On Soundness Notions for Interactive Oracle Proofs

Alexander R. Block[1], Albert Garreta[2], Pratyush Ranjan Tiwari[3], and Michał Zając[2]

[1]Georgetown University and UMD College Park, alexander.r.block@gmail.com
[2]Nethermind, {albert,michal}@nethermind.io
[3]Johns Hopkins University, pratyush@cs.jhu.edu

Abstract

Interactive oracle proofs (IOPs) (Ben-Sasson et al., TCC 2016) have emerged as a powerful model for proof systems which generalizes both Interactive Proofs (IPs) and Probabilistically Checkable Proofs (PCPs). While IOPs are not any more powerful than PCPs from a complexity theory perspective, their potential to create succinct proofs and arguments has been demonstrated by many recent constructions achieving better parameters such as total proof length, alphabet size, and query complexity. In this work, we establish new results on the relationship between various notions of soundness for IOPs. First, we formally generalize the notion of round-by-round soundness (Canetti et al., STOC 2019) and round-by-round knowledge soundness (Chiesa et al., TCC 2019). Given this generalization, we then examine its relationship to the notions of generalized special soundness (Attema et al., CRYPTO 2021) and generalized special unsoundness (Attema et al., TCC 2022). We show that:

1. generalized special soundness implies generalized round-by-round soundness;
2. generalized round-by-round knowledge soundness implies generalized special soundness;
3. generalized special soundness does not imply generalized round-by-round knowledge soundness;
4. generalized round-by-round soundness (resp., special unsoundness) is an upper bound (resp., a lower bound) on standard soundness, and that this relationship is tight when the round-by-round soundness and special unsoundness errors are equal; and
5. any special sound IOP can be transformed via (a variant of) the Fiat-Shamir transformation into a non-interactive proof that is adaptively sound in the Quantum Random Oracle Model.

1 Introduction

Probabilistic proof systems live at the heart of complexity theory and cryptography. Improvements in the practical efficiency of these proof systems have led to breakthroughs in zero-knowledge, delegation of computation, and other areas. Interactive oracle proofs (IOPs) were recently proposed [BCS16, RRR21] and have emerged as a powerful model for proof systems. Many recent constructions [KPV19, BCS16, CMS19, COS20, BBHR18, BGKS20, GWC19, Pol22, CBBZ23] of highly efficient and succinct proofs and arguments are compiled from IOPs. IOPs combine aspects of both probabilistically checkable proofs (PCPs) and interactive proofs (IPs), allowing a multi-round interaction between the prover and the verifier. A μ-round IOP can be viewed as a μ-round interactive proof (IP) where the verifier has PCP-like access to each prover message.

With the emergence of IOPs, the quest for understanding the security of IOPs has also started. Fortunately, a recent fruitful line of work has introduced many tools to understand this: these include the notions of state-restoration soundness [BCS16], round-by-round soundness [CCH+19], and (generalized) special soundness [CDS94, Wik21, AFK22] albeit some of these are in the context of multi-round IPs instead of IOPs. Another exciting line of work has attempted to establish relationships between these soundness notions for IPs [Hol19] and studied similar notions for IOPs [CMS19, COS20, KPV19].

This work formally analyzes and establishes the relationship between various soundness notions for IOPs. The first such notion of soundness is round-by-round (RBR) soundness. RBR soundness captures the idea of "persistent

falsehood" in an interactive oracle proof: if the protocol starts off in a situation where a statement is false and should be rejected by the verifier (i.e., "doomed state"), then no matter how cleverly the prover responds in subsequent rounds, the protocol will "forever remain doomed" (except with negligible probability). RBR *knowledge* soundness captures the idea that if there was such a prover that could escape the "doomed state" with a higher (e.g., non-negligible) probability, then there exists an extractor that can extract a valid witness given the (partial) transcript of this interaction. Other key soundness notions we consider are *special soundness* and *special unsoundness* of IOPs. Special soundness was originally introduced in the context of Σ-protocols [GMR89, Bab85] and was later generalized in [CDS94, Wik18, Wik21]. A protocol is considered special sound if given a tree of accepting (see Definition 3.5) transcripts for an input, there exists an extractor that can output a valid witness for the input. Special unsoundness [AFK22], on the other hand, argues about certain verifier challenges being extremely "lucky" for a malicious prover.

Establishing clear relationships among the various notions of soundness for IOPs is interesting because these soundness notions play a critical role in proving the Fiat-Shamir security [FS87] of multi-round protocols. Consequently, understanding these interrelations offers multiple avenues to ensure Fiat-Shamir security, significantly enhancing such cryptographic protocols' applicability. This is demonstrated distinctly by our results (Figure 1), filling several of previously empty space in the realm of relationships between these notions of soundness.

1.1 Our Results

In this paper, we formally establish new relations among generalized round-by-round (RBR) soundness and generalized special soundness. More formally, let $\mathbf{R}$ represent a relation (e.g., an NP relation) for which a μ-round interactive protocol is executed; $\mathcal{L}_{\mathbf{R}}$ then represents the language corresponding to the relation $\mathbf{R}$. Then for any statement $x \notin \mathcal{L}_{\mathbf{R}}$, RBR soundness is then defined with respect a series of "doomed" sets $\mathcal{D}_i$ for all $i = 0, ..., \mu$. These sets represent states of the protocol (comprising of the statement and the transcript so far) from which the prover *cannot* possibly convince the verifier that $x \in \mathcal{L}_{\mathbf{R}}$, except with small probability. Intuitively, the "doomed" set indicates that the protocol is in a point of no return for the prover: no matter what the prover does, except with small probability, the verifier will reject at the end of the interaction. Slightly more formally, RBR soundness (Definition 4.1) requires the following:

- If a statement $x \notin \mathcal{L}_{\mathbf{R}}$, then the protocol begins in this "doomed" state.
- If the current state (comprising the statement and the transcript of all messages so far) is "doomed", then no matter what the prover's next message is, the probability that the next state (including the prover's next message and the verifier's next message) is not doomed is at most $\varepsilon_i(|\mathbb{x}|)$, where ε_i are predefined error functions. This means that once a state is doomed, it is highly likely that all future states will remain doomed, no matter what the prover does.
- After all the interaction rounds, if the interaction ends in a doomed state, then the verifier should indeed reject the statement.

Note that this generalization considers the errors in each round individually, unlike previous definitions as noted in Definition 4.1.[1] RBR *knowledge soundness* is defined with respect to the same framework: a protocol is RBR knowledge sound if there exists an efficient extractor such that if during any round i of the interaction, if the prover can escape the doomed set $\mathcal{D}_i$ with probability larger than ε_i, then the extractor can extract a valid witness from the transcript of this interaction thus far. For RBR knowledge soundness, the protocol *always* begins in a doomed state, even if $x \in \mathcal{L}_{\mathbf{R}}$.

The generalized special soundness notion we consider is due to Attema, Cramer, and Fehr [ACF21] and is defined with respect to a tree of protocol transcripts. For a μ-round IOP $\Pi = (\mathsf{P}, \mathsf{V})$, and $(k_1, \ldots, k_\mu) \in \mathbb{N}^\mu$, a $(k_1, \ldots, k_\mu)$-*tree of accepting transcripts* for $\mathbb{x}$ is a set of $k = \prod_i k_i$ accepting transcripts transcripts $\{\tau_1, \ldots, \tau_k\}$, each with common first message m, arranged in a tree of depth $\mu + 1$ as follows. The nodes in each tree correspond to the prover's messages and the edges correspond to the verifier's challenges, every node at depth $i - 1$ (for $1 \leq i \leq \mu$) has k_i children corresponding to pairwise distinct challenges, and every complete transcript corresponds to exactly one path from the root node to a leaf node in the tree. The protocol Π is $(k_1, \ldots, k_\mu)$-*special sound* if there exists a polynomial time algorithm Ext that, on input $\mathbb{x}$ and any $(k_1, \ldots, k_\mu)$-tree of accepting transcripts for $\mathbb{x}$, outputs a witness $\mathbb{w}$ such that $(\mathbb{x}, \mathbb{w}) \in \mathbf{R}$.

[1] To the best of our knowledge, we are the first to formally define and analyze this generalized notion of round-by-round soundness. It is likely that this notion of RBR soundness has been implicit in prior works (e.g., in RBR soundness proofs of [KPV19] as one example), but we were unable to find prior work formally defining and analyzing this notion of RBR soundness.

The notion of *special unsoundness* [AFK22] says that an IOP $\Pi = (P, V)$ is *ℓ-special unsound* if there exists a dishonest prover strategy P^* such that during any round of the protocol, for any message m sent by P^*, there exists a "lucky" set of verifier challenges $L \subset C$ of size $|L| = \ell$ such that if the verifier V responds with $c \in L$, then P^* can "behave honestly" for the remainder of the protocol and V will accept at the end of the protocol execution; here, C represents the set of verifier challenges.

For these notions of soundness and their generalizations, we then prove the following results; see Figure 1 for a high-level overview of all of our results. First we show that special soundness implies round-by-round soundness.

Theorem 1.1 (Special Soundness Implies RBR Soundness)**.** *Let $\Pi = (\mathsf{P}, \mathsf{V})$ be a μ-round IOP for a relation $\mathbf{R}$. Let C_i be the set of verifier challenges for round $i \in \{1, \ldots, \mu\}$, and let $(k_1, \ldots, k_\mu) \in \mathbb{N}^\mu$. Assume that Π is $(k_1, \ldots, k_\mu)$-special sound. Then Π is RBR sound with errors*

$$\left(\frac{k_1 - 1}{|C_1|}, \ldots, \frac{k_\mu - 1}{|C_\mu|}\right). \tag{1}$$

Next, we show that round-by-round *knowledge* soundness implies special soundness.

Theorem 1.2 (RBR Knowledge Implies Special Soundness)**.** *Let $\Pi = (\mathsf{P}, \mathsf{V})$ be a μ-round IOP for a relation $\mathbf{R}$. Let C_i be the set of verifier challenges for round $i \in \{1, \ldots, \mu\}$. Assume Π has round-by-round knowledge with errors $\varepsilon_1, \ldots, \varepsilon_\mu$, and let*

$$(k_1, \ldots, k_\mu) = (\lceil |C_1|\varepsilon_1 \rceil + 1, \ldots, \lceil |C_\mu|\varepsilon_\mu \rceil + 1).$$

Suppose $\sum_{i \in [\mu]} \prod_{j \in [i]} k_j$ is upper bounded by a polynomial (on the lengths of inputs). Then Π is $(k_1, \ldots, k_\mu)$-special sound.

We follow this up with a negative result: special soundness does not imply round-by-round knowledge soundness.

Theorem 1.3 (Special Soundness does not Imply RBR Knowledge)**.** *Assume $\mathsf{NP} \neq \mathsf{P}$. Then for any polynomial function $\mu(|\mathbb{x}|)$, there exists a μ-round IOP Π with the following properties:*

- *Π is $(1, \overset{\mu}{\ldots}, 1)$-special sound.*
- *If Π is RBR knowledge sound with errors $(\varepsilon_1, \ldots, \varepsilon_\mu)$, then $\varepsilon_i(\ell) = 1$ for some input length ℓ and some $i \in [\mu]$.*

Finally, we show tight relationships between standard soundness, generalized round-by-round soundness, and special unsoundness.

Theorem 1.4 (Relation Between Soundness, Round-by-round Soundness, and Special Unsoundness)**.** *Let Π be a μ-round IOP for a relation $\mathbf{R}$. Assume Π has soundness error ε. Then the following hold:*

- ***RBR soundness is an upper bound for soundness.*** *If Π is round-by-round sound with errors $\varepsilon_1, \ldots, \varepsilon_\mu$, then*

$$\varepsilon \leq 1 - \prod_{i \in [\mu]} (1 - \varepsilon_i)$$

  for all $\mathbb{x} \notin \mathcal{L}_{\mathbf{R}}$.

- ***Special unsoundness is a lower bound for soundness.*** *If Π is special unsound with errors $\varepsilon'_1, \ldots, \varepsilon'_\mu$, then*

$$1 - \prod_{i \in [\mu]} \left(1 - \varepsilon'_i\right) \leq \varepsilon$$

  for all $\mathbb{x} \notin \mathcal{L}_{\mathbf{R}}$. Moreover, there exists a dishonest unbounded prover P^* that, given any input $\mathbb{x}$, manages to make the verifier accept with probability at least $1 - \prod_{i \in [\mu]} \left(1 - \varepsilon'_i\right)$.

- ***Tightness of RBR soundness, soundness, and special unsoundness.*** *Suppose Π is round-by-round sound with errors $\varepsilon_1, \ldots, \varepsilon_\mu$ and that Π is special unsound with the same errors $\varepsilon_1, \ldots, \varepsilon_\mu$. Then*

$$\varepsilon = 1 - \prod_{i \in [\mu]} (1 - \varepsilon_i)\,.$$

  Moreover, the error is tight in the sense that there exists a dishonest prover P^* that, given any input $\mathbb{x}$, manages to have the verifier accept with probability at least ε.

1.1.1 Special Soundness and State-restoration Soundness

Our results on the relationship between special soundness and round-by-round soundness (i.e., Theorem 1.1) gives us new results on the relationship between special soundness and *state-restoration soundness* [BCS16]. Informally, state-restoration (SR) soundness roughly states that an IOP remains secure (i.e., cannot convince a verifier of a false statement) even if a malicious prover is allowed to rewind the verifier to any prior state at most $b \geq 1$ times (see [BCS16] for complete details). It is known that state-restoration soundness and (non-generalized) round-by-round soundness are equivalent [Hol19]; moreover, state-restoration soundness error $\varepsilon_{sr}(b)$ and round-by-round soundness error ε_{rbr} must satisfy $\varepsilon_{\mathsf{sr}}(b) \leq b\varepsilon_{\mathsf{rbr}}$ [CMS19, COS20, KPV19], and this relation holds between state-restoration knowledge soundness and round-by-round knowledge soundness [COS20].

As a direct corollary of the above results (i.e., RBR soundness implies SR soundness) and Theorem 1.1, we obtain the following result.

Corollary 1.5 (Special Soundness Implies State-restoration Soundness)**.** *Let Π be a μ-round $(k_1, \ldots, k_\mu)$-special sound IOP with verifier challenge sets $C_1, \ldots C_\mu$. Then for $b \geq 1$, Π has state-restoration soundness error*

$$\varepsilon_{\mathsf{sr}}(b) \leq b \cdot \max_i \left\{ \frac{k_i - 1}{|C_i|} \right\}.$$

1.1.2 Special Soundness and Quantum-secure Fiat-Shamir

Recent cryptographic research put forth significant effort toward achieving post-quantum security of various cryptographic primitives, which include post-quantum security of non-interactive proofs obtain via the Fiat-Shamir transformation (or variants of this transformation) [CMS19, LZ19, DFMS19, CMSZ21]. With respect to post-quantum security, of interest to us is the so-called BCS transformation due to Ben-Sasson et al. [BCS16]; informally, this transformation compiles any IOP into a non-interactive proof via a variant of the Fiat-Shamir transformation in the random oracle model. In [BCS16] it is shown that applying this transformation to any state-restoration sound IOP results in an adaptively secure non-interactive proof in the random oracle model. The follow up work due to Chiesa et al. [CMS19] extend this work to show that compiling any round-by-round (knowledge) sound IOP with the BCS transformation yields an adaptively (knowledge) sound non-interactive proof in the random oracle model (ROM); we refer the reader to prior work (e.g., [BCS16, CMS19]) for complete details on this transformation. Furthermore, [CMS19] show that this transformation yields a non-interactive proof that is secure in the quantum random oracle model (QROM) (i.e., secure against quantum adversaries that are allowed to query the random oracle in superposition).

