Rogue-Instance Security for Batch Knowledge Proofs

Gil Segev[1(✉)], Amit Sharabi[2], and Eylon Yogev[2]

[1] School of Computer Science and Engineering, Hebrew University of Jerusalem, 91904 Jerusalem, Israel
segev@cs.huji.ac.il

[2] Department of Computer Science, Bar-Ilan University, Ramat Gan, Israel
amit.sharabi1@live.biu.ac.il, eylon.yogev@biu.ac.il

Abstract. We propose a new notion of knowledge soundness, denoted *rogue-instance security*, for interactive and non-interactive *batch* knowledge proofs. Our notion, inspired by the standard notion of rogue-key security for multi-signature schemes, considers a setting in which a malicious prover is provided with an honestly-generated instance $\mathbb{x}_1$, and may then be able to maliciously generate related "rogue" instances $\mathbb{x}_2, \ldots, \mathbb{x}_k$ for convincing a verifier in a batch knowledge proof of corresponding witnesses $\mathbb{w}_1, \ldots, \mathbb{w}_k$ for all k instances – without actually having knowledge of the witness $\mathbb{w}_1$ corresponding to the honestly-generated instance. This setting provides a powerful security guarantee for batch versions of a wide variety of practically-relevant protocols, such as Schnorr's protocol and similar ones.

We present a highly-efficient generic construction of a batch proof-of-knowledge applicable to any *algebraic* Sigma protocols. The algebraic property refers to a homomorphic structure of the underlying group and includes Schnorr's protocol and others. We provide an almost tight security analysis for our generic batch protocol, which significantly improves upon the previously known security bounds even for the specific case of batch Schnorr protocol. We extend our results beyond algebraic Sigma protocols. We analyze the rogue-instance security of a general batch protocol with plus-one special soundness (a generalization of standard special soundness) and achieve improved security bounds in the generic case.

Our results use a particular type of *high-moment* assumptions introduced by Rotem and Segev (CRYPTO 2021). These assumptions consider the hardness of a relation against algorithms with bounded *expected*

Gil Segev is supported by the Israel Science Foundation (Grant No. 1336/22) and by the European Union (ERC, FTRC, 101043243). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.
Amit Sharabi is sponsored by the Israel Science Foundation (Grant No. 2439/20).
Eylon Yogev is supported by an Alon Young Faculty Fellowship, by the Israel Science Foundation (Grant No. 2302/22), and by the BIU Center for Research in Applied Cryptography and Cyber Security in conjunction with the Israel National Cyber Bureau in the Prime Minister's Office.

G. Rothblum and H. Wee (Eds.): TCC 2023, LNCS 14369, pp. 121–157, 2023.
https://doi.org/10.1007/978-3-031-48615-9_5

running time. Although Rotem and Segev introduced these assumptions, they did not provide evidence to support their hardness. To substantiate and validate the high-moment assumptions, we present a new framework for assessing the concrete hardness of cryptographic problems against oracle algorithms with bounded expected runtime. Our framework covers generic models, including the generic group model, random oracle model, and more. Utilizing our framework, we achieve the first hardness result for these high-moment assumptions. In particular, we establish the second-moment hardness of the discrete-logarithm problem against expected-time algorithms in the generic group model.

1 Introduction

A zero-knowledge proof-of-knowledge protocol is a powerful cryptographic tool with diverse applications. It enables a prover to convincingly demonstrate to a verifier, who holds an instance $\mathbb{x}$, that it possesses knowledge of a valid witness $\mathbb{w}$ for $\mathbb{x}$. The fundamental power of such protocols lies in the ability to *extract* a witness from a given prover, a property that varies in its precise formulation across different protocols. Proofs of knowledge play a pivotal role in cryptographic protocols, both from a theoretical standpoint and in practical implementations.

One notable example is Schnorr's protocol [31,32], which serves as a zero-knowledge proof-of-knowledge for the knowledge of the discrete-logarithm of a group element. In its interactive form, this protocol offers an efficient identification scheme, while in its non-interactive form, it translates into a signature scheme via the Fiat-Shamir transformation. The widespread influence of the Schnorr identification and signature schemes stems from their conceptual simplicity and practical efficiency. Another compelling example is a proof-of-knowledge for a Pedersen commitment or hash function, which is the product of two Schnorr instances. In this scenario, the prover demonstrates the ability to "open" the commitment without actually revealing its contents, thus maintaining the privacy of the committer [27]. The wide-ranging applicability of these protocols within the field of cryptography has garnered substantial attention and interest in a tight analysis of their security bounds.

Extraction from Special Soundness. Both of the examples presented above exemplify Sigma protocols, which are three-move protocols that exhibit the unique soundness notion called "special soundness". This property plays a vital role in the construction of an extractor. Specifically, the property states that it is possible to extract a witness when provided with two accepting transcripts that share the same first message but differ in the second message. Consequently, to establish the protocol's security based on the hardness of the underlying relation, the extractor must successfully extract two such valid transcripts from a potentially malicious prover.

To achieve this goal, existing approaches employ a strategy of executing the protocol multiple times. The analysis of these approaches draws upon the

classic "forking lemma" introduced by Pointcheval and Stern [28] (see also [1, 7,10,21]). These different approaches showcase a trade-off between the success probability and the running time of the extractor. To provide a concrete example, let us examine the Schnorr identification scheme and signature scheme, which derive their security from the hardness of the discrete-logarithm problem. For the Schnorr identification scheme, suppose we have a malicious prover who runs in time t and succeeds in impersonating with probability ϵ. We can transform this malicious prover into a discrete-logarithm algorithm that runs in time $2t$ and succeeds with probability ϵ^2. Similarly, for the Schnorr signature scheme, suppose the attacker additionally performs at most q queries to the random oracle. We can transform this attacker into a discrete-logarithm algorithm that runs in time $2t$ and succeeds with probability ϵ^2/q. For any group of order p, where generic hardness of discrete-log is believed to hold [33], this leads to the bound $\epsilon \leq (t^2/p)^{1/2}$ for the Schnorr identification scheme, and a bound of $\epsilon \leq (q \cdot t^2/p)^{1/2}$ for the Schnorr signature scheme. Other trade-offs that were established lead to the same bound [5,19]. In idealized models, such as the generic group model [22,33] and the algebraic group model [2,4,14,15,25,29], it is possible to achieve an optimal bound of $\epsilon \leq t^2/p$ (see [15,33]).

