Conference on Computer and Communications Security, Proceedings of CCS 1993*, pages 62–73. ACM Press, 1993.

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

[7] M. Bellare, D. Hofheinz, and S. Yilek. Possibility and impossibility results for encryption and commitment secure under selective opening. In A. Joux, editor, *Advances in Cryptology, Proceedings of EUROCRYPT 2009*, number 5479 in Lecture Notes in Computer Science, pages 1–35. Springer-Verlag, 2009.

[8] M. Blum. Coin flipping by telephone. In A. 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.

[9] R. Canetti. Universally composable security: A new paradigm for cryptographic protocols. In *42th Annual Symposium on Foundations of Computer Science, Proceedings of FOCS 2001*, pages 136–145. IEEE Computer Society, 2001.

[10] R. Canetti and M. Fischlin. Universally composable commitments. In J. Kilian, editor, *Advances in Cryptology, Proceedings of CRYPTO 2001*, number 2139 in Lecture Notes in Computer Science, pages 19–40. Springer-Verlag, 2001.

[11] R. Canetti, U. Feige, O. Goldreich, and M. 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.

[12] R. Canetti, C. Dwork, M. Naor, and R. Ostrovsky. Deniable encryption. In B. 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.

[13] R. Canetti, J. Kilian, E. Petrank, and A. Rosen. Concurrent zero-knowledge requires $\tilde{\Omega}(\log n)$ rounds. In *33th Annual ACM Symposium on Theory of Computing, Proceedings of STOC 2001*, pages 570–579. ACM Press, 2001.

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

[15] R. Canetti, Y. Dodis, R. Pass, and S. Walfish. Universally composable security with global setup. In S. Vadhan, editor, *Theory of Cryptography, Proceedings of TCC 2007*, number 4392 in Lecture Notes in Computer Science, pages 61–85. Springer-Verlag, 2007.

[16] R. Cramer and V. Shoup. Design and analysis of practical public-key encryption schemes secure against chosen ciphertext attack. *SIAM Journal on Computing*, 33(1):167–226, 2003.

[17] I. Damgård. A "proof-reading" of some issues in cryptography. In L. Arge, C. Cachin, T. Jurdzinski, and A. 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.

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

[19] I. B. Damgård, T. P. Pedersen, and B. Pfitzmann. On the existence of statistically hiding bit commitment schemes and fail-stop signatures. In D. R. Stinson, editor, *Advances in Cryptology, Proceedings of CRYPTO '93*, number 773 in Lecture Notes in Computer Science, pages 250–265. Springer-Verlag, 1994.

[20] Y. Dodis, R. Oliveira, and K. Pietrzak. On the generic insecurity of the full domain hash. In V. Shoup, editor, *Advances in Cryptology, Proceedings of CRYPTO 2005*, number 3621 in Lecture Notes in Computer Science, pages 449–466. Springer-Verlag, 2005.

[21] D. Dolev, C. Dwork, and M. Naor. Non-malleable cryptography. In *Twenty-Third Annual ACM Symposium on Theory of Computing, Proceedings of STOC 1991*, pages 542–552. ACM Press, 1991. Extended abstract.

[22] C. Dwork, M. Naor, O. Reingold, and L. Stockmeyer. Magic functions. *Journal of the ACM*, 50(6):852–921, 2003.

[23] C. Dwork, M. Naor, and A. Sahai. Concurrent zero-knowledge. *Journal of the ACM*, 51(6):851–898, 2004.

[24] U. Feige and A. Shamir. Witness indistinguishability and witness hiding protocols. In *Twenty-Second Annual ACM Symposium on Theory of Computing, Proceedings of STOC 1990*, pages 416–426. ACM Press, 1990.

[25] R. Gennaro and S. Micali. Independent zero-knowledge sets. In M. Bugliese, B. Preneel, V. Sassone, and I. Wegener, editors, *Automata, Languages and Programming, 33th International Colloquium, Proceedings of ICALP 2006*, number 4052 in Lecture Notes in Computer Science, pages 34–45. Springer-Verlag, 2006.

[26] O. Goldreich. *Foundations of Cryptography – Volume 1 (Basic Tools)*. Cambridge University Press, Aug. 2001.

[27] O. Goldreich and H. Krawczyk. On the composition of zero-knowledge proof systems. *SIAM Journal on Computing*, 25(1):169–192, 1996.

[28] O. Goldreich, S. Micali, and A. 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.

[29] I. Haitner and T. Holenstein. On the (im)possibility of key dependent encryption. In O. Reingold, editor, *Theory of Cryptography, Proceedings of TCC 2009*, Lecture Notes in Computer Science. Springer-Verlag, 2009. To be published.

[30] I. Haitner and O. 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.

[31] I. Haitner, J. J. Hoch, O. Reingold, and G. Segev. Finding collisions in interactive protocols – a tight lower bound on the round complexity of statistically-hiding commitments. In *48th Annual Symposium on Foundations of Computer Science, Proceedings of FOCS 2007*, pages 669–679. IEEE Computer Society, 2007.

[32] B. Hemenway and R. Ostrovsky. Re-randomizable encryption implies selective opening security. IACR ePrint Archive, Feb. 2009.

[33] D. Hofheinz, J. Müller-Quade, and D. Unruh. Universally composable zero-knowledge arguments and commitments from signature cards. *Tatra Mountains Mathematical Publications*, pages 93–103, 2007.

[34] C.-Y. Hsiao and L. Reyzin. Finding collisions on a public road, or do secure hash functions need secret coins? In M. K. Franklin, editor, *Advances in Cryptology, Proceedings of CRYPTO 2004*, number 3152 in Lecture Notes in Computer Science, pages 92–105. Springer-Verlag, 2004.