As a direct consequence of our results, the work of [CMS19] and Theorem 1.1 shows that any special sound IOP can be compiled via the BCS transformation to obtain an adaptively sound non-interactive proof in the QROM. This gives the following corollary.

Corollary 1.6 (Special Soundness Implies FS Security in the QROM)**.** *Let Π be a μ-round $(k_1, \ldots, k_\mu)$-special sound IOP. Let $\mathsf{BCS}(\Pi)$ be the non-interactive proof obtained by applying the BCS transformation to Π, and let $\varepsilon = \max_{i \in [\mu]}\{(k_i - 1)/|C_i|\}$. Then $\mathsf{BCS}(\Pi)$ has adaptive soundness error $O(t^2\varepsilon + t^3/2^\lambda)$ against quantum attackers that make at most $t - O(q \log \ell)$ queries to the random oracle, where λ is the output length of the random oracle in bits, q is (an upper bound on) the total number of queries made by the verifier during any execution of Π, and ℓ is the total number of symbols sent by both the prover and verifier during any execution of Π.*

Remark 1.7. Note that Corollary 1.6 directly implies that the Fiat-Shamir transformation of any special sound interactive proof is a secure non-interactive proof in the QROM.

To the best of our knowledge, the above corollary is the first result relating the special soundness of a protocol and its security versus quantum adversaries when rendered non-interactive via the BCS transformation (i.e., a variant of the Fiat-Shamir transformation). Thus Corollary 1.6 is the first result to our knowledge relating special soundness of multi-round protocols to quantum security. Notably, due to Theorem 1.3, we *do not obtain* that the BCS transformation of a special sound IOP is knowledge sound versus quantum adversaries; we leave it as an interesting open problem to examine whether special soundness implies non-interactive knowledge soundness versus quantum adversaries. We remark that the Fiat-Shamir transformation of quantum-secure Σ-protocols (i.e., 1-round interactive arguments) was shown to be sound in the QROM [DFMS19].

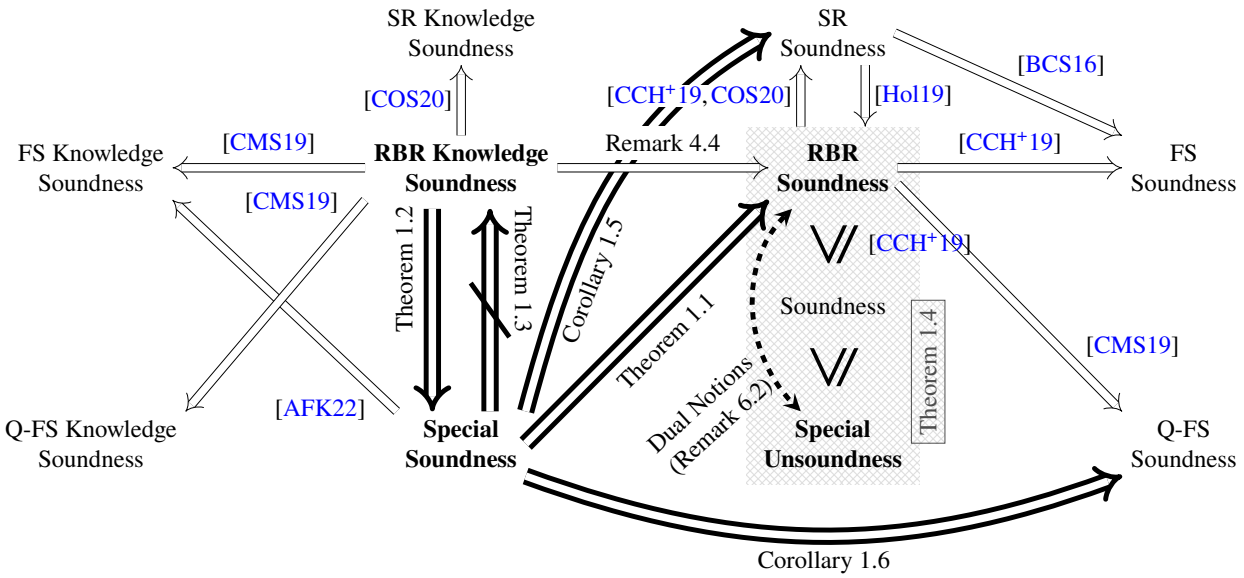

Figure 1: Pictorial overview of the relations between soundness notions. FS and Q-FS denote non-interactive adaptive security of the BCS transformation (i.e., a variant of the Fiat-Shamir transformation) in the random oracle model and quantum random oracle model, respectively. SR denotes state-restoration soundness. " $\Longrightarrow$ " arrows represent implications; " $\not\Longrightarrow$ " represents there is no implication; and the dashed $\leftrightarrow$ represents a relationship between notions (described by the text). Text in bold indicate the main soundness notions we study in this work. Thick arrows, lines, and background shading indicate our contributions. We remark that Corollary 1.5 follows from Theorem 1.1 and [CCH+19, COS20], and Corollary 1.6 follows from Theorem 1.1 and [CMS19].

1.2 Related Work

Interactive oracle proofs were introduced by Ben-Sasson et al. [BCS16], and along with it the notion of state-restoration soundness. This notion of soundness was introduced to formally show the Fiat-Shamir [FS87] security of multi-round IOPs in the random oracle model. The notion of round-by-round soundness was later introduced by Canetti et al. [CCH+19] for a similar reason: to prove Fiat-Shamir security of multi-round protocols, but in the *plain* model. Note, however, that round-by-round soundness readily implies Fiat-Shamir security in the random oracle model [CCH+19]. It was widely known that round-by-round soundness implies state-restoration soundness (e.g., [CCH+19, COS20]), and it was recently shown that state-restoration soundness implies round-by-round soundness [Hol19]. Special soundness was recently shown to also imply Fiat-Shamir security of multi-round protocols [AFK22]; moreover, this work also shows that special unsoundness of multi-round protocols readily admits an attack on the Fiat-Shamir transformed protocol.

Prior to these soundness tools, a variety of work [KRR17, CCRR18, HL18] circumvented the impossibility results of [BDG+13] by using stronger hardness assumptions to construct Fiat-Shamir compatible hash function families. Another line of work [GKR08, CMT12, BCGT13, Tha13, BTVW14, WTs+18, Set20, RR20] follows the frameworks of Kilian [Kil92] and Micali [Mic94] to compile interactive oracle proofs [BCS16] into efficient arguments and SNARKs via collision-resistant hash functions [BCS16, Kil92] or in the random oracle model [BCS16, Mic94].

2 Technical Overview

We give an overview of our main contributions in this section. Before we begin, we informally fix some notations. Informally, for a μ-round IOP $\Pi = (P, V)$ for a relation $\mathcal{R}$ and vector $(k_1, \ldots, k_\mu) \in \mathbb{N}^\mu$, we say that tree T is a $(k_1, \ldots, k_\mu)$-*tree of transcripts* for Π on any input $\mathbb{x} \in \mathcal{L}_\mathcal{R}$ if T is a depth $\mu + 1$ tree such that nodes at level i have k_i outgoing edges, and the tree is labeled in the following way: every node at level i is labeled with a prover message and every edge is labeled with the corresponding challenge; moreover, the root of the tree is a single message m sent by the prover, and all leaves are prover messages. We say that T is an accepting tree of transcripts if all root to leaf paths are accepting transcripts (i.e., they are accepted by V. Now given the notion of a tree of accepting transcripts, the protocol Π is $(k_1, \ldots, k_\mu)$-*special sound* [ACK21] if there exists a polynomial time extractor algorithm that when given as input any $(k_1, \ldots, k_\mu)$ tree of accepting transcripts for an instance $\mathbb{x}$, the extractor outputs a witness w such that $(x, w) \in \mathcal{R}$.

2.1 Generalizing Round-by-round Soundness

A first step in our work is generalizing the notions of round-by-round (RBR) soundness and RBR knowledge soundness. We first recall the notion of RBR soundness, introduced by Canetti et al. [CCH+19]. Informally, a public-coin interactive protocol for a language $\mathcal{L}$ is *round-by-round sound* (RBR sound) if at any point during the execution of the protocol, the protocol is in a well-defined state (depending on the protocol execution so far) and some of these states are "doomed", where being "doomed" means that no matter what message the prover sends, with overwhelming probability over the verifier messages, the protocol remains "doomed". A bit more formally, RBR soundness error ε states that: (1) if $\mathbb{x} \notin \mathcal{L}$ the initial state of the protocol is "doomed"; (2) if the protocol is in a "doomed" state during any non-final round of the protocol, then for any message sent by the prover, the protocol remains doomed with probability at least $1 - \varepsilon$ over the verifier messages; and (3) if the protocol terminates in a "doomed" state, then the verifier rejects. To generalize RBR soundness, we consider separate errors ε_i for each round of the protocol. That is, for item (2) above, if the protocol is in a "doomed" state at the start of round i, then for any message sent by the prover, the protocol remains doomed with probability at least $1 - \varepsilon_i$ over the verifier messages. We use the notation $\mathcal{D}$ to denote the set of "doomed" states below.[2]

Chiesa et al. [CMS19] et al. extend RBR soundness to the notion of *RBR knowledge soundness*, which roughly says that if (1) the protocol is in a "doomed" state during any round of interaction, and (2) *every* prover message can force the protocol to leave this "doomed" state with probability at least ε_k (over the verifier randomness), then an extractor can efficiently extract a witness (with probability 1) simply by examining the current protocol state and the prover's next message. Again, we extend the notion of RBR knowledge to allow for separate errors $\varepsilon_{\mathsf{k},i}$ during any intermediate round i of the protocol. The formal definitions of RBR soundness and RBR knowledge can be found in Section 3.

[2] One can also consider separate doomed states $\mathcal{D}_i$ for each round i of the protocol, but setting $\mathcal{D} = \cup_i \mathcal{D}_i$ still captures the RBR soundness.

2.2 Special Soundness Implies Round-by-round Soundness

In this section, we give an overview of our first main result: special soundness implies round-by-round soundness. More formally, we show that any $(k_1, \ldots, k_\mu)$ special sound IOP is generalized RBR sound with errors $\varepsilon_i = (k_i - 1)/|C_i|$. As a warm up, we first consider the non-generalized version of special soundness and RBR soundness (i.e., single error bound ε for all rounds of RBR soundness, and only k-ary trees of accepting transcripts).

2.2.1 Warm Up: 1-round IOPs

As a warm-up, consider any 1-round (i.e., 3-move) public coin IOP $\Pi = (P, V)$ with message space $\mathcal{M}$ and challenge space C for some relation $\mathcal{R}$. Suppose that Π is k-special sound. That is, there is an extractor Ext such that on any input $\mathbb{x}$ and any k accepting transcripts of the form (m, c_i, z_i) for $c_i \in C$ distinct for all i and $m, z_1, \ldots, z_k \in \mathcal{M}$, the extractor outputs a witness $\mathbb{w}$ such that $(\mathbb{x}, \mathbb{w}) \in \mathcal{R}$.

Now we claim that Π is round-by-round sound with error $\varepsilon = (k-1)/|C|$. This can be seen as follows. Fix an input $\mathbb{x}$ and consider any first message $m \in \mathcal{M}$ sent by the prover and any challenge c sent by the verifier. We say that $(\mathbb{x}, m, c)$ is *completable* if there exists $z \in \mathcal{M}$ such that $V(\mathbb{x}, m, c, z) = 1$; i.e., the verifier accepts. Intuitively speaking, for Π to be round-by-round sound, then the number of compleatable transcripts should be small for any $\mathbb{x} \notin \mathcal{L}$; in other words, for any first message m sent by the prover, for $c \xleftarrow{\$} C$ the probability that $(\mathbb{x}, m, c)$ is completable should be at most ε defined above. Let $P(\mathbb{x}, m)$ be the probability that $(\mathbb{x}, m, c)$ is completable for $c \xleftarrow{\$} C$, and suppose that $P(\mathbb{x}, m) > \varepsilon$. This implies that there exist $\alpha = \lceil |C| \cdot P(\mathbb{x}, m) \rceil > \varepsilon |C| = (k-1)$ distinct challenges $c_1, \ldots, c_\alpha \in C$ such that $(\mathbb{x}, m, c_i)$ is completable for all $i \in \{1, \ldots, \alpha\}$. Notice that $\alpha > k - 1$ and so $\alpha \geq k$. This tell us that any k of the α completable transcripts $(\mathbb{x}, m, c_i)$ form a k-tree of accepting transcripts of the form $T = \{(\mathbb{x}, m, c_i, z_i)\}_i$ where z_i completes $(\mathbb{x}, m, c_i)$. Thus by k-special soundness of Π, we have that $(\mathbb{x}, \mathsf{Ext}(\mathbb{x}, T)) \in \mathcal{R}$; however, this is a contradiction since $\mathbb{x} \notin \mathcal{L}$. Then it must be the case that at most $\alpha \leq (k-1) = \varepsilon |C|$ distinct challenges can result in a completable transcript, which implies that $P(\mathbb{x}, m) \leq (k-1)/|C|$ as desired.

More formally, to show that Π has round-by-round soundness error $\varepsilon = (k-1)/|C|$, we define a suitable doomed set $\mathcal{D}$ as follows. First consider the following two sets:

- Define $\mathcal{D}_0$ to be the set of all $(\mathbb{x}, \emptyset)$ such that $\mathbb{x} \notin \mathcal{L}_\mathcal{R}$.
- Define $\mathcal{D}_1$ to be the set of all $(\mathbb{x}, m, c)$ for $m \in \mathcal{M}$ and $c \in C$ that are *not completable*.

Then we set $\mathcal{D} = \mathcal{D}_0 \cup \mathcal{D}_1$. Under this definition of $\mathcal{D}$, clearly we have $(\mathbb{x}, \emptyset) \in \mathcal{D}$ for all $\mathbb{x} \notin \mathcal{L}_\mathcal{R}$ by definition; moreover, for any transcript $(\mathbb{x}, m, c)$ that is not completable, then for all $z \in \mathcal{M}$ we have $V(\mathbb{x}, m, c, z) = 0$. Now supposing that $(\mathbb{x}, \emptyset)$ is doomed (by definition), consider any potential prover message m. Let $P(\mathbb{x}, m)$ denote the probability that $(\mathbb{x}, m, c) \notin \mathcal{D}$, where the probability is taken over $c \xleftarrow{\$} C$. Suppose there exists a message $m \in \mathcal{M}$ such that $P(\mathbb{x}, m) > \varepsilon$. Then by our above argument, we reach a contradiction as we can construct a k-tree of accepting transcripts and output a witness $\mathbb{w}$, violating the assumption that $\mathbb{x} \notin \mathcal{L}$. Thus it must hold that $P(\mathbb{x}, m) \leq \varepsilon = (k-1)/|C|$, as desired.