High-Moment Forking Lemma. The extractor runs the given adversary for the second time, only if the first time succeeded. Thus, it is convenient to analyze the *expected* running-time of the extractor, rather than its strict running-time [20]. In this case, the result is an algorithm for solving discrete-logarithm with a bound on its expected running time. Recently, Segev and Rotem [30] have leveraged this type of analysis to derive tighter bounds for Schnorr's protocols (and similar Sigma protocols). Towards this end, they established a hardness of discrete-logarithm for excepted time algorithms.

In simple terms, their *second-moment assumption* states that the success probability ϵ of any algorithm A solving discrete-logarithm for a group of order p satisfies $\epsilon \leq \mathbb{E}\left[T_A^2\right]/p$, where T_A denotes the random variable corresponding to A's running time.[1] Under this assumption, Segev and Rotem were able to derive the bound of $\epsilon \leq (t^2/p)^{2/3}$, which is the best-known bound for Schnorr in the standard model. Achieving the optimal bound in the standard model remains an open problem that continues to drive ongoing research and exploration.

Batch Protocols. The Schnorr protocol and the Pedersen protocol both admit efficient *batch* versions [16]. A batch protocol is given k instances, $\mathbb{x}_1, \ldots, \mathbb{x}_k$, and allows to prove the knowledge of *all* corresponding k witnesses with a communication complexity that is approximately the same as that of a single proof of knowledge. The efficiency gain provided by batch protocols is a highly desirable property in many domains. In the context of blockchain, batching is a widely adopted practice aimed at reducing costs and optimizing resource utilization, the instances are usually public-keys and the witnesses are private-keys. By grouping

[1] They originally stated their assumption for a general d-moment but, in this paper, we focus on the second-moment.

multiple transactions or operations into a single batch, the associated overhead, such as communication and computation costs (which affect the transaction fees), can be significantly reduced.

However, the security analysis of batch protocols raises several concerns. The security bounds vary depending on how the instances are chosen in the security game (a modeling issue that does not appear with a single instance). For example, in a permissionless blockchain network, the attacker can choose the instances (its public-keys) adaptively as a function of existing instances sampled by honest parties. In such a case, the security reduction cannot assume hardness of the instances chosen by the adversary. These types of security games are known in the context of multi-signatures and are called rogue-key attacks (see [6,7,9,23,26] and the many references therein).

The special soundness property extends to the multiple instance case. In this setting, the extractor must extract $k+1$ valid transcripts from which it can compute all k corresponding witnesses (actually, it needs all $k+1$ transcripts even if it aims to compute a single witness). This is a generalization of the standard special soundness property, which we call *plus-one* special soundness. However, deriving tight security bounds for the batch setting is even more challenging than the single case. A straightforward extension of the single extractor to the batch version would run the malicious prover $k+1$ times and would yield an extractor that runs in approximately $(k+1) \cdot t$ time, but with a success probability of ϵ^{k+1}, i.e., an exponential decay in the number of instances. This is indeed the case in the batch Schnorr protocol given in [16]. Furthermore, the tighter bound of Segev and Rotem [30] does not seem to extend to the multiple instance case (regardless of the precise security game definition). This raises the question of how to derive tight security bounds for batch protocols.

1.1 Our Contributions

We give several contributions towards a better understanding of batch proof-of-knowledge protocols.

Rogue-Instance Soundness. Our first contribution is a strong security notion for batch protocols, denoted *rogue-instance security*, for interactive and non-interactive *batch* knowledge proofs. Our notion is inspired by the standard notion of rogue-key security for multi-signature schemes. We consider a setting in which a malicious prover is provided with an honestly-generated instance $\mathbb{x}_1$ (according to some distribution), and is then allowed to maliciously generate related "rogue" instances $\mathbb{x}_2, \ldots, \mathbb{x}_k$ for convincing a verifier in a batch knowledge proof of corresponding witnesses $\mathbb{w}_1, \ldots, \mathbb{w}_k$ for all k instances. This is done without the malicious prover having knowledge of the witness $\mathbb{w}_1$ corresponding to the honestly-generated instance. This setting provides a powerful security guarantee for batch versions of numerous practical and relevant protocols, such as Schnorr's protocol and similar ones. See Sect. 4 for the precise definition.

Batching Algebraic Sigma Protocols We construct batch protocols for a large family of Sigma protocols and provide a relatively tight analysis. Our construction works for *algebraic* Sigma protocols, which captures the proof-of-knowledge protocol for discrete-logarithm (Schnorr) [31,32], Pedersen commitment [27], Guillou-Quisquater identification scheme [17] and more. The algebraic property refers to a homomorphic structure of the underlying group. Algebraic Sigma protocols consist of an algebraic one-way function f such that the prover aims to prove knowledge of a preimage under f. The notion of algebraic one-way function introduced by Catalano et al. [11] which relates to the notion of group-homomorphic one-way generators introduced by Cramer and Damgård [13]. We analyze the security of our construction in the rogue-instance game and achieve the bound $\epsilon \leq (t^2/p)^{2/3}$ (for groups of order p) which matches the state-of-the-art bound of Segev and Rotem [30] for a single instance. In particular, our bound does not depend on the number of rogue instances. In more general form, our theorem is as follows.

Theorem 1 (Informal). *Let Π be an algebraic Sigma protocol for a relation $\mathcal{R} \subseteq \mathcal{X} \times \mathcal{W}$. If $\mathcal{R}$ is second-moment hard with respect to a distribution $\mathcal{D}$, then $\mathcal{R}$ has a batch protocol with rogue soundness error $\epsilon(t) \leq (t^2/|\mathcal{W}|)^{2/3}$.*

In particular, our theorem gives us tighter security bounds for the batch version of Schnorr and Pederson protocols. Specifically, the batch version of Schnorr's protocols immediately implies the same bounds for the corresponding batch identification scheme.

Corollary 1. *Assuming that the discrete-logarithm problem is second-moment hard, any adversary that runs in time t wins in the rogue soundness game for the batch Schnorr and Okamoto identification schemes with probability at most $(t^2/p)^{2/3}$, where p is the order of the underlying group.*

We extend our results for general batch Sigma protocols. We analyze the rogue-instance security of a general batch protocol with plus-one special soundness and achieve the bound of $\epsilon \leq (k^2 \cdot t^2/p)^{1/2}$, which is inferior to our bound for the specific case of algebraic protocols, but superior to previously known bounds.