[35] R. Impagliazzo and S. 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.

[36] J. Kilian and E. Petrank. Concurrent and resettable zero-knowledge in poly-logarithmic rounds. In *33th Annual ACM Symposium on Theory of Computing, Proceedings of STOC 2001*, pages 560–569. ACM Press, 2001.

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

[38] M. Naor and M. Yung. Universal one-way hash functions and their cryptographic applications. In *Twenty-First Annual ACM Symposium on Theory of Computing, Proceedings of STOC 1989*, pages 33–43. ACM Press, 1989.

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

[40] M. Prabhakaran, A. Rosen, and A. Sahai. Concurrent zero knowledge with logarithmic round complexity. In *43th Annual Symposium on Foundations of Computer Science, Proceedings of FOCS 2002*, pages 366–375. IEEE Computer Society, 2002.

[41] O. Reingold, L. Trevisan, and S. P. Vadhan. Notions of reducibility between cryptographic primitives. In M. Naor, editor, *Theory of Cryptography, Proceedings of TCC 2004*, number 2951 in Lecture Notes in Computer Science, pages 1–20. Springer-Verlag, 2004.

[42] R. Richardson and J. Kilian. On the concurrent composition of zero-knowledge proofs. In J. Stern, editor, *Advances in Cryptology, Proceedings of EUROCRYPT '99*, number 1592 in Lecture Notes in Computer Science, pages 415–431. Springer-Verlag, 1999.

[43] V. Shoup. OAEP reconsidered. In J. Kilian, editor, *Advances in Cryptology, Proceedings of CRYPTO 2001*, number 2139 in Lecture Notes in Computer Science, pages 239–259. Springer-Verlag, 2001.

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

[45] H. Wee. One-way permutations, interactive hashing and statistically hiding commitments. In S. Vadhan, editor, *Theory of Cryptography, Proceedings of TCC 2007*, number 4392 in Lecture Notes in Computer Science, pages 419–433. Springer-Verlag, 2007.

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

The intuitive contradiction. The formulations of Theorem 1 and Theorem 3 seem intuitively much too general: essentially they claim impossibility of black-box 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., Definition 10 (i.e., IND-SO-COM security)? After all, Definition 10 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 Definition 10, in the sense that

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

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 (in particular, using a trivial reduction), and this seems to contradict Theorem 3. 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-black-box 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-black-box. In other words, this way of achieving IND-SO-COM security cannot be black-box, 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 he 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[18], $\mathcal{P}$ 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 Definition 10, *can* access all available oracles.

So for the above modeling of IND-SO-COM as a property $\mathcal{P}$ in the sense of Definition 9, our impossibility results still hold, but become meaningless (since basically using property $\mathcal{P}$ makes the proof non-black-box). In a certain sense, this comes from the fact that the modeling of IND-SO-COM as a property $\mathcal{P}$ is inherently non-black-box. A similar argument holds for the message distribution in the SIM-SO-COM experiment; there, however, we face the additional problem of modeling the existence of a simulator in a property.

What computational assumptions can be formalized as properties in a "black-box" way? Fortunately, most standard computational assumptions can be modeled in a black-box 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 black-box 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 Theorem 1 and Theorem 3 is very general and useful.

(Non-)programmable random oracles. We stress that the black-box requirement for random oracles (when used in the role of $\mathcal{X}$) corresponds to "non-programmable random oracles" (as used by, e.g., Bellare and Rogaway [6]) as opposed to "programmable random oracles" (as used by, e.g., Nielsen [39]). 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 black-box. And indeed, with programmable random oracles, even non-interactive

[18] by Definition 9, $\mathcal{P}$ must be specified independently of additional oracles; if we did allow $\mathcal{P}$ to access additional oracles, this would break our impossibility proofs

SIM-SO-COM secure commitment schemes can be built relatively painless. As an example, [39] proves a simple encryption scheme (which can be interpreted as a non-interactive commitment scheme) secure under selective openings.

What if we change the definition of IND-SO-COM? One referee raised the natural question of what would happen if we changed the IND-SO-COM definition such that it *is* the property-based definition discussed above. (In other words, in the modified IND-SO-COM definition, we do not quantify over all $\mathcal{M}$; instead, A initially transmits a description of $\mathcal{M}$ to the security experiment.) Let us call this definition IND-SO-COM$'$. It is clear that there is a trivial black-box reduction of IND-SO-COM$'$ security to the property $\mathcal{P}$ that models IND-SO-COM security. Correspondingly, the proof of Theorem 1 would cease to hold for IND-SO-COM$'$, since we could not express message distributions that depend on auxiliary oracles (such as the oracle $\mathcal{RO}$ from that proof). However, a trivial reduction shows that IND-SO-COM$'$ security implies IND-SO-COM security in the sense of Definition 10.

However, now even a (technically) fully black-box reduction that shows or employs IND-SO-COM$'$ security might use the *code* of the message distribution $\mathcal{M}$ (simply because that code is transmitted in the clear). Specifically, the straightforward reduction of IND-SO-COM security to IND-SO-COM$'$ security makes use of the code of $\mathcal{M}$. Hence, we can say that our impossibility results implicitly refer to reductions that are black-box *with respect to the message distribution.* In fact, one could hope for a, say, IND-SO-COM secure commitment scheme whose proof circumvents Theorem 3 merely by using the code of the message distribution. However, as soon as a black-box proof of IND-SO-COM$'$ security only makes black-box use of $\mathcal{M}$, it gives rise to a black-box proof of IND-SO-COM security, and Theorem 3 applies.