2.2.2 Extending to μ-round IOPs

Notice that the above argument crucially relies on the fact that if you can leave the doomed set in the first round with probability larger than $\varepsilon = (k-1)/|C|$, then you can manifest a k-tree of accepting transcripts. To extend the above argument to $\mu > 1$ round protocols, we want to preserve this above fact. Recall the notion of a completable transcript from above, which states that for a given partial transcript $(\mathbb{x}, m, c)$, there exists a prover message $z \in \mathcal{M}$ that causes the verifier to accept the entire transcript $(\mathbb{x}, m, c, z)$. Now we extend this notion to any μ-round IOP: informally, consider any partial transcript of the form $\tau_i := (\mathbb{x}, m_1, c_1, \ldots, m_i, c_i)$, where $i \leq \mu$ and message $m_j \in \mathcal{M}$ is sent by the prover at the start of round j and the verifier responds with challenge $c_j \xleftarrow{\$} C$, for all $j \leq i$. Now intuitively, we want to say that τ_i is completable if there is a sequence of messages and challenges $\sigma_{i+1} = (m_{i+1}, c_{i+1}, \ldots, m_\mu, c_\mu, m_{\mu+1})$ such that $V(\tau_i, \sigma_{i+1}) = 1$; i.e., (τ_i, σ_{i+1}) is a complete accepting transcript. However, under this definition we would not be able to show our result.

In our above argument for the 1-round IOP case, we crucially relied on fact that the completable transcripts $(\mathbb{x}, m, c)$ form an *entire* k-tree of accepting transcripts. Under our proposed extended definition of a completable transcript, this

fact would no longer hold. To see this, suppose that $(\mathbb{x}, m_1, c_{1,i})$ is completable for at least k challenges $c_{1,1}, \ldots, c_{1,k}$. Under our current definition of completable, this only implies that there exist k sequences of messages and challenges $\sigma_{1,i}$ such that $(\mathbb{x}, m_1, c_{1,i}, \sigma_{1,i})$ is an accepting transcript. Clearly, this is *not* a k-tree of accepting transcripts, so we can no longer derive our contradiction as with the 1-round IOP case.

We address the above issue by extending the definition of a completable transcript to a k-completable transcript. For any partial transcript τ_i for as defined above, we say that τ_i is *k-completable* if there exists a k-tree of transcripts T of depth $\mu + 1 - i$ such that for every $\sigma \in T$, the transcript (τ_i, σ) is accepted by the verifier, where σ is of the form $(m_{i+1}, c_{i+1}, \ldots, m_\mu, c_\mu, m_{\mu+1})$. Now under this definition, we can now proceed to extend our proof to μ-round IOPs.

Suppose that Π is a k-special sound IOP. We argue that Π has round-by-round soundness error $\varepsilon = (k-1)/|C|$. Define a doomed set $\mathcal{D}$ as follows. First, consider the following sets:

- Define $\mathcal{D}_0$ to be the set of all $(\mathbb{x}, \emptyset)$ such that $\mathbb{x} \notin \mathcal{L}_\mathcal{R}$.
- For all $i \in \{1, \ldots, \mu\}$, define $\mathcal{D}_i$ to be the set of partial transcripts $\tau_i = (\mathbb{x}, m_1, c_1, \ldots, m_i, c_i)$ such that τ_i is *not completable*.

Now take $\mathcal{D} = \cup_{i=0}^{\mu} \mathcal{D}_i$. Clearly, as required we have $(\mathbb{x}, \emptyset) \in \mathcal{D}$ for all $\mathbb{x} \notin \mathcal{L}_\mathcal{R}$. Moreover, if $\tau_\mu = (\mathbb{x}, m_1, c_1, \ldots, m_\mu, c_\mu)$ is not k-completable, then for all $m_{\mu+1} \in \mathcal{M}$ we have $V(\tau_\mu, m_{\mu+1}) = 0$ by definition of k-completable.

Let $\tau_0 = (\mathbb{x}, \emptyset)$. We now argue that the probability that $\tau_1 = (\tau_0, m_1, c_1)$ is k-completable is at most ε for all $m_1 \in \mathcal{M}$ and $c_1 \xleftarrow{\$} C$. Our argument now proceeds identically to the 1-round IOP case. Let $P_1(\mathbb{x}, m_1)$ denote the probability over $c_1 \xleftarrow{\$} C$ that $\tau_1 = (\mathbb{x}, m_1, c_1)$ is k-completable; in particular, $P_1(\mathbb{x}, m_1) = \Pr[\tau_1 \notin \mathcal{D}]$ by our definition. Suppose there exists $m_1 \in \mathcal{M}$ such that $P_1(\mathbb{x}, m_1) > \varepsilon$. Now this implies that there exist at $\alpha > \varepsilon|C| = (k-1)$ distinct challenges $c_{1,1}, \ldots, c_{1,\alpha} \in C$ such that $(\mathbb{x}, m_1, c_{1,j})$ is k-completable. By definition of completable, this implies that for every $j \in \{1, \ldots, \alpha\}$ and partial transcript $\tau_{1,j} = (\mathbb{x}, m_1, c_{1,j})$, there exists a k-tree of transcripts $T_{2,j}$ of depth μ such that for all $\sigma \in T_{1,j}$, the complete transcript $(\tau_{1,j}, \sigma)$ is accepted by the verifier. Since $\alpha \geq k$, taking any k out of α of the partial transcripts $(\mathbb{x}, m_1, c_{1,j})$ along with their k-tree of transcripts $T_{2,j}$ forms a k-tree of accepting transcripts of depth $\mu + 1$. Now by special soundness of Π, there exists an efficient extractor that extracts a witness $\mathbb{w}$ such that $(\mathbb{x}, w) \in \mathcal{R}$ when given this k-tree of accepting transcripts, contradicting our assumption that $\mathbb{x} \notin \mathcal{L}$. Thus it must be the case that $P_1(\mathbb{x}, m_1) \leq \varepsilon$ as required.

Now consider any intermediate round $i \leq \mu - 1$ and let $\tau_i = (\mathbb{x}, m_1, c_1, \ldots, m_1, c_i)$ be a partial transcript for this round. Suppose that $\tau_i \in \mathcal{D}$; we now show that for all $m_{i+1} \in \mathcal{M}$, the probability that $\tau_{i+1} = (\tau_i, m_{i+1}, c_{i+1})$ is k-completable is at most ε for $c_{i+1} \xleftarrow{\$} C$. Let $P_{i+1}(\tau_i, m_{i+1}) = \Pr_{c_{i+1}}[(\tau_i, m_{i+1}, c_{i+1}) \notin \mathcal{D}]$ denote this probability and suppose there exists $m_{i+1} \in \mathcal{M}$ such that $P_{i+1}(\tau_i, m_{i+1}) > \varepsilon$. By the same argument as above and by definition of ε, there exist at least k distinct challenges $c_{i+1,1}, \ldots, c_{i+1,k}$ such that the partial transcript $\tau_{i+1,j} = (\tau_i, m_{i+1}, c_{i+1,j})$ is k-completable. By definition of k-completable, this implies for each $\tau_{i+1,j}$ there exists a k-tree of transcripts $T_{i+2,j}$ of depth $\mu - i$ such that for all $\sigma \in T_{i+1,j}$, the transcript $(\tau_{i+1,j}, \sigma)$ is a complete and accepting transcript. Then we can construct a k-tree of transcripts T_{i+1} of depth $\mu - i + 1$ as $T_{i+1} = \{(\tau_i, m_{i+1}, c_{i+1,j}, \sigma)_{j \in \{1,\ldots,k\}} : \sigma \in T_{i+1,j}\}$. Then clearly T_{i+1} forms a k-tree of transcripts that completes the partial transcript τ_i; this contradicts the assumption that $\tau_i \in \mathcal{D}$, i.e., τ_i is not k-completable. Thus it must hold that $P_{i+1}(\tau_i, m_{i+1}) \leq \varepsilon$ for all $m_{i+1} \in \mathcal{M}$, establishing the result.

2.2.3 Extending to Generalized Special Soundness and Generalized Round-by-round Soundness

The above argument naturally generalizes to $(k_1, \ldots, k_\mu)$ special soundness and $(\varepsilon_1, \ldots, \varepsilon_\mu)$ round-by-round soundness, with message spaces $\mathcal{M}_1, \ldots, \mathcal{M}_{\mu+1}$ and challenge spaces $C_1, \ldots, C_\mu$. First, we define a partial transcript τ_i to be $(k_{i+1}, \ldots, k_\mu)$-*completable* if there exists a $(k_{i+1}, \ldots, k_\mu)$-tree of transcripts T of depth $\mu - i + 1$ such that for all $\sigma \in T$, (τ_i, σ) is a complete transcript and $V(\tau_i, \sigma) = 1$. Then under this definition, we define our doomed set $\mathcal{D} = \cup_{i=0}^{\mu} \mathcal{D}_i$ where $\mathcal{D}_0$ is defined identically as above and $\mathcal{D}_i$ is the set of all partial transcripts τ_i that are not $(k_{i+1}, \ldots, k_\mu)$-completable. Then setting $\varepsilon_i = (k_i - 1)/|C_i|$, we can proceed with the above argument in an identical fashion, deriving our contradictions and showing round-by-round soundness of the IOP Π.

2.2.4 Special Soundness Implies Soundness in the Quantum Random Oracle Model

A direct consequence of the previous result is that the BCS transformation [BCS16] of a special sound IOP is sound in the Quantum Random Oracle Model (QROM). Indeed, by the previous result, any special sound IOP is RBR sound, and, due to a result of [CMS19], the BCS transformation of any RBR sound IOP is a sound non-interactive proof in the QROM. In particular, this implies that the Fiat-Shamir transformation of any special sound Interactive Proof is sound in the QROM.

2.3 Round-by-round Knowledge Implies Special Soundness

In this section, we give an overview of our second main result: round-by-round knowledge soundness implies special soundness.

2.3.1 Warm Up: 1-round IOPs

We again consider a warm up to the proof by analyzing a 1-round public coin IOP $\Pi = (P, V)$ with message space $\mathcal{M}$ and challenge space C for some relation $\mathcal{R}$. Let ε be the round-by-round knowledge error of Π. Let $\mathbb{x}$ be arbitrary and let $\mathcal{D}$ be the doomed set for round-by-round knowledge error. By round-by-round (RBR) knowledge, we know there exists a polynomial time extractor Ext such that that if for all $m \in \mathcal{M}$ it holds that $P(\mathbb{x}, m) = \Pr_c[(\mathbb{x}, m, c) \notin \mathcal{D}] > \varepsilon$, then $\mathsf{Ext}(\mathbb{x}, m)$ outputs a witness $\mathbb{w}$ such that $(\mathbb{x}, \mathbb{w}) \in \mathcal{R}$, where the probability is taken over $c \xleftarrow{\$} C$. Crucially, for all such transcripts $(\mathbb{x}, m, c) \notin \mathcal{D}$, there exists $z \in \mathcal{M}$ such that $(\mathbb{x}, m, c, z)$ is an accepting transcript; we leverage this fact below.

Setting $k = \lceil |C|\varepsilon \rceil + 1$, we now construct a polynomial-time extractor Ext' such that given any k-tree of accepting transcripts $T = \{(\mathbb{x}, m, c_i, z_i) \colon c_i \in C, z_i \in \mathcal{M}\}_{i \in \{1,\ldots,k\}}$, $(\mathbb{x}, \mathsf{Ext}(T)) \in \mathcal{R}$. Suppose that indeed for all $m \in \mathcal{M}$ it holds that $P(\mathbb{x}, m) > \varepsilon$. This implies that for any $m \in \mathcal{M}$, there are $\alpha > \varepsilon|C|$ distinct challenges $c_1, \ldots, c_\alpha$ such that $\mathsf{Ext}(\mathbb{x}, m, c_i)$ outputs a valid witness. This implies that $\alpha \geq \lceil \varepsilon|C| \rceil + 1 = k$, and thus there are at least k distinct challenges $c_1, \ldots, c_k$ such that $\mathsf{Ext}(\mathbb{x}, m, c_i)$ outputs a valid witness, for all $m \in \mathcal{M}$. Now given input a tree of accepting transcripts T defined above, our new extractor $\mathsf{Ext}'(T)$ simply runs Ext on input $(\mathbb{x}, m)$, followed by $(\mathbb{x}, m, c_i, z_i)$ for all $i \in [k]$. Then Ext' outputs the same witness $\mathbb{w}$ that is given by Ext. Note that by assumption, we have that $(\mathbb{x}, m, c_i, z_i) \notin \mathcal{D}$, and we have $k = \lceil \varepsilon|C| \rceil + 1$ such transcripts, so by round-by-round knowledge it must be the case that $\mathsf{Ext}(\mathbb{x}, m, c_i, z_i)$ outputs a valid witness. Otherwise, if $(\mathbb{x}, m, c_i, z_i) \in \mathcal{D}$, then the verifier would reject this transcript, contradicting T being a tree of accepting transcripts.

2.3.2 Extending to μ-round IOPs

Extending the above intuition to μ-round IOPs introduces some subtleties that must be handled. Suppose that Π is again RBR knowledge sound with error ε and let $k = \lceil |C|\varepsilon \rceil + 1$. Suppose we are given a k-tree of accepting transcripts T of depth $\mu + 1$ with root m_1. We now construct a special soundness extractor that extracts a valid witness given the tree T.

Consider any partial transcript $\tau_i = (\mathbb{x}, m_1, c_1, \ldots, m_i, c_i)$ such that τ_i is a path in T. Then there is a unique message m^*_{i+1} such that (τ_i, m^*_{i+1}) is a path in T. Now suppose that τ_i has the following properties:

- $\tau_i \in \mathcal{D}$ (i.e., τ_i is a doomed transcript);
- For all outgoing edges $c_{i+1,1}, \ldots, c_{i+1,k}$ connected to m_{i+1}, we have $(\tau_i, m^*_{i+1}, c_{i,j}) \notin \mathcal{D}$ for all j.

Clearly if τ_i has this property, then our extractor Ext' when given tree T as input, if Ext' runs the RBR knowledge extractor Ext on input $(\tau_i, m^*_{i+1}, c_{i,j})$ then Ext outputs a valid witness. This follows since by or definition of k and by RBR knowledge, the fraction of challenges $c \in C$ such that $(\tau_i, m^*_{i+1}, c) \notin \mathcal{D}$ is strictly larger than ε; i.e., $k/|C| > \varepsilon$ by definition.

Now we claim that for *any* tree of accepting transcripts T, there exists some partial transcript τ_i that satisfies the above stated properties, where $i \in [\mu]$. Supposing this claim is true, then the extractor Ext' on input any k-tree of accepting transcripts T, simply runs Ext on input (τ_j, m^*_{j+1}) over all possible partial transcripts τ_j that are in T and start at the root of T and connect to node m^*_{j+1}. Then clearly by the above claim, Ext outputs a valid witness $\mathbb{w}$ such that $(\mathbb{x}, \mathbb{w}) \in \mathcal{R}$; else this would violate RBR knowledge. Thus Ext' outputs this same witness $\mathbb{w}$.