Theorem 2 (Informal). *Let Π be k-batch Sigma protocol for a relation $\mathcal{R} \subseteq \mathcal{X} \times \mathcal{W}$ with plus-one special soundness. If $\mathcal{R}$ is second-moment hard with respect to a distribution $\mathcal{D}$, then Π has rogue soundness error $\epsilon(t) \leq (k^2 \cdot t^2/|\mathcal{W}|)^{1/2}$.*

In Table 1 we exemplify the concrete improvements we get in Theorem 1 and Theorem 2 for various parameter settings.

Non-interactive Proof-of-Knowledge. We construct non-interactive batch arguments from algebraic Sigma protocols by applying the Fiat-Shamir paradigm to the batch Sigma protocols. Given Theorem 1, the generic analysis of the Fiat-Shamir yields a bound on the rogue-instance game of $\epsilon \leq q \cdot (t^2/p)^{2/3}$ when considering malicious prover who runs in time t and performs at most q queries

Table 1. A comparison of the security guarantees for the batch Schnorr scheme provided by [16] compared to our bounds given in Theorem 2 and in Theorem 1.

Attacker's running time t	Security parameter λ	Batch parameter k	Bound of [16] $(t^2/p)^{1/(k+1)}$	Generic bound Theorem 2 $(k^2 \cdot t^2/p)^{1/2}$	Algebraic bound Theorem 1 $(t^2/p)^{2/3}$
2^{64}	256	2	$2^{-42.67}$	2^{-63}	$2^{-85.33}$
2^{64}	256	4	$2^{-25.6}$	2^{-62}	$2^{-85.33}$
2^{80}	256	6	$2^{-13.71}$	$2^{-45.42}$	2^{-64}
2^{80}	512	8	$2^{-39.11}$	2^{-173}	$2^{-234.66}$
2^{100}	512	16	$2^{-18.35}$	2^{-152}	2^{-208}
2^{100}	512	24	$2^{-12.48}$	$2^{-151.42}$	2^{-208}
2^{128}	512	24	$2^{-10.24}$	$2^{-123.42}$	$2^{-170.66}$
2^{128}	512	32	$2^{-7.76}$	2^{-123}	$2^{-170.66}$

to the random oracle. However, direct analysis of the rogue-instance game yields a bound of $\epsilon \leq (kq \cdot t^2/p)^{2/3}$ which is again matches the bound of Rotem and Segev [30], for a single instance. Informally, we show the following.

Theorem 3 (Informal). *Let Π be an algebraic Sigma protocol for a relation $\mathcal{R} \subseteq \mathcal{X} \times \mathcal{W}$. If $\mathcal{R}$ is second-moment hard with respect to a distribution $\mathcal{D}$, then $\mathcal{R}$ has a non-interactive batch argument with rogue soundness error $\epsilon(t) \leq (kq \cdot t^2/|\mathcal{W}|)^{2/3}$.*

Establishing Hardness for High-Moment Assumptions. Theorem 1 and Theorem 3 rely on the second-moment-hardness of a relation, an assumption introduced in [30]. While the use of these assumptions is beneficial, there is no evidence to support their hardness. To remedy the situation, we present a new framework that allows to establish bounds for oracle-algorithms with expected running time. Utilizing our framework, we achieve the first hardness result for these high-moment assumptions, relative to a oracle. The general statement of our framework is somewhat technical and is given in Theorem 2. Thus, we present two main implications of our framework, which are easier to state.

First, we establish the second-moment hardness of the discrete-logarithm problem against expected-time algorithms in the generic group model. Shoup [33] analyzed the generic hardness of the discrete-logarithm problem with respect to strict time algorithms. He showed that any generic t-time algorithm that solves the discrete-logarithm problem has success probability at most $\epsilon \leq t^2/p$. Applying our framework yields a bound of $\epsilon \leq \mathbb{E}\left[T_A^2\right]/p$ when considering *unbounded* algorithms where T_A denotes the random variable indicating the algorithm's running time.

Theorem 4 (second-moment hardness in generic group model; Informal). *For any query algorithm A, let $T_A = T_A(\lambda)$ be a random variable indicating the number of queries performed by A until he stops. For every algorithm*

A that solves the discrete-logarithm problem in a generic group of prime order p and succeeds with probability ϵ_A it holds that

$$\epsilon_A \leq \frac{\mathbb{E}\left[T_A^2\right]}{p} \quad .$$

Our framework is inspired by [19] which showed a generic framework to prove bounds with respect to expected-time algorithms when considering only the first-moment of the expected running time. Their result proves the first-moment assumption (Definition 1), but cannot be used to derive second-moment hardness. Moreover, our framework achieves tighter bounds than theirs and is arguably easier to use (see Corollary 3).

Second, we derive expected-time bounds for SNARKs in the random oracle model (ROM). We focus on the construction of Micali [24], which compiles a PCP to a SNARK in the ROM. It is known that if the underlying PCP has soundness error ϵ_{PCP}, then every malicious prover that makes at most t-queries to the random oracle can convince the verifier of a false statement with probability at most $\epsilon \leq t \cdot \epsilon_{\mathsf{PCP}} + \frac{3}{2} \cdot \frac{t^2}{2^\lambda}$ (see analysis in [8]). Using our framework, we derive the following bound.

Theorem 5 (second-moment hardness of SNARKs; Informal) *Suppose the Micali construction is instantiated for a relation $\mathcal{R}$ with a PCP with error ϵ_{PCP}, and random oracle with output length λ. Then, for every $\mathbb{x} \notin \mathcal{L}(\mathcal{R})$ and every malicious argument prover $\tilde{\mathcal{P}}$ that performs $T_{\tilde{\mathcal{P}}}$ oracle queries (as a random variable) and outputs a proof $\tilde{\pi}$ it holds that*

$$\Pr\left[\mathcal{V}^f(\mathbb{x}, \tilde{\pi}) = 1\right] \leq \mathbb{E}\left[T_{\tilde{\mathcal{P}}}\right] \cdot \epsilon_{\mathsf{PCP}} + 4 \cdot \frac{\mathbb{E}\left[T_{\tilde{\mathcal{P}}}^2\right]}{2^\lambda} \quad .$$

In Sect. 2.6, we further discuss the type of cryptographic problems relative to an oracle captured by our framework. A formal treatment of the framework, including definitions, statements, and further examples, is given in Sect. 6.1.

2 Our Techniques

We summarize the main ideas behind our results.

- In Sect. 2.1 we discuss the computational assumptions we consider in this work.
- In Sect. 2.2 we define batch Sigma protocols and extend the notion of rogue-key security for multi-signature, to rogue-instance security of batch proof-of-knowledge.
- In Sect. 2.3 we first show a general compiler from a large family of Σ-protocols to a batch Σ-protocol. Then, we show the high-level proof of the rogue-security of batch Σ-protocols constructed via the general compiler.
- In Sect. 2.4 we start by showing how to construct non-interactive batch arguments using the general compiler, then, we bound their rogue-security.

– In Sect. 2.5 we show how to apply our techniques on a general batch Σ-protocol and derive a concrete bound on their rogue-soundness error.
– In Sect. 2.6 we describe our framework for establishing high-moment hardness assumptions.

2.1 High-Moment Hardness

We begin by describing the computational assumptions that underlie our work. Let $\mathcal{R} \subseteq \mathcal{X} \times \mathcal{W}$ be a relation, where $\mathcal{X}$ is the set of instances and $\mathcal{W}$ is the set of witnesses. We note that the relation (and in fact all algorithms that will be described later on) are with respect to a setup algorithm that produces public parameters. For the simplicity of this high-level overview, we omit the public parameters (where formal definitions take them into account).

We consider distribution $\mathcal{D}$ over instance-witness pairs such that $(\mathbb{x}, \mathbb{w}) \in \mathcal{R}$. For example, the distribution can sample a discrete-logarithm challenge. Typically, the hardness of the distribution is stated with respect to strict-time algorithms, that is, algorithms that run in some fixed time t. Here, we consider hardness with respect to an algorithm where the running time, t, is a random variable. We denote by $T_{\mathcal{A},\mathcal{D}}$ the random variable indicating the running time of A on input $\mathbb{x}$ where $(\mathbb{x}, \mathbb{w}) \leftarrow \mathcal{D}$. Informally, we say that $\mathcal{R}$ is *first-moment hard* with respect to the distribution $\mathcal{D}$ if for every algorithm A, it holds that

$$\textbf{first-moment hardness:} \quad \Pr\left[(\mathbb{x}, A(\mathbb{x})) \in \mathcal{R}\right] \leq \frac{\mathbb{E}\left[T_{A,\mathcal{D}}\right]}{|\mathcal{W}|^{0.5}} \ , \tag{1}$$

where the probability is taken over $(\mathbb{x}, \mathbb{w}) \leftarrow \mathcal{D}$ and over A. The first-moment assumption is justified by the work of Jaeger and Tessaro [19]. They developed a framework for proving tight bounds on the advantage of an adversary with expected-time guarantees in generic models (a.k.a. "bad flag analysis"). In particular, they prove the first-moment hardness of the discrete-logarithm problem in the generic group model. That is, they show that every algorithm A with an expected running time $\mathbb{E}[T_A]$ computes the discrete-logarithm problem in the generic group model with probability at most $\mathbb{E}[T_A]/p^{1/2}$ (where p is the group size).

Recently, Rotem and Segev [30] have generalized this assumption for higher moments, where most important for our work is the second-moment assumption. We say that a relation is *second-moment hard* with respect to a distribution $\mathcal{D}$ if for every algorithm A it holds that

$$\textbf{second-moment hardness:} \quad \Pr\left[(\mathbb{x}, A(\mathbb{x})) \in \mathcal{R}\right] \leq \frac{\mathbb{E}\left[T^2_{A,\mathcal{D}}\right]}{|\mathcal{W}|} \ , \tag{2}$$

where the probability is taken over $(\mathbb{x}, \mathbb{w}) \leftarrow \mathcal{D}$ and the algorithm A. The hardness of the second-moment assumption does not follow from the framework of [19], and has no justification even in generic models. In order to validate this assumption, we develop a framework (see Sect. 2.6), in the spirit of [19] which

does allow us to establish bounds for second-moments. In particular, it allows us to prove the second-moment hardness of the discrete-logarithm problem in the generic group model. That is, we show that every algorithm A with an expected running time $\mathbb{E}[T_A]$ computes the discrete-logarithm problem in the generic group model with probability at most $\mathbb{E}\left[T_A^2\right]/p$.

2.2 Rogue-Instance Security for Batch Protocols

We move on to describe our notion of rogue-instance soundness for batch protocols. In a batch Σ-protocol, we are given k instance-witness pairs $(\mathbb{x}_1, \mathbb{w}_1), \ldots, (\mathbb{x}_k, \mathbb{w}_k)$. The prover consists of two algorithms $\mathbf{P} = (\mathbf{P}_1, \mathbf{P}_2)$, where $\mathbf{P}_1$ sends a message α, the verifier $\mathbf{V}$ sends a random challenge $\beta \in \mathcal{C}$, $\mathbf{P}_2$ responds with a message γ, and the verifier $\mathbf{V}$ decides whether to accept.

The standard adaptive soundness requirement considers the case where a malicious prover wishes to convince the verifier on k instances of its choice. However, we consider batch Σ-protocols with rogue-instance security, where one instance $\mathbb{x}_1$ is sampled according to a given hard distribution, and the rest of the instances $\mathbb{x}_2, \ldots, \mathbb{x}_k$ are chosen adaptively as a function of $\mathbb{x}_1$.

Specifically, a batch Σ-protocol Π has ϵ rogue-soundness error if for every malicious prover $\tilde{\mathbf{P}} = (\tilde{\mathbf{P}}_1, \tilde{\mathbf{P}}_2)$ that runs in time t it holds that

$$\Pr\left[\mathsf{RogueExp}_{\Pi}(\tilde{\mathbf{P}}, \lambda) = 1\right] \leq \epsilon(t),$$

where the experiment $\mathsf{RogueExp}_{\Pi}(\tilde{\mathbf{P}}, \lambda)$ defined as follows:

1. $(\mathbb{x}_1, \mathbb{w}_1) \leftarrow \mathcal{D}_\lambda$
2. $((\tilde{\mathbb{x}}_2, \ldots, \tilde{\mathbb{x}}_k), \alpha, \mathsf{st}) \leftarrow \tilde{\mathbf{P}}_1(\mathbb{x}_1)$
3. $\beta \leftarrow \mathcal{C}$
4. $\gamma \leftarrow \tilde{\mathbf{P}}_2(\mathsf{st}, \beta)$
5. Output $\mathbf{V}(\mathbb{x}_1, \tilde{\mathbb{x}}_2, \ldots, \tilde{\mathbb{x}}_k, \alpha, \beta, \gamma)$.