All that remains to be shown is the above claim. We give a full proof of this claim in Section 5. At a high level, we show the claim by reverse induction on $i \in [\mu]$. For $i = \mu$, if the claim does not hold, then clearly T is no longer a tree of accepting transcripts as there is some complete transcript that remains in the doomed set, violating the assumption that T is an accepting tree of transcripts. Then fixing $i < \mu$ and assuming the claim is true for all j such that $i < j \leq \mu$, suppose τ_{i-1} is a partial transcript in the tree T. If m_i^* is the unique node in T such that (τ_{i-1}, m_i^*) is a path in T, if for all outgoing edges $c_{i,1}, \ldots, c_{i,k}$ of m_i^* we have $(\tau_{i-1}, m_i^*, c_{i,j}) \notin \mathcal{D}$, then we are done. Otherwise, there exists some outgoing edge c_{i,j^*} of m_i^* such that $(\tau_{i-1}, m_i^*, c_{i,j^*}) \in \mathcal{D}$. Now if this happens, by our induction hypothesis, we have that for m_{i+1}^* which has incoming edge c_{i,j^*} and all outgoing edges $c_{i+1,j}$ from m_{i+1}^*, the partial transcript $(\tau_{i-1}, m_i^*, c_{i,j^*}, m_{i+1}^*, c_{i+1,j}) \notin \mathcal{D}$. This completes the induction step and the proof.

2.3.3 Extending to Generalized Round-by-round Knowledge and Generalized Special Soundness

The above argument again naturally extends to the generalized RBR knowledge and generalized special soundness cases. Indeed, taking $k_i = \lceil |C_i|\varepsilon_i \rceil + 1$ for RBR knowledge errors ε_i, the above argument holds by replacing all k-tree of transcripts with $(k_1, \ldots, k_\mu)$-tree of transcripts, and by considering the RBR knowledge error ε_i for any partial transcript τ_i. See Section 5 for full details.

2.4 Special Soundness Does Not Imply Round-by-round Knowledge

Roughly, one of our main results states that an IOP may be $(k_1, \ldots, k_\mu)$-special sound with small k_i's, but at the same time only have RBR knowledge with errors $(\varepsilon_1, \ldots, \varepsilon_\mu)$ if $\varepsilon_i = 1$ for some i.

We begin by providing intuition on why this is the case. Afterward we will give a more technical explanation. To this end, observe that the extractor $\mathsf{Ext}_{\mathsf{spec}}$ given by the definition of special soundness is given access to a tree of accepting *complete* transcripts, while the extractor $\mathsf{Ext}_{\mathsf{RBR}}$ from the RBR knowledge definition only receives partial transcripts as inputs. Thus, in a sense, $\mathsf{Ext}_{\mathsf{spec}}$ has more information to work with than $\mathsf{Ext}_{\mathsf{RBR}}$. More precisely, suppose an IOP Π is built in a way that accepting complete transcripts "contain full information" about a valid witness, but that no partial transcript contains such information. Then, given a tree of accepting transcripts, the extractor $\mathsf{Ext}_{\mathsf{spec}}$ is able to extract a witness from it, since the tree contains complete transcripts that include "full information" about a valid witness. However, $\mathsf{Ext}_{\mathsf{RBR}}$ is never given a complete accepting transcript, and instead only sees partial transcripts. Consequently, by our assumptions on Π, the extractor $\mathsf{Ext}_{\mathsf{RBR}}$ may be unable to reconstruct a witness from these partial transcripts, no matter how likely it is that the partial transcript leaves certain doomed set. Under this informally described scenario, Π would be special sound, but it would not have RBR knowledge soundness.

We next formalize the ideas above. For simplicity, we restrict ourselves to IOP's where the prover and verifier perform only one round of communication, i.e. $\mu(|\mathbb{x}|) = 1$ for all input $\mathbb{x}$. Our full result deals with IOP's with arbitrarily (polynomialy) many rounds of interaction.

Fix a relation $\mathbf{R}$ so that the language $\mathcal{L}_{\mathbf{R}}$ is in NP but not in P. Now we construct an IOP $\Pi = (\mathsf{P}, \mathsf{V})$ for $\mathbf{R}$ such that, as hinted above, its partial transcripts contain "no information" about witnesses, while its complete transcript do. To this end, for any $(\mathbb{x}, \mathbb{w}) \in \mathbf{R}$, we have P send the 0 bit as its first message m_1. Then we have V reply with a random string c, to which the honest prover P replies by sending the entire witness $\mathbb{w}$. Then V accepts the proof if and only if $(\mathbb{x}, \mathbb{w}) \in \mathbf{R}$ and $m_1 = 0$.

Clearly, Π is (1)-special sound, since any complete accepting transcript for $\mathbb{x}$ contains a witness $\mathbb{w}$. On the other hand, all non-complete partial transcripts for Π are of the form $(\mathbb{x})$, $(\mathbb{x}, 0)$, or $(\mathbb{x}, 0, c)$. Neither of these "contains information" about a witness for $\mathbb{x}$ (other than the information provided by the input $\mathbb{x}$ itself).

Now assume Π has RBR knowledge with error ε_1, and let $\mathcal{D}$ and Ext be a corresponding doomed set and an extractor algorithm. Then, by definition of RBR knowledge, whenever we have

$$\Pr_c[(\mathbb{x}, 0, c) \notin \mathcal{D}] > \varepsilon_1, \tag{2}$$

the extractor Ext is able to find a valid witness for $\mathbb{x}$ just from seeing $(\mathbb{x}, 0)$. However, if $\mathbb{x} \in \mathcal{L}_{\mathbf{R}}$, then the probability $\Pr_c[(\mathbb{x}, 0, c) \notin \mathcal{D}]$ must be 1, because for all c, the partial transcript $(\mathbb{x}, 0, c)$, which comprises all μ rounds of interaction, can be extended into a complete accepting transcript $(\mathbb{x}, 0, c, m)$. Then, by definition of doomed set, we must have $(\mathbb{x}, 0, c) \notin \mathcal{D}$. It follows from this and Eq. (2), that if $\varepsilon_1(|\mathbb{x}|) < 1$, then the extractor Ext is able to output a

witness for $\mathbb{x}$ just from seeing $(\mathbb{x}, 0)$. With these arguments in mind, and assuming $\varepsilon_1(|\mathbb{x}|) < 1$, it is straightforward to build a deterministic polynomial time algorithm that recognizes the language $\mathcal{L}_{\mathbf{R}}$, contradicting our assumption that $\mathcal{L}_{\mathbf{R}}$ is in NP but not in P. Thus, Π cannot have RBR knowledge soundness, i.e. if it has RBR knowledge, then the error ε_1 is not negligible —in fact, it is 1.

In Section 5 we formulate these ideas in full formality. Moreover, instead of restricting ourselves the 1-round IOPs, we generalize our arguments to construct a μ-round IOP that is special sound but does not have RBR knowledge soundness, for any polynomial $\mu = \mu(|\mathbb{x}|)$.

2.5 Special Unsoundness and Round-by-round Soundness are Dual

Finally, we relate round-by-round soundness and the notion of *special unsoundness* [AFK22]. Informally, an IOP $\Pi = (P, V)$ is *ℓ-special unsound* if there exists a dishonest prover strategy P^* such that during any round of the protocol, for any message m sent by P^*, there exists a "lucky" set of verifier challenges $L \subset C$ such that $|L| = \ell$ such that if the verifier V responds with $c \in L$, then P^* can "behave honestly" for the remainder of the protocol and V will accept at the end of the protocol execution.

Given the above notion of special unsoundness, it is immediate on an intuitive level that special unsoundness and RBR soundness are "dual" notions: RBR soundness states that for any dishonest prover strategy, the probability the prover gets "lucky" is upper bounded by some ε, whereas special unsoundness says that there exists a prover strategy where the probability the prover gets "lucky" is lower bounded by some ε'. We formalize this relationship in Section 6. Assume that Π is an ε-sound μ-round IOP, and assume that Π is generalzied round-by-round sound with errors $(\epsilon_1, \ldots, \epsilon_\mu)$ and special unsound with errors $(\epsilon'_1, \ldots, \epsilon'_\mu)$. Then for $\epsilon = 1 - \prod_i (1 - \epsilon_i)$ and $\epsilon' = 1 - \prod_i (1 - \epsilon'_i)$, it holds that $\epsilon' \leq \varepsilon \leq \epsilon$. Moreover, we show that if $\epsilon'_i = \epsilon_i$ for all i, then we have $\varepsilon = \epsilon$. See Section 6 for complete details.

3 Preliminaries

A *relation* $\mathbf{R}$ is a subset of pairs $(\mathbb{x}; \mathbb{w}) \in \{0,1\}^* \times \{0,1\}^*$. The strings $\mathbb{x}$ are called *inputs* (these are often called also *statements* or *instances*), and the strings $\mathbb{w}$ are called *witnesses*. To each relation $\mathbf{R}$ there corresponds a language $L_{\mathbf{R}} \subseteq \{0,1\}^*$ consisting of all statements $\mathbb{x}$ such that $(\mathbb{x}, \mathbb{w}) \in \mathbf{R}$ for some $\mathbb{w}$. When $(\mathbb{x}, \mathbb{w}) \in \mathbf{R}$, we say that $\mathbb{w}$ is a *valid witness for* $\mathbb{x}$. We assume our relations to be in the class NP.

We parameterize our security functions either by the length $|\mathbb{x}|$ of an input, but for ease of exposition we omit these from our notation, i.e. we write expressions such as "soundness error ε" instead of "soundness error $\varepsilon(|\mathbb{x}|)$". We proceed similarly for other types of functions.

We denote by $\mathbb{N}$ be the set of all positive non-zero integers. For any $m \in \mathbb{N}$, we let $[m]$ denote the set $\{1, \ldots, m\}$. For any finite set S, we let $s \xleftarrow{\$} S$ denote the process of sampling an element of S uniformly and independently at random.

3.1 Interactive Oracle Proofs

Given a map $f \in A^B$ for some sets A, B, we denote by $[\![f]\!]$ an *oracle* to the map f. This is a hypothetical algorithm that takes elements $a \in A$ as input, and outputs $f(a)$ instantaneously.

Definition 3.1 (Interactive Proofs (IP))**.** *A* μ-round interactive proof *for a relation* $\mathbf{R}$ *is a pair of interactive algorithms* $\Pi = (\mathsf{P}, \mathsf{V})$ *such that:*

- *For* $\mathbb{x} \in L_{\mathbf{R}}$ *and* $\mathbb{w}$ *such that* $(\mathbb{x}, \mathbb{w}) \in \mathbf{R}$*, before the start of the protocol,* P *receives both* $(\mathbb{x}, \mathbb{w})$ *as input and* V *receives* $\mathbb{x}$ *as input.*
- $\mathsf{P}(\mathbb{x}, \mathbb{w})$ *and* $\mathsf{V}(\mathbb{x})$ *exchange* $2\mu(|\mathbb{x}|) + 1$ *messages, where* P *sends the first and last message, and during any round of interaction* P *sends message* m_i *to* V*. After* P *sends* $m_{\mu(|\mathbb{x}|)+1}$*,* V *either accepts or rejects.*

We require the following properties to hold:

- ***Completeness:*** *for all* $(\mathbb{x}, \mathbb{w}) \in \mathbf{R}$*, we have*

$$\Pr\left[\langle \mathsf{P}(\mathbb{w}), \mathsf{V}\rangle(\mathbb{x}) = \mathsf{accept}\right] = 1,$$

  where $\langle \mathsf{P}(\mathbb{w}), \mathsf{V}\rangle(\mathbb{x})$ *denotes the output of* P *and* V *interacting on common input* $\mathbb{x}$ *where* P *is additionally given* $\mathbb{w}$ *as input, and the above probability holds over the random coins of* V*.*

- ϵ***-Soundness:*** *for any* $\mathbb{x} \notin L_{\mathbf{R}}$ *and any unbounded interactive algorithm* P^**, we have*

$$\Pr\left[\langle \mathsf{P}^*, \mathsf{V}\rangle(\mathbb{x}) = \mathsf{accept}\right] \leq \epsilon,$$

  where the probability is taken over the random coins of V*.*

We say that Π *is* public-coin *if all messages sent by* V *are independent uniform random strings of some bounded length and the output of* V *does not depend on any secret state.*

Remark 3.2. In this paper, all IP's and IOP's are assumed to be public-coin. As seen above, μ is a function that depends on $|\mathbb{x}|$, however we omit explicitly referring to this dependence, writing μ to refer both to the function and to the value $\mu(|\mathbb{x}|)$.

Definition 3.3 (Interactive Oracle Proof)**.** *An Interactive Oracle Proof (IOP) for a relation* $\mathbf{R}$ *is a* μ*-round IP* (P, V) *for* $\mathbf{R}$ *in which, for all* $\mathbb{x}$*, at each round of interaction* $i \in [\mu(\mathbb{x})]$*,* P *sends* m_i *and* V *receives oracle access to* m_i *via* $[\![m_i]\!]$*. Crucially, at the end of the interactive phase,* V *does not necessarily need to read the whole* m_i *in order to decide whether to accept or reject.*

Definition 3.4 (Message Spaces and i-round Partial Transcripts)**.** *Let* Π *be a* μ*-round IOP* $\Pi = (\mathsf{P}, \mathsf{V})$ *for a relation* $\mathbf{R}$*. We denote by* $\mathcal{M}_1, \ldots, \mathcal{M}_{\mu+1}$ *the sets of all potential prover's messages, so that, at round* $i \in [\mu]$*,* P *sends a message from* $\mathcal{M}_i$*. The set* $\mathcal{M}_{\mu+1}$ *constitutes the set of all potential prover's last messges. Similarly, we let* $C_1, \ldots, C_\mu$ *be the sets of all potential verifier's messages, which we refer to as* challenges*, so that at Round* $i \in [\mu]$ V *replies with a challenge from* C_i*. These sets do not depend on* $\mathbb{x}$ *or any message/challenge previously exchanged between the prover and the verifier.*

Given $i \in [\mu]$*, we define a* i-round partial transcript *as a vector of the form* $\tau = (\mathbb{x}, m_1, c_1, \ldots, m_i, c_i)$ *where* $m_i \in \mathcal{M}_j$ *and* $c_i \in C_j$ *for all* $j \in [i]$*. We also let a* 0-round partial transcript *be a "vector" of the form* $(\mathbb{x})$*. We write* $PartTr(i)$ *to denote the set of all* i*-round partial transcripts. A* complete transcript *is a transcript of the form* $(\mathbb{x}, \tau, m)$ *where* $(\mathbb{x}, \tau) \in PartTr(\mu)$ *and* $m \in \mathcal{M}_{\mu+1}$*. Such a transcript is* accepting *if* $\mathsf{V}(\mathbb{x}, \tau, m) = \mathsf{accept}$*.*

3.2 Special Soundness

Let P be a potentially dishonest prover for Π, let $\mathbb{x}$ be a statement for a relation $\mathbf{R}$, and let $\tau \in \mathrm{PartTr}(i)$. We write $\mathsf{P}(\mathbb{x}, \tau)$ to denote the state of P at the beginning of Round $i+1$ if $(\mathbb{x}, \tau)$ is the transcript so far.