Recall that the definition above omits the setup phase, see Sect. 4 for the precise definition.

2.3 Batching Algebraic Sigma Protocols

We first describe our general compiler for batching algebraic Σ-protocols. This compiler takes an algebraic protocol (which we define next) and outputs a batch version of it (for the same relation). Then, we show the high-level proof of our (almost tight) rogue-security for the batch protocol.

Algebraic Sigma Protocols. Algebraic Σ-protocols are defined with respect to an algebraic one-way function F. The protocol is a proof-of-knowledge of a preimage of $\mathsf{F}(r)$, for randomly sampled r. It is a generalization of the *preimage protocol* presented by Cramer and Damgård [13]. Algebraic one-way functions were introduced by [11], a closely related notion to group-homomorphic one-way functions introduced by [13].

Informally, we say that a one-way function $\mathsf{F}\colon \mathcal{A}^m \to \mathcal{B}$ is algebraic if $\mathcal{A}$ and $\mathcal{B}$ are abelian cyclic groups and for every $x, x' \in \mathcal{A}^m$ it holds that $\mathsf{F}(x + x') = \mathsf{F}(x) \cdot \mathsf{F}(x')$. We say that a Σ-protocol $\Pi = (\mathbf{P}_1, \mathbf{P}_2, \mathbf{V})$ is algebraic if the protocol has the following general recipe:

1. The prover $\mathbf{P}_1$ produces a message $\alpha = \mathsf{F}(r)$ for $r \in \mathcal{A}$.
2. A challenge β is sampled from $\mathbb{Z}_p$ where p is the order of $\mathcal{A}$.
3. The prover $\mathbf{P}_2$ produces a message $\gamma = r + \beta \cdot \mathbb{w}$.
4. The verifier checks correctness by checking whether $\mathsf{F}(\gamma) \stackrel{?}{=} \alpha \cdot \mathbb{x}^{\beta}$.

General Compiler to Batch Sigma Protocols. We construct a batch Σ protocol $\Pi^* = (\mathbf{P}_\mathbf{1}^*, \mathbf{P}_\mathbf{2}^*, \mathbf{V}^*)$ from algebraic Σ-protocol by invoking the Σ-protocol k times. Specifically, given k instances, $\mathbf{P}_\mathbf{1}^*$ invokes $\mathbf{P}_1(\mathbb{x}_i)$ and produces the message α which is the multiplication of all α_i's. Then, given k challenges, $\mathbf{P}_\mathbf{2}^*$ invokes $\mathbf{P}_2$ for each challenge and produces the compressed message γ by summing the messages γ_i. More formally, given an algebraic Σ-protocol $\Pi = (\mathbf{P}_1, \mathbf{P}_2, \mathbf{V})$, we construct a batch Σ-protocol $\Pi^* = (\mathbf{P}_\mathbf{1}^*, \mathbf{P}_\mathbf{2}^*, \mathbf{V}^*)$ as follows:

1. The prover $\mathbf{P}_\mathbf{1}^*$ invokes $\alpha_i \leftarrow \mathbf{P}_1(\mathbb{x}_i)$ and produces the message $\alpha = \Pi_{i=1}^{k} \alpha_i$.
2. k challenges β_i are sampled from $\mathbb{Z}_p$ where p is the order of $\mathcal{A}$.
3. The prover $\mathbf{P}_\mathbf{2}^*$ invokes $\gamma_i \leftarrow \mathbf{P}_2(\beta_i)$ for each challenge β_i and produces the compressed message $\gamma = \sum_{i=1}^{k} \gamma_i$.
4. The verifier checks correctness by checking whether $\mathsf{F}(\gamma) \stackrel{?}{=} \alpha \cdot \Pi_{i=1}^{k} \mathbb{x}_i^{\beta_i}$.

One can observe that the completeness of Π^* follows from the homomorphic property of F. The prover-to-verifier communication is two group elements. The verifier sends k elements, but since they are all uniformly random strings, they can be easily compressed to a single group element using any pseudo-random generator (e.g., using a random oracle).

Our objective is now to bound the rogue-soundness error of Π^*. To achieve this, we consider a malicious prover $\tilde{\mathbf{P}}$ that given as input an instance $\mathbb{x}_1$ which is sampled from a distribution $\mathcal{D}$, and chooses the rest of the instances $\mathbb{x}_2, \ldots, \mathbb{x}_k$ as a function of $\mathbb{x}_1$. Its goal is to convince the verifier on $\mathbb{x}_1, \ldots, \mathbb{x}_k$. We construct an algorithm that given as input an instance $\mathbb{x}$, invokes $\tilde{\mathbf{P}}$ on $\mathbb{x}$ in order to obtain a witness for $\mathbb{x}$. Combined with the second-moment assumption, it allows us to bound $\tilde{\mathbf{P}}$'s success probability (which is the rogue-soundness error).

In order to construct A, we make use of the special soundness property of Σ-protocols. Note that if a Σ-protocol has special soundness, then our construction yields a batch protocol which has *plus-one* special soundness (i.e., given $k + 1$ accepting transcripts on k instances with a common first message and pairwise distinct challenges, one can extract all k witnesses). Obtaining $k + 1$ valid transcripts from the adversary is very costly. However, in our case, we are only interested in extracting a single witness. Thus, we define a relaxed notion called *local special soundness* that allows to extract a single witness from two specifically designed transcripts.

Local Special Soundness. Informally, a batch Σ-protocol has *local special soundness* if there exists an extractor E such that given k instances $\mathbb{x}_1, \ldots, \mathbb{x}_k$ and a pair of accepting transcripts with a common first message and only one different challenge $\beta_i \neq \beta'_i$, outputs a valid witness for $\mathbb{x}_i$. We now show that every batch Σ-protocol constructed from algebraic Σ-protocol as above, has local special soundness.

Claim 1 (Informal). *The batch Σ-protocol Π^* constructed above from algebraic Σ-protocol has local special soundness.*

Proof (Proof sketch). Consider the algorithm E which takes as input a pair of accepting transcripts $(\alpha, \beta_1, \ldots, \beta_k, \gamma)$, $(\alpha, \beta'_1, \ldots, \beta'_k, \gamma')$ such that there exists only one index j on which $\beta_j \neq \beta'_j$, defined as follows:

1. Let i^* be the index on which $\beta_{i^*} \neq \beta'_{i^*}$.
2. Output $(\gamma - \gamma')/(\beta_{i^*} - \beta'_{i^*})$ on the group $\mathbb{Z}_p$ where p is the order of $\mathcal{A}_{\mathsf{pp}}$.

The proof follows from the homomorphic property of F (see Sect. 5.1 for a complete proof).

Due to the local special soundness property, it is sufficient to construct an algorithm A that invokes $\tilde{\mathbf{P}}$ on $\mathbb{x}$ and outputs two accepting transcripts $(\alpha, \beta_1, \ldots, \beta_k, \gamma), (\alpha, \beta'_1, \ldots, \beta'_k, \gamma')$ such that $\beta_1 \neq \beta'_1$.

We reduce the problem of finding two such transcripts to the "collision game" first introduced in [12]. In more detail, we show that given an algorithm that succeeds in the collision game, we can construct an algorithm that outputs two such transcripts, which conclude extracting a witness.

The Collision Game. We consider the collision game first introduced in [12] and used in [3,18] which consists of a binary matrix $H \in \{0,1\}^{R \times N}$. The output of the game is 1 if and only if two 1-entries in the same row have been found.

Informally, the R rows correspond to the prover's randomness and the N columns correspond to the verifier's randomness. An entry of H equals 1 if and only if the corresponding transcript is accepting. Then, finding two 1-entries in the same row corresponds to finding two accepting transcripts with a common first message and distinct challenges. Therefore, an algorithm for the collision game can be transformed into an algorithm that finds two accepting transcripts, which by the local special soundness, allows extracting a witness (see Sect. 5.3 for a complete proof).

We now focus on constructing an algorithm for the collision game. In contrast to the collision game algorithm of [12] which runs in strict polynomial time, our algorithm runs in expected polynomial time. A similar approach can be found in [3,18], however, their algorithm minimizes only the first-moment of the expected running time. The collision game algorithm of [3,18] samples an entry of H, if this entry equals 1, the algorithm continues to sample the entire row till it finds another 1-entry. One can observe that the second-moment of the expected running time of this algorithm is too high to get improved bounds.

Our goal is to construct an algorithm that maximizes the trade-off between the success probability and the second-moment of the expected running time, in order to use the second-moment assumption.

Lemma 1 (Informal). *Let $H \in \{0,1\}^{R \times N}$ be a binary matrix and let ϵ be the fraction of 1-entries in H. Then, there exists an algorithm A with oracle access to H such that the following holds:*

1. *The expected number of queries performed by A to H is at most 2.*
2. *The second-moment of the expected number of queries performed by A to H is at most 4.*
3. *The probability that A succeeds in the collision game is at least $\epsilon^{1.5}$.*

Proof (Proof sketch). Let $B = \frac{1}{\sqrt{\epsilon}}$ and consider the following algorithm A:

A^H

1. Sample an entry (ρ, β) in H. If $H[\rho, \beta] = 0$, abort. Let $F = \emptyset$.
2. For every $i \in [B]$: sample without replacement entries in the same row ρ. If $H[\rho, \beta_i] = 1$, set $F \leftarrow F \cup \{\beta_i\}$.
3. If $F = \emptyset$, abort. Otherwise, choose uniformly at random an index $\beta' \in F$ and output ρ, β, β'.

Let $\mathcal{Q}_A$ be a random variable indicating the number of queries performed by A to H. For this section only, we omit the bound on the expected number of queries and refer to the second-moment only. A complete proof of the formal lemma can be found in Sect. 5.2.

By the description of A it performs 1 query to H with probability $(1 - \epsilon)$ and $(1 + B)$ queries with probability ϵ. Therefore,

$$\mathbb{E}\left[\mathcal{Q}_A^2\right] = (1-\epsilon) \cdot 1^2 + \epsilon \cdot (1+B)^2 \leq 1 + 2\sqrt{\epsilon} + 1 \leq 4 \ .$$

For now, we give a high-level overview of the proof of A's success probability. A complete proof can be found in Sect. 5.2. Assuming the first query to H was 1-entry, the algorithm continues to sample entries in the same row. Thus, if it hit a row with only one 1-entry, it succeeds in the game with probability zero. Therefore, we divide the rows by the number of 1-entries in it and look at the probability to sample such a row. Formally, for every $0 \leq d \leq N$, we let δ_d be the fraction of rows with exactly d 1-entries. Assuming the first query was 1-entry, A succeeds in the game if it finds at least one more 1-entry with B draws. Let X_d be a random variable indicating the number of 1-entries found in B draws in a row with exactly d 1-entries. Overall,

$$\Pr\left[\mathsf{CollGame}(A, H) = 1\right] \geq \sum_{d=2}^{N} \delta_d \cdot \frac{d}{N} \cdot \Pr\left[X_d \geq 1\right] \ .$$

In Sect. 5.2, we show that the above term is bounded by $\approx \epsilon^{1.5}$.

2.4 Non-interactive Batch Arguments

In the previous subsection we showed a general compiler for batching algebraic Σ-protocols and bound their rogue-soundness error. Similarly, in this subsection we refer to the non-interactive analog. We first construct non-interactive batch arguments from algebraic Σ-protocols and then bound their rogue-instance security.

Non-interactive Batch Arguments from Sigma Protocols. We show how to construct non-interactive batch arguments from algebraic Σ-protocols.

The construction is given by applying the Fiat-Shamir paradigm on the batch Σ-protocol constructed in Sect. 2.3 except for one minor change. Recall that in the construction of batch Σ-protocols, the prover is given as input k different challenges for each input. We wish to keep this property in the non-interactive analog. Specifically, we construct a non-interactive batch argument $\mathsf{NARG} = (\mathcal{P}, \mathcal{V})$ from algebraic Σ-protocol by invoking the Σ-protocol k times and obtaining the challenges from a random oracle function $f \in \mathcal{U}(\lambda)$. In more detail, given k instances, the prover $\mathcal{P}$ invokes $\alpha_i \leftarrow \mathbf{P}_1(\mathbb{x}_i)$ and computes α as the multiplication of α_i's. Then, it obtains each challenge β_i by querying $f(\mathbb{x}_1, \ldots, \mathbb{x}_k, \alpha, i)$. Finally, it invokes $\mathbf{P}_2$ for each challenge and computes γ by summing the messages γ_i. The prover $\mathcal{P}$ outputs the proof string (α, γ). The verifier $\mathcal{V}$ computes β_i by querying the random oracle f and checking whether $\mathsf{F}(\gamma) \stackrel{?}{=} \alpha \cdot \Pi_{i=1}^{k} \mathbb{x}_i^{\beta_i}$. One can observe that the completeness of NARG follows from the homomorphic property of F and that the proof size is two group elements.