Definition 3.5 (Tree of Transcripts)**.** *Let* $\Pi = (\mathsf{P}, \mathsf{V})$ *be a* μ*-round IOP. Let* $(k_1, \ldots, k_\mu) \in \mathbb{N}^\mu$*. A* $(k_1, \ldots, k_\mu)$-tree of transcripts *for* $\mathbb{x}$ *is a set of* $k = \prod_i k_i$ *complete transcripts* $(\tau_1, \ldots, \tau_k)$ *with common first message* m*, arranged in a tree of depth* $\mu + 1$ [3] *and arity* $k_1, \ldots, k_\mu$*, respectively. The nodes in the tree correspond to the prover's messages, and the edges correspond to the verifier's challenges. Every internal node at depth* $i - 1$ *(*$1 \leq i \leq \mu$*) has* k_i *children with distinct challenges. Every* τ_j*,* $j \in [k]$ *corresponds to one path from the root to the leaf node.*

Finally, we say that the tree is a $(k_1, \ldots, k_\mu)$-tree of accepting transcripts *for* $\mathbb{x}$ *if every transcript is accepted by* V*.*

Definition 3.6 (Special Soundness)**.** *Let* Π *be a* μ*-round IOP for a relation* $\mathbf{R}$*, and let* $(k_1, \ldots, k_\mu) \in \mathbb{N}^\mu$*. We say* Π *is* $(k_1, \ldots, k_\mu)$-special sound *if there exists a polynomial time algorithm* Ext *that, on input* $\mathbb{x}$ *and any* $(k_1, \ldots, k_\mu)$*-tree of accepting transcripts for* $\mathbb{x}$*, outputs a witness* $\mathbb{w}$ *such that* $(\mathbb{x}, \mathbb{w}) \in \mathbf{R}$*.*

[3] We set the root vertex of a tree to be of depth 1 (as opposed to 0).

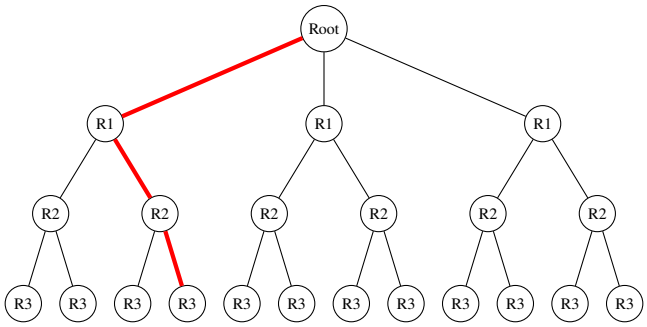

Figure 2: 3-round IOP tree as a $(3, 2, 3)$-tree of transcripts with a highlighted complete/accepting transcript.

3.3 Special Unsoundness

The notion of special unsoundness was introduced in [AFK22] in the context of analysing the security of the Fiat-Shamir transformation of the parallel repetition of an interactive proof. The authors describe an attack to such protocol and analyse its security. Later in Section 6 we will showcase the interest of this definition in a more general context, unrelated to the parallel repetition of IP's.

Definition 3.7 (Special Unsoundness [AFK22])**.** *Let Π be a μ-round IOP, and let $(\ell_1, \ldots, \ell_\mu) \in \mathbb{N}^\mu$. We say that Π has* $(\ell_1, \ldots, \ell_\mu)$-special unsoundness *if there exists a dishonest prover $\mathcal{A}$ of the following form and so that in the execution with V and input $\mathbb{x}$ the following holds:*

- *$\mathcal{A}$ starts off in* active mode*, which is so that in every round i, when $\mathcal{A}$ sends the message m_i, there exists a subset $\mathcal{L}_i \subseteq C_i$ such that $|\mathcal{L}_i| = \ell_i$ (defined as a function of the state of $\mathcal{A}$ at shit point) such that if the subsequent challenge c_i is in $\mathcal{L}_i$, then $\mathcal{A}$ switches into* passive mode*.*
- *If $\mathcal{A}$ switches into passive mode, then it remains in passive mode until the end of the protocol, and V accepts at the end of the protocol.*

4 Generalized Special Soundness and Unsoundness, and Round-by-round Soundness and Knowledge

In this section, we define the notions of *generalized round-by-round soundness* and *knowledge*, respectively. These are essentially the original definitions from [CCH+19] and [CMS19] with the modification that, instead of having a "one-size-fits-all-rounds" soundness/knowledge error, we consider an error for each round. Following [Hol19], we use a formalism based on "doomed" sets of partial transcripts, as opposed to using a "state function". We will often abbreviate the expression "round-by-round" as RBR.

Definition 4.1 ((Generalized) Round-by-round Soundness)**.** *An IOP $\Pi = (\mathsf{P}, \mathsf{V})$ for a relation $\mathbf{R}$ has* (generalized) round-by-round soundness *with errors $(\varepsilon_1, \ldots, \varepsilon_\mu)$ if there exists a (not necessarily efficiently computable) "doomed set" $\mathcal{D}$ of partial transcripts such that the following properties hold:*

1. *If $\mathbb{x} \notin \mathcal{L}_{\mathbf{R}_i}$, then $(\mathbb{x}) \in \mathcal{D}$.*
2. *For any $\mathbb{x}$ and any μ-round partial transcript $(\mathbb{x}, \tau) \in \mathit{PartTr}(\mu)$ and any last prover message $m \in \mathcal{M}_{\mu+1}$, if $(\mathbb{x}, \tau) \in \mathcal{D}$ then $V(\mathbb{x}, \tau, m) = \mathsf{reject}$.*
3. *If $(\mathbb{x}, \tau) \in \mathcal{D}$ and $(\mathbb{x}, \tau) \in \mathit{PartTr}(i-1)$ is a $(i-1)$-round partial transcript for some $i \in [\mu]$, then for all $m \in \mathcal{M}_i$ we have*

$$\Pr_{c \xleftarrow{\$} C_i}[(\tau, m, c) \notin \mathcal{D}] \leq \varepsilon_i.$$

If for all $i \in [\mu]$, $\varepsilon_i = \varepsilon_i(|\mathbb{x}|)$ is a negligible function of the input length $|\mathbb{x}|$, then we simply say that Π has RBR soundness.

The original definition of round-by-round soundness considers only the scenario in which $\varepsilon_1 = \ldots = \varepsilon_\mu$, and in that case one says that Π has *RBR soundness with error ε*, where $\varepsilon = \varepsilon_i$ for all i.

Definition 4.2 ((Generalized) Round-by-round Knowledge)**.** *We say an IOP $\Pi = (\mathsf{P}, \mathsf{V})$ for a relation $\mathbf{R}$ has* (generalized) round-by-round knowledge *with errors $(\varepsilon_1, \ldots, \varepsilon_\mu)$ if there exists a (not necessarily efficiently computable) "doomed set" $\mathcal{D}$ of partial transcripts such that the following properties hold:*

- *For all possible inputs $\mathbb{x}$, not necessarily in $\mathcal{L}_{\mathbf{R}}$, we have $(x) \in \mathcal{D}$.*
- *For any μ-round partial transcript $(\mathbb{x}, \tau) \in PartTr(\mu)$ and any last prover message $m \in \mathcal{M}_{\mu+1}$, if $(\mathbb{x}, \tau) \in \mathcal{D}$, then $\mathsf{V}(\mathbb{x}, \tau, m) =$ reject.*
- *There exists a polynomial time algorithm* Ext*, called* extractor*, with the following properties. If $(\mathbb{x}, \tau) \in \mathcal{D}$ and $(\mathbb{x}, \tau) \in PartTr(i-1)$ for some $i \in [\mu]$, and for all $m \in \mathcal{M}_i$ we have*

$$\Pr_{c \xleftarrow{\$} C_i}[(\mathbb{x}, \tau, m, c) \notin \mathcal{D}] > \varepsilon_i,$$

  then $\mathsf{Ext}(\mathbb{x}, \tau, m)$ outputs a witness $\mathbb{w}$ such that $(\mathbb{x}, \mathbb{w}) \in \mathbf{R}$.

  In this case, we say that the partial transcript $(\mathbb{x}, \tau, m)$ above is an RBR extractable *partial transcript.*

If for all $i \in [\mu]$, $\varepsilon_i = \varepsilon_i(|\mathbb{x}|)$ is a negligible function of the input length $|\mathbb{x}|$, then we simply say that Π has RBR knowledge soundness.

The original definitions of RBR soundness and knowledge consider only the case in which the errors are the same for each round, i.e. $\varepsilon_1 = \ldots = \varepsilon_\mu$.

Remark 4.3. Often in this paper we will drop the term "generalized" and talk simply of RBR soundness and RBR knowledge, referring always to the generalized definitions provided above.

Remark 4.4 (RBR Knowledge Implies RBR Soundness). An IOP with RBR knowledge with errors $(\varepsilon_1, \ldots, \varepsilon_\mu)$ is necessarily RBR sound with errors $(\varepsilon_1, \ldots, \varepsilon_\mu)$. This is because if $(\mathbb{x}, \tau, m)$ is an RBR extractable partial transcript, then the extractor from the definition of RBR knowledge outputs a witness $\mathbb{w}$ such that $(\mathbb{x}, \mathbb{w}) \in \mathbf{R}$, and so $\mathbb{x} \in \mathcal{L}_{\mathbf{R}}$. Hence, if $\mathcal{D}$ is the doomed set with which Π has RBR knowledge, the subset $\mathcal{D}' \subseteq \mathcal{D}$ consisting of all $(\mathbb{x}, \tau) \in \mathcal{D}$ such that $\mathbb{x} \notin \mathcal{L}_{\mathbf{R}}$ is a doomed set with which Π has RBR soundness with errors $(\varepsilon_1, \ldots, \varepsilon_\mu)$.

The following remark highlights the relation between the generalized and non-generalized versions of round-by-round soundness.

Remark 4.5. Let Π be an IOP with μ rounds. Suppose Π is round-by-round sound with errors $(\varepsilon_1, \ldots, \varepsilon_\mu)$. Then Π is round-by-round sound with error ε), where

$$\varepsilon = \max_{i \in [\mu]}\{\varepsilon_i\}.$$

Conversely, if Π is round-by-round sound with error ε, then it is round-by-round sound with errors $(\varepsilon_1, \ldots, \varepsilon_\mu)$ where $\varepsilon_i = \varepsilon$ for all $i \in [\mu]$.

5 Relating Special Soundness and Round-by-round Knowledge

We begin this section by stating our main results, and proof each of them afterwards. The first one states that special soundness implies RBR soundness.

Theorem 1.1 (Special Soundness Implies RBR Soundness)**.** *Let $\Pi = (\mathsf{P}, \mathsf{V})$ be a μ-round IOP for a relation $\mathbf{R}$. Let C_i be the set of verifier challenges for round $i \in \{1, \ldots, \mu\}$, and let $(k_1, \ldots, k_\mu) \in \mathbb{N}^\mu$. Assume that Π is $(k_1, \ldots, k_\mu)$-special sound. Then Π is RBR sound with errors*

$$\left(\frac{k_1 - 1}{|C_1|}, \ldots, \frac{k_\mu - 1}{|C_\mu|}\right). \tag{1}$$

As a corollary of Theorem 1.1 and [CMS19], we see that special sound IOPs can be compiled to secure non-interactive proofs in the quantum random oracle model via the BCS transformation [BCS16].

Corollary 1.6 (Special Soundness Implies FS Security in the QROM)**.** *Let Π be a μ-round $(k_1, \ldots, k_\mu)$-special sound IOP. Let $\mathsf{BCS}(\Pi)$ be the non-interactive proof obtained by applying the BCS transformation to Π, and let $\varepsilon = \max_{i\in[\mu]}\{(k_i - 1)/|C_i|\}$. Then $\mathsf{BCS}(\Pi)$ has adaptive soundness error $O(t^2\varepsilon + t^3/2^\lambda)$ against quantum attackers that make at most $t - O(q \log \ell)$ queries to the random oracle, where λ is the output length of the random oracle in bits, q is (an upper bound on) the total number of queries made by the verifier during any execution of Π, and ℓ is the total number of symbols sent by both the prover and verifier during any execution of Π.*

The next result states that RBR knowledge implies special soundness.

Theorem 1.2 (RBR Knowledge Implies Special Soundness)**.** *Let $\Pi = (\mathsf{P}, \mathsf{V})$ be a μ-round IOP for a relation $\mathbf{R}$. Let C_i be the set of verifier challenges for round $i \in \{1, \ldots, \mu\}$. Assume Π has round-by-round knowledge with errors $\varepsilon_1, \ldots, \varepsilon_\mu$, and let*

$$(k_1, \ldots, k_\mu) = (\lceil |C_1|\varepsilon_1 \rceil + 1, \ldots, \lceil |C_\mu|\varepsilon_\mu \rceil + 1).$$

Suppose $\sum_{i\in[\mu]} \prod_{j\in[i]} k_j$ is upper bounded by a polynomial (on the lengths of inputs). Then Π is $(k_1, \ldots, k_\mu)$-special sound.

Finally, the third result shows that under a mild assumption, in general, special soundness can only imply a non-interesting notion of RBR knowledge.

We can now state the third result of this section.

Theorem 1.3 (Special Soundness does not Imply RBR Knowledge)**.** *Assume $\mathsf{NP} \neq \mathsf{P}$. Then for any polynomial function $\mu(|\mathbb{x}|)$, there exists a μ-round IOP Π with the following properties:*

- *Π is $(1, \overset{\mu}{\ldots}, 1)$-special sound.*
- *If Π is RBR knowledge sound with errors $(\varepsilon_1, \ldots, \varepsilon_\mu)$, then $\varepsilon_i(\ell) = 1$ for some input length ℓ and some $i \in [\mu]$.*

We proceed to prove each of the results above. Before proving Theorem 1.1 we introduce some terminology.

Definition 5.1 (Complete paths, rooted paths, and identification of paths with their label)**.** *First, we say a path in a tree is* complete *if it starts at the root node and it ends in a leaf. We say a path is* rooted *if it start at the root node. All complete paths are rooted.*

For ease of presentation we introduce the following abuse of notation and terminology: we identify rooted paths with their labels, which are partial transcript.[4]

Definition 5.2 ($(k_i, \ldots, k_\mu)$-tree of Transcripts)**.** *Let Π be a μ-round IOP, let $(k_1, \ldots, k_\mu) \in \mathbb{N}^\mu$, and let $i \in [\mu]$. By a $(k_i, \ldots, k_\mu)$-tree of transcripts we refer to a subtree T' of a $(k_1, \ldots, k_\mu)$-tree T of transcripts such that the root node of T' is at depth i in T. In other words, T' is a tree of depth $\mu - i + 1$, where for depths $1, \ldots, \mu - i + 1$, nodes have $k_i, \ldots, k_\mu$ children, respectively. Each complete path in T' corresponds to a suffix of a complete transcript. Note that if $i = 1$ then we recover the original definition of $(k_1, \ldots, k_\mu)$-tree of transcripts.*

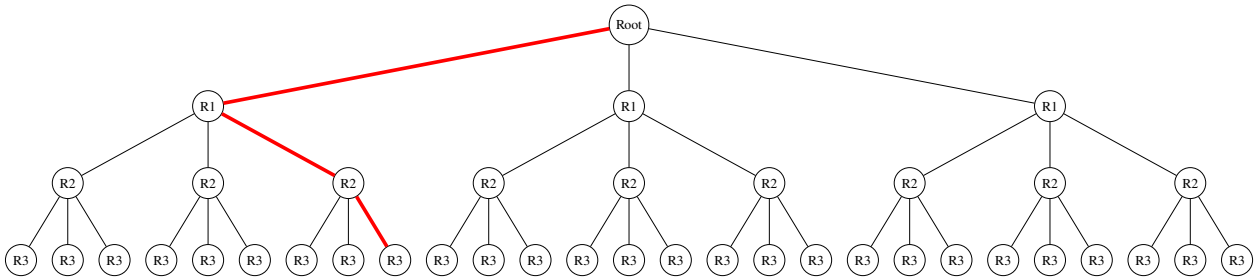

Figure 3: $(3, 3, 3)$-tree of transcripts with a highlighted complete transcript.