Our objective now is to bound the rogue-soundness error of NARG. Similarly to the interactive case, the NARG constructed above has local special soundness. Therefore, in order to extract a witness, it suffices to construct an algorithm that outputs a pair of transcripts with a common first message and only one different challenge $\beta_i \neq \beta_i'$.

Collision Game for the Non-interactive Analog. Similar to the interactive case, our goal is to reduce the task of finding two such transcripts to the collision game. However, this transformation presents certain challenges. First, in the interactive case, we have two elements of randomness - the prover's randomness and the verifier's randomness which can be straightforwardly represented as a matrix. In contrast, in the non-interactive settings, the verifier's randomness is replaced by random oracle queries. A malicious prover performs at most q queries to the random oracle in order to obtain the challenges. Each answer from the random oracle may affect the prover's algorithm.

Secondly, in the interactive case, a prover $\mathbf{P}$ can be represented by two algorithms $\mathbf{P}_1, \mathbf{P}_2$. The algorithm $\mathbf{P}_1$ outputs the first message α and a state st, and $\mathbf{P}_2$ given as input the challenges β_i and the state st. Consequently, in order to obtain a pair of transcripts with a common first message, we can invoke $\mathbf{P}_1$ and $\mathbf{P}_2$, followed by invoking $\mathbf{P}_2$ again, on the same state and different challenges. In the non-interactive analog, a prover $\mathcal{P}$ outputs the instances $\mathbb{x}_2, \ldots, \mathbb{x}_k$ along with (α, γ). We assume without loss of generality that $\mathcal{P}$ always outputs α that

it queried the random oracle f with $(\mathbb{x}_1, \tilde{\mathbb{x}}_2, \ldots, \tilde{\mathbb{x}}_k, \alpha)$. Then, in order to obtain two transcripts with a common first message, we need to "guess" which random oracle query the prover is going to output. We invoke the prover once to obtain $(\tilde{\mathbb{x}}_2, \ldots, \tilde{\mathbb{x}}_k, \alpha, \gamma)$ and let i^* be the random oracle on which the prover queried $(\mathbb{x}_1, \tilde{\mathbb{x}}_2, \ldots, \tilde{\mathbb{x}}_k, \alpha)$. Then, we invoke the prover, replicating the same random oracle responses up to the i^*-th query. With probability $\approx 1/\mathsf{q}$ the prover outputs the same instances and first message α.

Therefore, we reduce the problem of finding two such transcripts into the "tree game". In this game, we consider a fixed randomness for the prover and consider a tree of depth q and degree 2^λ. The depth corresponds to the number of queries performed by the prover and the degree corresponds to the possible answers from the random oracle f. Consequently, the execution of the prover corresponds to a random walk on the tree and a leaf corresponds to the output of the prover. We let the value of a leaf be the random oracle query on which the prover queried f with this output. More precisely, each leaf corresponds to an output $(\mathbb{x}_2, \ldots, \mathbb{x}_k, \alpha, \gamma)$, we consider the value of a leaf to be the random oracle query in which the prover queried f with $(\mathbb{x}_2, \ldots, \mathbb{x}_k, \alpha)$. Then, finding two transcripts with a common first message and distinct challenges corresponds to finding two leaves with the same value i such that their lowest common ancestor is an internal node v of height i. A formal proof of the reduction appears in the full version.

The Tree Game. We introduce a tree game where an algorithm is given oracle access to a tree T where the value of each leaf is a number. Consider a complete tree T of depth l and degree r. Let $\mathsf{Leaves}(T)$ be the leaves of T and for every $u \in \mathsf{Leaves}(T)$ let $\mathsf{val}(u)$ be the value "stored" in u. Note that not all leaves hold a number value, we consider the value of such a leaf as $\perp$. During the execution of the game, the algorithm A is given as input a number k and oracle access to the tree T and aims to find $k+1$ leaves $u_1, \ldots, u_{k+1}$ with the same value i that have the same lowest common ancestor v such that $\mathsf{height}(v) = i$.

Due to the local special soundness property, it is sufficient to construct an algorithm that outputs two accepting transcripts, then in this section, we consider the specific case where $k = 1$.

Lemma 2 (Informal). *Let T be a complete tree of depth l and degree r and let ϵ be the fraction of non-bot leaves in T. Then, there exists an algorithm A with oracle access to T such that on input $k = 1$ the following holds:*

1. *The expected number of queries performed by A to H is at most 2.*
2. *The second-moment of the expected number of queries performed by A to H is at most 4.*
3. *The probability that A succeeds in the collision game is at least $\epsilon^{1.5}/l$.*

Proof (Proof sketch). Let $B = \frac{1}{\sqrt{\epsilon}}$ and consider the following algorithm A:

A^T

1. Sample a leaf $u \in \mathsf{Leaves}(T)$. If $\mathsf{val}(u) = \perp$, abort.
2. Let v be the parent of u of height $\mathsf{val}(u)$ and let w be the parent of u of height $(\mathsf{val}(u) - 1)$. Let $F = \emptyset$.
3. For every $i \in [B]$: sample without replacement leaves from $T_v \setminus T_w$. If $\mathsf{val}(u_i) = \mathsf{val}(u)$, set $F \leftarrow F \cup \{u_i\}$.
4. If $F = \emptyset$, abort. Otherwise, choose uniformly at random a leaf $u' \in F$ and output u, u'.

Let $\mathcal{Q}_A$ be a random variable indicating the number of queries performed by A to T. For this section only, we omit the bound on the expected number of queries and refer to the second-moment only. A complete proof of the formal lemma appears in the full version.

By the description of A it performs 1 query to T with probability $(1-\epsilon)$ and $(1+B)$ queries with probability ϵ. Therefore,

$$\mathbb{E}\left[\mathcal{Q}_A^2\right] = (1-\epsilon)\cdot 1^2 + \epsilon\cdot(1+B)^2 \leq 1 + 2\sqrt{\epsilon} + 1 \leq 4 \ .$$

For now, we give an informal high-level overview of the proof of A's success probability. A complete proof appears in the full version. Assume A samples a leaf u with the value h, then, A continues to sample leaves from the same sub-tree in order to find another leaf with the value h. Let v be the parent of u of height h. Note that for every h and v, the number of leaves with the value h in T_v may be different, which affects its success probability. Therefore, for every value h, we "divide" the internal nodes to "buckets" by the probability to sample a leaf with the value h in its sub-tree, and then we look at the probability to "reach" each bucket.