[4]Technically, this is ambiguous because two distinct rooted paths could, theoretically, have the same label. However in our context there will never be risk of confusion.

Recall that we use $\mathrm{PartTr}(i)$ to denote the set of i-round partial transcripts of an IOP.

Definition 5.3 (Completable Transcripts)**.** *Following the notation above, let $i = 1, \ldots, \mu+1$, and let $(\mathbb{x}, \tau) \in \mathit{PartTr}(i-1)$. We say $(\mathbb{x}, \tau)$ is* completable *if one of the following holds:*

- *$i \in [\mu]$ and there exists a $(k_i, \ldots, k_\mu)$-tree of transcripts T' such that, for all complete path τ' in T', the transcript $(\mathbb{x}, \tau, \tau')$ is accepted by the verifier (following our convention from Definition 5.1, here we identify the path τ' with the sequence of messages and challenges associated to each of its nodes and edges (i.e. its label)). We say that T'* completes *$(\mathbb{x}, \tau)$.*
- *$i = \mu$ and there exists $m \in \mathcal{M}_{\mu+1}$ such that the complete transcript $(\mathbb{x}, \tau, m)$ is accepted by the verifier.*

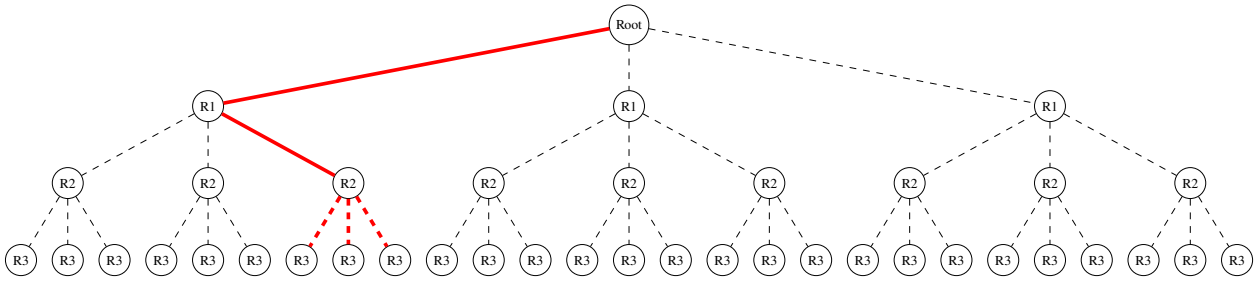

Figure 4: 3-round IOP tree with a highlighted completable partial transcript. In this figure, the highlighted path through the tree ends in the second round with node R2, and the hanging tree (with dashed red edges) at R2 represents the tree that completes the partial transcript.

Recall that, given a IOP with RBR knowledge $(\varepsilon_1, \ldots, \varepsilon_\mu)$, by a *RBR extractable transcript* we refer to a partial transcript $(\mathbb{x}, \tau, m)$ where $(\mathbb{x}, \tau) \in \mathrm{PartTr}(i-1)$ and $\Pr_{c_i \xleftarrow{\$} C_i}[(\mathbb{x}, \tau, m) \notin \mathcal{D}] > \varepsilon_i$.

Proof of Theorem 1.1. Assume Π is $(k_1, \ldots, k_\mu)$-special sound. We define a doomed set $\mathcal{D}$ of partial transcripts as the union $\mathcal{D} = \bigcup_{i=0}^{\mu} \mathcal{D}_i$, where each $\mathcal{D}_i \subseteq \mathrm{PartTr}(i)$, i.e. $\mathcal{D}_i$ consists of i-round partial transcripts. These sets $\mathcal{D}_i$ are defined as follows.

$$\begin{aligned} \mathcal{D}_0 &:= \{(\mathbb{x}) \mid \mathbb{x} \notin \mathcal{L}_{\mathbf{R}}\}, \\ \mathcal{D}_i &:= \{(\mathbb{x}, \tau) \in \mathrm{PartTr}(i) \mid (\mathbb{x}, \tau) \text{ is not completable}\}. \end{aligned}$$

We now prove that Π has RBR soundness with errors $\varepsilon_i = (k_i - 1)/|C_i|$ for all $i \in [\mu]$. Let Ext be the extractor from the definition of special soundness, and fix an input $\mathbb{x}$. Let $i =\in [\mu]$, and take a doomed $(i-1)$-round partial transcript $(\mathbb{x}, \tau) \in \mathrm{PartTr}(i-1) \cap \mathcal{D} = \mathcal{D}_{i-1}$. In particular, either $i = 1$ and $(\mathbb{x}, \tau) = (\mathbb{x}) \in \mathcal{D}_0$, or $(\mathbb{x}, \tau) \in \mathcal{D}_i$ for $i \geq 1$, and then by definition $(\mathbb{x}, \tau)$ is not completable. Given $m \in \mathcal{M}_i$, denote

$$P_i(\mathbb{x}, \tau, m) := \Pr_{c \xleftarrow{\$} C_i}[(\mathbb{x}, \tau, m, c) \notin \mathcal{D}_i].$$

Assume $P_i(\mathbb{x}, \tau, m) > \varepsilon_i$. Then there are at least k_i distinct challenges $c_1, \ldots, c_{k_i}$ such that $(\mathbb{x}, \tau, m, c_j) \notin \mathcal{D}_i$ for all $j \in [k_i]$. Thus each $(\mathbb{x}, \tau, m, c_j)$ is completable. We claim that then also $(\mathbb{x}, \tau)$ is completable. Indeed, let $T_1, \ldots, T_{k_i}$ be each a $(k_{i+1}, \ldots, k_\mu)$-tree of transcripts such that each T_j completes the i-round partial transcript $(\mathbb{x}, \tau, m, c_j)$, for all $j \in [k_i]$. By definition, for each complete path τ_j in T_j, the complete transcript $(\mathbb{x}, \tau, m, c_j, \tau_j)$ is accepted by the verifier. Now construct a $(k_i, \ldots, k_\mu)$-tree T of transcripts as follows: Create a a root node V_R with label m, and create k_i childs $V_{R_1}, \ldots, V_{R_{k_i}}$ for the root node V_R, labeling the edges with the challenges $c_1, \ldots, c_{k_i}$, respectively. Now, attach the trees $T_1, \ldots, T_{k_i}$ to the child vertices $V_{R_1}, \ldots, V_{R_{k_1}}$. The resulting tree T is a $(k_i, \ldots, k_\mu)$-tree of transcripts with root labeled as m. Moreover, each complete path in T has label of the form (m, c_j, τ') for some $j \in [k_i]$. Hence, T completes $(\mathbb{x}, \tau)$ since $(\mathbb{x}, \tau, m, c_j, \tau')$ is a complete transcript that is accepted by the verifier, for all choices of m, c_j and τ'. The claim is proved.

As a consequence if $i \geq 2$ then $P_i(\mathbb{x}, \tau, m)$ cannot be strictly larger than ε_i if $(\mathbb{x}, \tau) \in \mathcal{D}_{i-1}$. Now assume that $i = 1$, so that $(\mathbb{x}, \tau) = (\mathbb{x}) \in \mathrm{PartTr}(0)$ is a 0-round partial transcript, and assume $(\mathbb{x}) \in \mathcal{D}_0$. Suppose towards contradiction that $P_1(\mathbb{x}, m) > \varepsilon_1$ for some $m \in \mathcal{M}_1$. Then the tree T with root m that completes $(\mathbb{x})$ (obtained through the argument above) is a $(k_1, \ldots, k_\mu)$-tree of accepting transcripts for $\mathbb{x}$. Given $\mathbb{x}$ and T as input, the "special soundness extractor" Ext outputs a valid witness for $\mathbb{x}$, contradicting the fact that $\mathcal{D}_0$ consists precisely of those $(\mathbb{x})$ such that $\mathbb{x} \notin \mathcal{L}_{\mathrm{R}}$. Finally, notice that by definition, for any $(\mathbb{x}, \tau) \in \mathrm{PartTr}(\mu)$ with $(\mathbb{x}, \tau) \in \mathcal{D}$, and any $m \in \mathcal{M}_{\mu+1}$, we have that the verifier rejects $(\mathbb{x}, \tau, m)$. This proves that Π is RBR sound with the claimed errors. □

Proof of Corollary 1.6. This is a direct consequence of Theorem 1.1 and of Theorem 8.6 from [CMS19]. The former yields that any $(k_1, \ldots, k_\mu)$-special sound IOP has RBR soundness (in the usual, not generalized sense) $\leq \max_{i \in [\mu]}\{(k_i - 1)/|\mathcal{D}_i|\}$. The latter states that the BCS transformation of a RBR sound IOP is sound in the Quantum Random Oracle Model, with the parameters and soundness errors from the statement of the corollaries. □

Next, we prove Theorem 1.2.

Proof of Theorem 1.2. Assume Π has RBR knowledge with errors $(\varepsilon_1, \ldots, \varepsilon_\mu)$. Let $\mathcal{D}$ be a corresponding doomed set. Let $(k_1, \ldots, k_\mu) = (\lceil |C_1|\varepsilon_1 + 1 \rceil, \ldots, \lceil |C_\mu|\varepsilon_\mu + 1 \rceil)$, fix an input $\mathbb{x}$, and let T be a $(k_1, \ldots, k_\mu)$-tree of accepting transcripts for $\mathbb{x}$.

For convenience we introduce the following notation: let τ be a rooted path in T (i.e. a path starting at the root of T) with $(\mathbb{x}, \tau) \in \mathrm{PartTr}(i)$ for some $i = 0, \ldots, \mu$. Then, recall from the definition of i-round partial transcripts, τ ends with a challenge c from C_i. As such, the path τ has a uniquely defined node right after c in the tree T. We denote it by $\mathsf{EndNode}(\tau)$.

We claim there is a rooted path τ in T with the following properties:

- $(\mathbb{x}, \tau) \in \mathrm{PartTr}(i-1)$ and $(\mathbb{x}, \tau) \in \mathcal{D}$ for some $i \in [\mu]$.
- For each of the k_i edges with challenges $(c_{i,j})_{j \in [k_i]}$ departing from node $\mathsf{EndNode}(\tau)$, we have

$$(\mathbb{x}, \tau, \mathsf{EndNode}(\tau), c_{i,j}) \notin \mathcal{D}.$$

Observe that for such a path, the transcript $(\mathbb{x}, \tau, \mathsf{EndNode}(\tau))$ is RBR extractable since the fraction of challenges $c \in C_i$ such that $(\mathbb{x}, \tau, \mathsf{EndNode}(\tau), c) \notin \mathcal{D}$ is at least $k_i/|C_i| > (k_i - 1)/|C_i| = \varepsilon_i$. Accordingly, we also call the path *RBR extractable* (as we identify paths with labels).

We now prove the theorem under the assumption that the claim above is true. To do so, we define an extractor $\mathsf{Ext}_{\mathsf{spec}}$ for the special soundness of Π in the following way: given an input $\mathbb{x}$ and a $(k_1, \ldots, k_\mu)$-tree T' of accepting transcripts for $\mathbb{x}$, the extractor $\mathsf{Ext}_{\mathsf{spec}}$ enumerates all paths in T' of the form (τ, m) where $(\mathbb{x}, \tau) \in \mathrm{PartTr}(i)$ and $m = \mathsf{EndNode}(\tau)$ for some $i = 0, \ldots, \mu$. Then it successively runs the RBR knowledge extractor $\mathsf{Ext}_{\mathsf{rbr}}$ on all the transcripts $(\mathbb{x}, \tau, m)$. If $\mathsf{Ext}_{\mathsf{rbr}}$ eventually outputs a witness $\mathbb{w}$, then $\mathsf{Ext}_{\mathsf{spec}}$ outputs $\mathbb{w}$, otherwise it outputs $\perp$.

Assuming the claim above is true, we know that at least one of the transcripts $(\mathbb{x}, \tau, m)$ is RBR extractable, and so $\mathsf{Ext}_{\mathsf{spec}}$ eventually outputs a valid witness. Finally, the total number of transcripts to inspect is exactly the number of vertices in T, which is exactly $1 + \sum_{i \in [\mu]} \prod_{j \in [i]} k_j$. By hypothesis, this is bounded by a polynomial in $|\mathbb{x}|$, and so $\mathsf{Ext}_{\mathsf{spec}}$ runs in polynomial time, as needed. We conclude that, once our previous claim is proved, the theorem will also be proved.

For technical reasons we prove a more general claim. Intuitively, the claim says that if we are given a rooted path τ in the tree (not necessarily a complete path), then the RBR extractable rooted path we are looking for can be found "by continuing τ". Formally, our new claim says that if τ is a rooted path in T and $(\mathbb{x}, \tau) \in \mathrm{PartTr}(i-1)$, then there is a RBR extractable rooted path in T of the form (τ, τ', m), where $(\mathbb{x}, \tau, \tau') \in \mathrm{PartTr}(j-1)$ for some $j \geq i$ (if $j = i$, then $\tau' = \emptyset$), and $m = \mathsf{EndNode}(\tau, \tau')$. Of course, if this claim is true, then the original one is true as well.

To prove the claim we proceed by reverse induction on $i \in [\mu]$. Let $(\mathbb{x}, \tau)$ be as in the induction hypotheses.

Suppose $i = \mu$. Let $c_{\mu,1}, \ldots, c_{\mu,k_\mu}$ be all edges leaving from $m = \mathsf{EndNode}(\tau)$. Notice that, for all $t \in [k_\mu]$, the rooted path $(\tau, m, c_{\mu,t})$ is a μ-round partial transcript, and thus $(\mathbb{x}, \tau, m, c_{\mu,t}) \notin \mathcal{D}$, as otherwise the verifier would reject $(\mathbb{x}, \tau, m, c_{\mu,t}, m')$, where $m' = \mathsf{EndNode}(\tau, m, c_{\mu,t}) \in \mathcal{M}_{\mu+1}$, contradicting that T is a tree of accepting transcripts. Hence $(\mathbb{x}, \tau, m, c_{\mu,t}) \notin \mathcal{D}$ for all $t \in [k_\mu]$.

Now fix $i \in [\mu - 1]$ and assume the claim is true for all j with $i < j \leq \mu$. Let $m = \mathsf{EndNode}(\tau) \in \mathcal{M}_i$ and $c_{i,1}, \ldots, c_{i,k_i}$'s be as before. If $(\mathbb{x}, \tau, m, c_{i,t}) \notin \mathcal{D}$ for all $t \in [k_i]$, then we are done. Otherwise, there exists t_0 such that $(\mathbb{x}, \tau, m, c_{i,t_0}) \in \mathcal{D}$. Let $m' = \mathsf{EndNode}(\mathbb{x}, \tau, m, c_{i,t_0}) \in \mathcal{M}_{i+1}$. Then $(\mathbb{x}, \tau, m, c_{i,t_0}, m')$ is of the form $(\mathbb{x}, \tau', m')$ where $(\mathbb{x}, \tau') \in \mathrm{PartTr}(i+1)$ and $(\mathbb{x}, \tau') \in \mathcal{D}$. Now we can apply our induction hypothesis on $(\mathbb{x}, \tau')$. The claim then follows immediately. This completes the proof of the theorem. □

Finally, we prove Theorem 1.3.

Proof of Theorem 1.3. Let $\mathbf{R}$ be a relation such that $\mathcal{L}_{\mathbf{R}}$ is a language in NP but not in P. Let $\mu = \mu(|\mathbb{x}|)$ be a polynomial function on the length of inputs. We define a μ-round IOP Π for $\mathbf{R}$ as follows. Let $(\mathbb{x}, \mathbb{w}) \in \mathbf{R}$. The first μ messages sent by the honest P to the verifier V are all the single bit message 0. The last message sent by P is the witness $\mathbb{w}$. As for the verifier, as in all public-coin IOPs, all challenges sent by V are uniformly sampled random strings. At the end of the protocol, V checks whether $(\mathbb{x}, \mathbb{w}) \in \mathbf{R}$, and accepts only if this is the case and all of the messages sent by the prover are the single bit 0.

This IOP Π is sound with soundness error 0. Moreover, Π is $(1, \overset{\mu}{\ldots}, 1)$-special sound, since any $(1, \overset{\mu}{\ldots}, 1)$-tree of accepting transcripts for an input $\mathbb{x}$ has a valid witness $\mathbb{w}$ for $\mathbb{x}$ in its one leaf. Observe also that Π is perfectly complete, i.e. an honest prover convinces the verifier with probability 1.

Now assume Π has RBR knowledge with errors $(\varepsilon_1, \ldots, \varepsilon_\mu)$. Let $\mathcal{D}$ be the corresponding doomed set of partial transcripts, and let Ext be the corresponding extractor. We argue that for some input length ℓ and some $i \in [\mu]$ we have $\varepsilon_i(\ell) = 1$. Indeed, assume towards contradiction that this is not the case. We will describe a deterministic polynomial time algorithm that, given any $\mathbb{x} \in \mathcal{L}_{\mathbf{R}}$, outputs a valid witness $\mathbb{w}$ for $\mathbb{x}$.

To this end, fix first $(\mathbb{x}, \mathbb{w}) \in \mathbf{R}$. By our previous assumption, we have $\varepsilon_i(|\mathbb{x}|) < 1$ for all i. Let $i \in [\mu]$, let $(\mathbb{x}, \tau)$ be a $(i-1)$-round partial transcript with $(\mathbb{x}, \tau) \in \mathcal{D}$, and let $m \in \mathcal{M}_i$. Define

$$\rho(\mathbb{x}, \tau, m) := \Pr_{c \in C_i}[(\mathbb{x}, \tau, m, c) \notin \mathcal{D}]$$

Let $(\mathbb{x}, \tau')$ be a μ-round partial transcript generated from the interaction of the honest prover and the verifier, so that $(\mathbb{x}, \tau') = (\mathbb{x}, 0, c_1, 0, c_2, \ldots, 0, c_\mu)$ for some challenges c_i. Observe that $(\mathbb{x}, \tau', \mathbb{w})$ is a complete transcript that is accepted by the verifier. Hence, by definition of RBR knowledge, $(\mathbb{x}, \tau') \notin \mathcal{D}$. Let $\rho_{i-1} := \rho(\mathbb{x}, 0, c_1, \ldots, 0, c_{i-1}, 0)$. We claim that $\rho_{i-1} = 1$ for some $i \in [\mu]$. Indeed, if this was not the case, Π would not be perfectly complete, since, given $(\mathbb{x}, \mathbb{w})$ the honest prover would only convince the verifier with probability at most $1 - \prod_{i \in [\mu]}(1 - \rho_{i-1})$ (i.e. the probability that, for some i, a doomed partial transcript $(\mathbb{x}, 0, c_1, \ldots, 0, c_{i-1}, 0)$ escapes the doomed set after receiving a challenge c_i), but this latter probability is strictly less than 1 unless $\rho_{i-1} = 1$ for some $i \in [\mu]$. This proves our claim.

Now, let $i \in [\mu]$ be such that $\rho_{i-1} = 1$. Then, since $\rho_{i-1} > \varepsilon_i(|\mathbb{x}|)$, given $(\mathbb{x}, 0, c_1, \ldots, c_{i-1}, 0)$ the extractor Ext outputs a valid witness $\mathbb{w}'$ for $\mathbb{x}$ in polynomial time. Notice that, conversely, if $\mathbb{x} \notin \mathcal{L}_{\mathbf{R}}$, then given any partial transcript of the form $(\mathbb{x}, 0, c_1, \ldots, c_{i-1}, 0)$, Ext outputs $\perp$ in polynomial time.

This suggests the following polynomial algorithm Ext′ for the language $\mathcal{L}_{\mathbf{R}}$. Given an arbitrary input $\mathbb{x}$, generate μ partial transcripts of the form $(\mathbb{x}, 0, c_1, \ldots, 0, c_{i-1}, 0)$, for $i \in [\mu]$. Here the c_i are strings generated in an arbitrary deterministic manner. Give each of these as input to Ext. If, for some of these, Ext outputs a valid witness $\mathbb{w}'$ for $\mathbb{x}$, then Ext′ outputs $\mathbb{w}'$. Otherwise Ext′ outputs $\perp$.

The algorithm Ext′ is a deterministic algorithm. Moreover, it runs in polynomial time because 1) $\mu(|\mathbb{x}|)$ is polynomial and 2) Ext runs in polynomial time. This contradicts the fact that $\mathcal{L}_{\mathbf{R}}$ is in NP but not in P. Thus, our initial assumption that $\varepsilon_i(|\mathbb{x}|) < 1$ for all $\mathbb{x}$ and all $i \in [\mu]$ cannot hold. □

6 Special Unsoundness as the Dual of Round-by-round Soundness

In this section we discuss relations between the soundness, the RBR soundness, and the special unsoundness of an IOP Π. We will see that RBR soundness and special unsoundness are "dual" concepts of each other, and that the former upper bounds the soundness of Π, while the latter lower bounds it, and moreover allows an attack to Π whose success probability is this lower bound. We will also see that when the RBR soundness and the special unsoundness errors are the same, then these errors are tight and their combination equals the soundness of Π.

We begin by formulating a variation of the definition of special unsoundness (cf., Definition 3.7). The main differences are that, instead of having sets of "lucky challenges", we have sets of "lucky partial transcripts". The motivation behind this alternative formulation is that it highlights why special unsoundness acts as the dual notion of RBR soundness.

Definition 6.1 (Special Unsoundness [AFK22] – alternative formulation)**.** *Let Π be a μ-round IOP for a relation $\mathbf{R}$. We say Π is* special unsound *with errors $(\varepsilon_1, \ldots, \varepsilon_\mu)$ if there exist a set $\mathcal{L}$ of "lucky" partial transcripts, and an unbounded prover algorithm P^* such that, for all $\mathbb{x}$, the following hold.*

- *For all $\mathbb{x} \notin \mathcal{L}_{\mathbf{R}}$ we have that the 0-round partial transcript $(\mathbb{x})$ does not belong to $\mathcal{L}$.*
- *Let $(\mathbb{x}, \tau) \in PartTr(\mu)$ be a μ-round partial transcript. Assume $(\mathbb{x}, \tau) \in \mathcal{L}$. Then $\mathsf{V}(\mathbb{x}, \tau, m) = \mathsf{accept}$ for any last prover's message $m \in \mathcal{M}_{\mu+1}$.*

  Moreover, for all $i \in [\mu - 1]$ and $(\mathbb{x}, \tau) \in \mathcal{L} \cap PartTr(i-1)$, P^* is able to compute $m \in \mathcal{M}_i$ such that $(\mathbb{x}, \tau, m, c) \in \mathcal{L}$ for all $c \in C_i$.

  In words, given a $(i-1)$-round partial transcript that is "lucky", P^* is able to choose a message from $\mathcal{M}_i$ so that the subsequent i-round partial transcript is lucky, no matter what challenge the verifier sends.

  Consequently (by induction), in this case the prover is then able to find a lucky complete transcript, making the verifier accept.
- *Let $i \in [\mu]$ and let $(\mathbb{x}, \tau) \in PartTr(i-1)$ be a $(i-1)$-round partial transcript produced during the interaction of the prover and the verifier , and let $m \leftarrow \mathsf{P}^*(\mathbb{x}, \tau)$. Then it holds that*

$$\Pr_{c \xleftarrow{\$} C_i}[(\mathbb{x}, \tau, m, c) \in \mathcal{L}] \geq \varepsilon_i.$$

Remark 6.2 (Special Unsoundness as the Dual Notion of RBR Soundness). Definition 6.1 can be understood as the dual of RBR soundness in the sense that the sets $\mathcal{L}, \mathcal{D}$ of lucky and doomed transcripts from the respective definition have opposite properties:

- An input $\mathbb{x} \notin \mathcal{L}$ does not belong to $\mathcal{L}$, while it belongs to $\mathcal{D}$.
- The verifier rejects any doomed complete transcript, while it accepts any complete lucky transcript.

  On the other hand, given a non-lucky partial transcript, if after some subsequent round the partial transcript is lucky, the verifier will eventually accept.
- For all round $i \in [\mu - 1]$, and for all prover, the probability that a $(i-1)$-round doomed partial transcripts stops being doomed at Round i is at most ε_i. On the other hand, there exists a prover that for all non-lucky $(i-1)$-round partial transcript, the probability that the transcript becomes lucky in the next round is at least ε'_i.

Remark 6.3 (Equivalence between Definition 3.7 and Definition 6.1). Definitions 3.7 and 6.1 are indeed equivalent: the set of "lucky" challenges from Definition 3.7 depend on the partial transcript so far, hence one may as well define a set of "lucky" partial transcripts instead, as we did in Definition 6.1. Moreover, the active and passive modes of the adversary P^* from Definition 3.7 can be seen as analogues of P^* having or not produced a lucky partial transcript. If it has (passive mode), then P^* operates in a way that all subsequent transcripts are lucky, and so, following this analogy, P^* stays in passive mode until the end of the proof. As required, at the end of the proof, if the complete transcript is lucky (analogously, if P^* is in passive mode), the verifier accepts.

The next result relates the notions of soundness, RBR soundness, and special unsoundness of an IOP. The key observation is that the concepts of special unsoundness and round-by-round soundness are dual of each other. As a result, and intuitively speaking, we have that the special unsoundness "error" lower bounds the soundness of the protocol, while the round-by-round soundness "error" upper bounds it.

Theorem 1.4 (Relation Between Soundness, Round-by-round Soundness, and Special Unsoundness)**.** *Let Π be a μ-round IOP for a relation $\mathbf{R}$. Assume Π has soundness error ε. Then the following hold:*

- ***RBR soundness is an upper bound for soundness.*** *If Π is round-by-round sound with errors $\varepsilon_1, \ldots, \varepsilon_\mu$, then*

$$\varepsilon \leq 1 - \prod_{i \in [\mu]} (1 - \varepsilon_i)$$

for all $\mathbb{x} \notin \mathcal{L}_{\mathbf{R}}$.

- ***Special unsoundness is a lower bound for soundness.*** *If Π is special unsound with errors $\varepsilon'_1, \ldots, \varepsilon'_\mu$, then*

$$1 - \prod_{i \in [\mu]} \left(1 - \varepsilon'_i\right) \leq \varepsilon$$

for all $\mathbb{x} \notin \mathcal{L}_{\mathbf{R}}$. Moreover, there exists a dishonest unbounded prover P^* that, given any input $\mathbb{x}$, manages to make the verifier accept with probability at least $1 - \prod_{i \in [\mu]} \left(1 - \varepsilon'_i\right)$.

- ***Tightness of RBR soundness, soundness, and special unsoundness.*** *Suppose Π is round-by-round sound with errors $\varepsilon_1, \ldots, \varepsilon_\mu$ and that Π is special unsound with the same errors $\varepsilon_1, \ldots, \varepsilon_\mu$. Then*

$$\varepsilon = 1 - \prod_{i \in [\mu]} (1 - \varepsilon_i)\,.$$

Moreover, the error is tight in the sense that there exists a dishonest prover P^* that, given any input $\mathbb{x}$, manages to have the verifier accept with probability at least ε.

Remark 6.4. The quantity $\rho := 1 - \prod_{i \in [\mu]}(1 - x_i)$ is, for small x_i's, approximately $\sum_{i \in [\mu]} x_i$, as this is the first-order term in the Taylor approximation of ρ around the point $(0, \ldots, 0)$.

Proof. We begin by proving Item 1 of the theorem. Assume $\mathbb{x} \notin \mathcal{L}_{\mathbf{R}}$ and let P^* be any unbounded dishonest prover. Let $(\mathbb{x}, \tau) \in \mathrm{PartTr}(\mu)$ be a μ-round partial transcript produced during the interaction of P^* and V, and let $m \in \mathcal{M}_\mu$, so that $(\mathbb{x}, \tau, m)$ is a complete transcript. Let $\rho := = 1 - \Pi_{i \in [\mu]}(1 - \varepsilon_i)$.

Let $P := \Pr_{\leftarrow \langle \mathsf{P}^*, \mathsf{V} \rangle}[\mathsf{V}(\mathbb{x}, \tau) = \mathsf{accept}]$. $P \leq \rho$ (for all input length ℓ and all $\mathbb{x}, \tau$, with $|\mathbb{x}| = \ell$).

In order to have $\mathsf{V}(\mathbb{x}, \tau, m) = \mathsf{accept}$, it is necessary (but not sufficient) that $(\mathbb{x}, \tau) \notin \mathcal{D}_\mu$. Hence

$$\Pr_{(\tau, m) \leftarrow \langle \mathsf{P}^*, \mathsf{V} \rangle(\mathbb{x})}[\mathsf{V}(\mathbb{x}, \tau) = \mathsf{accept}] \leq \Pr_{(\tau, m) \leftarrow \langle \mathsf{P}^*, \mathsf{V} \rangle(\mathbb{x})}[(\mathbb{x}, \tau) \notin \mathcal{D}_\mu].$$

Let E_i be the event that the prefix of $(\mathbb{x}, \tau)$ that is a i-round partial transcript, is not in a doomed set. Assume for now that P^* has the capacity to, if E_i has happened, make $E_{i+1}, \ldots, E_\mu$ occur with probability 1. Then $\Pr_{(\tau, m) \leftarrow \langle \mathsf{P}^*, \mathsf{V} \rangle(\mathbb{x})}[(\mathbb{x}, \tau) \notin \mathcal{D}_\mu]$ is the probability that, among μ trials (each trial corresponding to "leaving the doomed set" at one of the μ rounds of interaction), each with success probability at most ε_i ($i \in [\mu]$), respectively, one of them succeeds. This is because once Π^* manages to produce a partial transcript that is not in $\mathcal{D}$, it is able to operate (by our assumption), in a way that the subsequent partial transcripts never belong to $\mathcal{D}$ again. The aforementioned probability is precisely $\rho = 1 - \Pi_{i \in [\mu]}(1 - \varepsilon_i)$. Hence

$$\Pr_{(\tau, m) \leftarrow \langle \mathsf{P}^*, \mathsf{V} \rangle(\mathbb{x})}[(\mathbb{x}, \tau) \notin \mathcal{D}_\mu] \leq \rho. \tag{3}$$

Now consider any other attacker $\mathsf{P}^{*\prime}$ that, once it has obtained a partial transcript that does not belong to $\mathcal{D}$, it operates in a way that makes it possible for a further partial transcript to ends back up in $\mathcal{D}$. This attacker always produces a non-doomed complete transcript with at most the probability that the previous attacker P^* does. This is because once a previously doomed partial transcript has been extended so that it is not in a doomed set, P^* ends with a non-doomed complete transcript with probability 1, while the probability of $\mathsf{P}^{*\prime}$ may be smaller. This completes the proof of Item 1.

Next we prove Item 2. Fix an input $\mathbb{x} \notin \mathcal{L}_{\mathbf{R}_i}$. Let P^* be the dishonest prover from Definition 6.1. As before, let $(\mathbb{x}, \tau) \in \mathrm{PartTr}(\mu)$ be a μ-round transcript produced during the interaction of the prover and the verifier, and let $m \in \mathcal{M}_\mu$.

In order to have $\mathsf{V}(\mathbb{x}, \tau, m) = \mathsf{accept}$, it is sufficient (but not necessary) that there exist a prefix $(\mathbb{x}, \tau_i)$ of $(\mathbb{x}, \tau)$ that is a i-round partial transcript and belongs to the lucky set $\mathcal{L}$. Let E be the event that such a prefix exists. Similarly as in Item 1, the probability of E occurring is at least $\rho' := 1 - \prod_{i \in [\mu]}(1 - \varepsilon'_i)$. i.e. the probability that after μ trials, each with success probability ε'_i, at least one trial was successful. Hence

$$\Pr_{(\tau,m) \leftarrow \langle \mathsf{P}^*, \mathsf{V} \rangle (\mathbb{x})} [\mathsf{V}(\mathbb{x}, \tau, m) = \mathsf{accept}] \geq \rho', \tag{4}$$

proving Item 2.

Item 3 follows from Items 1 and 2, since under its hypotheses we have that any malicious prover convinces the verifier with probability at most $1 - \prod_{i \in [\mu]}(1 - \varepsilon_i)$, and that there exists a prover that convinces the verifier with at least this probability. □

Remark 6.5. In [CCH+18], the authors prove that, given a μ-round interactive proof/argument Π with soundness $\varepsilon_{\mathsf{sound}}$ and round-by-round soundness $\varepsilon_{\mathsf{rbr}}$ (in the non-generalized sense of RBR soundness), one has $\varepsilon_{\mathsf{sound}} \leq \mu \varepsilon_{\mathsf{rbr}}$. The previous Theorem 1.4 yields a slight improvement in this formula, giving $\varepsilon_{\mathsf{sound}} \leq 1 - (1 - \varepsilon_{\mathsf{rbr}})^{\mu}$. We remark however that we expect this improvement to be known already (at least in folklore), and that when the errors are small, the improvement is negligible (cf., Remark 6.4).

Acknowledgements

Alexander R. Block was supported by DARPA under Contract No. HR00112020022 and No. HR00112020025. Albert Garreta was supported by the Ethereum Foundation's grant FY23-0885. Pratyush Ranjan Tiwari was partly supported by the Ethereum Foundation's grant FY23-1087, a Security & Privacy research gift from Google, and a research gift from Cisco. Michał Zając was supported by the Ethereum Foundation's grant FY23-0885. The views, opinions, findings, conclusions and/or recommendations expressed in this material are those of the authors and should not be interpreted as reflecting the position or policy of DARPA or the United States Government, and no official endorsement should be inferred.

References

[ACF21] Thomas Attema, Ronald Cramer, and Serge Fehr. Compressing proofs of k-out-of-n partial knowledge. In Tal Malkin and Chris Peikert, editors, *CRYPTO 2021, Part IV*, volume 12828 of *LNCS*, pages 65–91, Virtual Event, August 2021. Springer, Heidelberg. doi:10.1007/978-3-030-84259-8_3. 2

[ACK21] Thomas Attema, Ronald Cramer, and Lisa Kohl. A compressed Σ-protocol theory for lattices. In Tal Malkin and Chris Peikert, editors, *CRYPTO 2021, Part II*, volume 12826 of *LNCS*, pages 549–579, Virtual Event, August 2021. Springer, Heidelberg. doi:10.1007/978-3-030-84245-1_19. 6

[AFK22] Thomas Attema, Serge Fehr, and Michael Klooß. Fiat-shamir transformation of multi-round interactive proofs. In Eike Kiltz and Vinod Vaikuntanathan, editors, *Theory of Cryptography - 20th International Conference, TCC 2022, Chicago, IL, USA, November 7-10, 2022, Proceedings, Part I*, 2022. 1, 2, 3, 5, 6, 11, 13, 19

[Bab85] László Babai. Trading group theory for randomness. In Robert Sedgewick, editor, *Proceedings of the 17th Annual ACM Symposium on Theory of Computing, May 6-8, 1985, Providence, Rhode Island, USA*, pages 421–429. ACM, 1985. 2

[BBHR18] Eli Ben-Sasson, Iddo Bentov, Yinon Horesh, and Michael Riabzev. Fast reed-solomon interactive oracle proofs of proximity. In *45th International Colloquium on Automata, Languages, and Programming, ICALP 2018, July 9-13, 2018, Prague, Czech Republic*, 2018. 1

[BCGT13] Eli Ben-Sasson, Alessandro Chiesa, Daniel Genkin, and Eran Tromer. Fast reductions from RAMs to delegatable succinct constraint satisfaction problems: extended abstract. In Robert D. Kleinberg, editor, *ITCS 2013*, pages 401–414. ACM, January 2013. doi:10.1145/2422436.2422481. 6

[BCS16] Eli Ben-Sasson, Alessandro Chiesa, and Nicholas Spooner. Interactive oracle proofs. In Martin Hirt and Adam D. Smith, editors, *TCC 2016-B, Part II*, volume 9986 of *LNCS*, pages 31–60. Springer, Heidelberg, October / November 2016. doi:10.1007/978-3-662-53644-5_2. 1, 4, 5, 6, 9, 15

[BDG+13] Nir Bitansky, Dana Dachman-Soled, Sanjam Garg, Abhishek Jain, Yael Tauman Kalai, Adriana López-Alt, and Daniel Wichs. Why "Fiat-Shamir for proofs" lacks a proof. In Amit Sahai, editor, *TCC 2013*, volume 7785 of *LNCS*, pages 182–201. Springer, Heidelberg, March 2013. doi:10.1007/978-3-642-36594-2_11. 6

[BGKS20] Eli Ben-Sasson, Lior Goldberg, Swastik Kopparty, and Shubhangi Saraf. DEEP-FRI: sampling outside the box improves soundness. In *Innovations in Theoretical Computer Science Conference, ITCS*, 2020. 1

[BTVW14] Andrew J. Blumberg, Justin Thaler, Victor Vu, and Michael Walfish. Verifiable computation using multiple provers. Cryptology ePrint Archive, Report 2014/846, 2014. https://eprint.iacr.org/2014/846. 6

[CBBZ23] Binyi Chen, Benedikt Bünz, Dan Boneh, and Zhenfei Zhang. Hyperplonk: Plonk with linear-time prover and high-degree custom gates. In *Advances in Cryptology - EUROCRYPT*, 2023. 1

[CCH+18] Ran Canetti, Yilei Chen, Justin Holmgren, Alex Lombardi, Guy N. Rothblum, and Ron D. Rothblum. Fiat-Shamir from simpler assumptions. Cryptology ePrint Archive, Report 2018/1004, 2018. https://eprint.iacr.org/2018/1004. 21

[CCH+19] Ran Canetti, Yilei Chen, Justin Holmgren, Alex Lombardi, Guy N. Rothblum, Ron D. Rothblum, and Daniel Wichs. Fiat-Shamir: from practice to theory. In Moses Charikar and Edith Cohen, editors, *51st ACM STOC*, pages 1082–1090. ACM Press, June 2019. doi:10.1145/3313276.3316380. 1, 5, 6, 13

[CCRR18] Ran Canetti, Yilei Chen, Leonid Reyzin, and Ron D. Rothblum. Fiat-Shamir and correlation intractability from strong KDM-secure encryption. In Jesper Buus Nielsen and Vincent Rijmen, editors, *EUROCRYPT 2018, Part I*, volume 10820 of *LNCS*, pages 91–122. Springer, Heidelberg, April / May 2018. doi:10.1007/978-3-319-78381-9_4. 6

[CDS94] Ronald Cramer, Ivan Damgård, and Berry Schoenmakers. Proofs of partial knowledge and simplified design of witness hiding protocols. In Yvo Desmedt, editor, *Advances in Cryptology - CRYPTO '94, 14th Annual International Cryptology Conference, Santa Barbara, California, USA, August 21-25, 1994, Proceedings*, volume 839 of *Lecture Notes in Computer Science*, pages 174–187. Springer, 1994. 1, 2

[CMS19] Alessandro Chiesa, Peter Manohar, and Nicholas Spooner. Succinct arguments in the quantum random oracle model. In Dennis Hofheinz and Alon Rosen, editors, *TCC 2019, Part II*, volume 11892 of *LNCS*, pages 1–29. Springer, Heidelberg, December 2019. doi:10.1007/978-3-030-36033-7_1. 1, 4, 5, 6, 9, 13, 15, 17

[CMSZ21] Alessandro Chiesa, Fermi Ma, Nicholas Spooner, and Mark Zhandry. Post-quantum succinct arguments: Breaking the quantum rewinding barrier. In *62nd IEEE Annual Symposium on Foundations of Computer Science, FOCS 2021, Denver, CO, USA, February 7-10, 2022*, pages 49–58. IEEE, 2021. doi:10.1109/FOCS52979.2021.00014. 4

[CMT12] Graham Cormode, Michael Mitzenmacher, and Justin Thaler. Practical verified computation with streaming interactive proofs. In Shafi Goldwasser, editor, *ITCS 2012*, pages 90–112. ACM, January 2012. doi:10.1145/2090236.2090245. 6

[COS20] Alessandro Chiesa, Dev Ojha, and Nicholas Spooner. Fractal: Post-quantum and transparent recursive proofs from holography. In Anne Canteaut and Yuval Ishai, editors, *EUROCRYPT 2020, Part I*, volume 12105 of *LNCS*, pages 769–793. Springer, Heidelberg, May 2020. doi:10.1007/978-3-030-45721-1_27. 1, 4, 5, 6

[DFMS19] Jelle Don, Serge Fehr, Christian Majenz, and Christian Schaffner. Security of the Fiat-Shamir transformation in the quantum random-oracle model. In Alexandra Boldyreva and Daniele Micciancio, editors, *CRYPTO 2019, Part II*, volume 11693 of *LNCS*, pages 356–383. Springer, Heidelberg, August 2019. doi:10.1007/978-3-030-26951-7_13. 4

[FS87] Amos Fiat and Adi Shamir. How to prove yourself: Practical solutions to identification and signature problems. In Andrew M. Odlyzko, editor, *CRYPTO'86*, volume 263 of *LNCS*, pages 186–194. Springer, Heidelberg, August 1987. doi:10.1007/3-540-47721-7_12. 2, 6

[GKR08] Shafi Goldwasser, Yael Tauman Kalai, and Guy N. Rothblum. Delegating computation: interactive proofs for muggles. In Richard E. Ladner and Cynthia Dwork, editors, *40th ACM STOC*, pages 113–122. ACM Press, May 2008. doi:10.1145/1374376.1374396. 6

[GMR89] Shafi Goldwasser, Silvio Micali, and Charles Rackoff. The knowledge complexity of interactive proof systems. *SIAM J. Comput.*, 1989. 2

[GWC19] Ariel Gabizon, Zachary J. Williamson, and Oana Ciobotaru. PLONK: Permutations over lagrange-bases for oecumenical noninteractive arguments of knowledge. Cryptology ePrint Archive, Report 2019/953, 2019. https://eprint.iacr.org/2019/953. 1

[HL18] Justin Holmgren and Alex Lombardi. Cryptographic hashing from strong one-way functions (or: One-way product functions and their applications). In Mikkel Thorup, editor, *59th FOCS*, pages 850–858. IEEE Computer Society Press, October 2018. doi:10.1109/FOCS.2018.00085. 6

[Hol19] Justin Holmgren. On round-by-round soundness and state restoration attacks. Cryptology ePrint Archive, Report 2019/1261, 2019. https://eprint.iacr.org/2019/1261. 1, 4, 5, 6, 13

[Kil92] Joe Kilian. A note on efficient zero-knowledge proofs and arguments (extended abstract). In *24th ACM STOC*, pages 723–732. ACM Press, May 1992. doi:10.1145/129712.129782. 6

[KPV19] Assimakis Kattis, Konstantin Panarin, and Alexander Vlasov. Redshift: Transparent snarks from list polynomial commitments. Cryptology ePrint Archive, Paper 2019/1400, 2019. https://eprint.iacr.org/2019/1400. URL: https://eprint.iacr.org/2019/1400, doi:10.1145/548606.3560657. 1, 2, 4

[KRR17] Yael Tauman Kalai, Guy N. Rothblum, and Ron D. Rothblum. From obfuscation to the security of Fiat-Shamir for proofs. In Jonathan Katz and Hovav Shacham, editors, *CRYPTO 2017, Part II*, volume 10402 of *LNCS*, pages 224–251. Springer, Heidelberg, August 2017. doi:10.1007/978-3-319-63715-0_8. 6

[LZ19] Qipeng Liu and Mark Zhandry. Revisiting post-quantum Fiat-Shamir. In Alexandra Boldyreva and Daniele Micciancio, editors, *CRYPTO 2019, Part II*, volume 11693 of *LNCS*, pages 326–355. Springer, Heidelberg, August 2019. doi:10.1007/978-3-030-26951-7_12. 4

[Mic94] Silvio Micali. CS proofs (extended abstracts). In *35th FOCS*, pages 436–453. IEEE Computer Society Press, November 1994. doi:10.1109/SFCS.1994.365746. 6

[Pol22] Polygon Zero Team. Plonky2: Fast recursive arguments with plonk and fri, 2022. https://github.com/mir-protocol/plonky2/tree/main/plonky2. 1

[RR20] Noga Ron-Zewi and Ron D. Rothblum. Local proofs approaching the witness length [extended abstract]. In *61st FOCS*, pages 846–857. IEEE Computer Society Press, November 2020. doi:10.1109/FOCS46700.2020.00083. 6

[RRR21] Omer Reingold, Guy N. Rothblum, and Ron D. Rothblum. Constant-round interactive proofs for delegating computation. *SIAM J. Comput.*, 2021. 1

[Set20] Srinath Setty. Spartan: Efficient and general-purpose zkSNARKs without trusted setup. In Daniele Micciancio and Thomas Ristenpart, editors, *CRYPTO 2020, Part III*, volume 12172 of *LNCS*, pages 704–737. Springer, Heidelberg, August 2020. doi:10.1007/978-3-030-56877-1_25. 6

[Tha13] Justin Thaler. Time-optimal interactive proofs for circuit evaluation. In Ran Canetti and Juan A. Garay, editors, *CRYPTO 2013, Part II*, volume 8043 of *LNCS*, pages 71–89. Springer, Heidelberg, August 2013. doi:10.1007/978-3-642-40084-1_5. 6

[Wik18] Douglas Wikström. Special soundness revisited. Cryptology ePrint Archive, Report 2018/1157, 2018. https://eprint.iacr.org/2018/1157. 2

[Wik21] Douglas Wikström. Special soundness in the random oracle model. Cryptology ePrint Archive, Report 2021/1265, 2021. https://eprint.iacr.org/2021/1265. 1, 2

[WTs+18] Riad S. Wahby, Ioanna Tzialla, abhi shelat, Justin Thaler, and Michael Walfish. Doubly-efficient zkSNARKs without trusted setup. In *2018 IEEE Symposium on Security and Privacy*, pages 926–943. IEEE Computer Society Press, May 2018. doi:10.1109/SP.2018.00060. 6