Formally, for every $0 \leq d \leq l \log r$ and $0 \leq h \leq l-1$, we let

$$\delta_{d,h} = \Pr_{v:\mathsf{height}(v)=h}\left[\frac{|\{u \in \mathsf{Leaves}(T_v) : \mathsf{val}(u) = h\}|}{|\mathsf{Leaves}(T_v)|} \in \left[2^{-d}, 2^{-d+1}\right]\right] \ .$$

Note that a node v is in the d-th "bucket" if the probability to sample a leaf with the value h in the sub-tree T_v is in $\left[2^{-d}, 2^{-d+1}\right]$. Assuming the first query to the tree is a leaf u with the value h, the remainder of the game can be modeled by a hypergeometric distribution. Informally, B elements from a population of size $|T_v \setminus T_w|$ containing $\approx 2^{-d}$ successes are drawn without replacement. Let $X_{\delta_{d,h}}$ be a random variable indicating the number of leaves with the value h found in B draws in a sub-tree T_v such that v is in the d-th "bucket". Thus,

$$\Pr\left[\mathsf{TreeCollGame}(A,T) = 1\right] \geq \sum_{h=0}^{l-1}\sum_{d=2}^{N} \delta_{d,h}\cdot 2^{-d}\cdot \Pr\left[X_{\delta_{d,h}} \geq 1\right] \ .$$

In the full version, we show that the above term is bounded by $\approx \epsilon^{1.5}/l$.

2.5 General Batch Sigma Protocols

Batch Sigma protocols. In the general case, we consider batch Σ-protocols where given k instance-witness pairs $(\mathbb{x}_i, \mathbb{w}_i)$, the prover $\mathbf{P}_1$ sends a message α, the verifier $\mathbf{V}$ samples a challenge β and sends it, the prover $\mathbf{P}_2$ responds with a message γ, and the verifier $\mathbf{V}$ decides whether to accept or reject by applying a predicate to $(\mathbb{x}_1, \ldots, \mathbb{x}_k, \alpha, \beta, \gamma)$. In order to bound the rogue-soundness error of batch Σ-protocols, we make use of the special soundness property. In particular, we consider the *plus-one special soundness* which guarantees the existence of an extractor E. When it is given as input $k+1$ transcripts of an execution of a batch Sigma protocol on k instances, the extractor outputs k corresponding witnesses. More precisely, the extractor is given as input $k+1$ transcripts with a common first message and distinct pairwise challenges.

We construct an algorithm A that given as input an instance $\mathbb{x}$ invokes a malicious prover on input $\mathbb{x}$ to obtain $k+1$ transcripts, which by the plus-one special soundness allows extracting k witnesses, specifically, to output a witness for $\mathbb{x}$. Note that the algorithm needs to invoke the prover multiple times in order to achieve approximately the same probability as in the specific case of batch protocols constructed from algebraic Σ-protocols. Unfortunately, it appears that finding a good trade-off between the second-moment of the expected running time and the success probability of the algorithm is challenging in this context. As a result, in the general case, we rely on the first-moment assumption.

Similarly, we reduce the problem of finding $k+1$ accepting transcripts to a generalized version of the collision game first introduced in [12]. In more detail, we construct an algorithm for the collision game and then use it in order to obtain $k+1$ accepting transcripts (with a common first message and pairwise distinct challenges), which conclude extracting a witness.

General Collision Game. We provide a general version of the collision game first introduced in [12] and used in [3,18], which consists of a binary matrix $H \in \{0,1\}^{R \times N}$. We generalize the collision game by an additional input, a number $k \in \mathbb{N}$. The output of the game is 1 if and only if $k+1$ entries with the value 1 in the same row have been found. An algorithm for the collision game is given as input a number $k \in \mathbb{N}$ and an oracle access to the matrix H.

Informally, the R rows correspond to the prover's randomness and the N columns correspond to the verifier's randomness. An entry of H equals 1 if and only if the corresponding transcript is accepting. Then, finding $k+1$ entries with the value 1 in the same row corresponds to finding $k+1$ accepting transcripts with a common first message and pairwise distinct challenges. Therefore, an algorithm for the collision game can be transformed into an algorithm that finds $k+1$ accepting transcripts, which as discussed above, allows extracting a witness (see the full version for a complete proof).

Lemma 3 (Informal). *Let $H \in \{0,1\}^{R \times N}$ be a binary matrix and let ϵ be the fraction of 1-entries in H. Then, there exists an algorithm A with oracle access to H such that on input k the following holds:*

1. *The expected number of queries performed by A to H is at most $k+1$.*
2. *The probability that A succeeds in the game is at least ϵ.*

Proof (Proof sketch). We consider the following algorithm:

$A^H(k)$

1. Sample an entry (ρ, β) in H. If $H[\rho, \beta] = 0$, abort.
2. Sample without replacement entries in the same row ρ, until $k+1$ entries with the value 1 are found or the row has been exhausted.

Let $\mathcal{Q}_A$ be a random variable indicating the number of queries performed by A to H. Note that the number of 1-entries in each row affects the expected number of queries performed by A. Thus, we let ϵ_ρ be the fraction of 1-entries in row ρ. Assuming the first query to H lies in row ρ and equals 1, the remainder of the algorithm can be modeled by a negative hypergeometric distribution. Elements from a population of size $N-1$ containing $\epsilon_\rho N - 1$ successes are drawn without replacement till k successes are counted. Thus, assuming that the first query lies in a row ρ and equals 1, the expected number of queries performed by A is $\frac{k(N-1+1)}{\epsilon_\rho N - 1 + 1} = \frac{k}{\epsilon_\rho}$. Overall,

$$\mathbb{E}\left[\mathcal{Q}_A\right] = 1 + \frac{1}{R}\sum_{1}^{R} \epsilon_\rho \cdot \frac{k}{\epsilon_\rho} = k+1 \ .$$

As discussed in Sect. 2.3, in order to bound the success probability we divide the rows by the number of 1-entries in it. Formally, for every $0 \le d \le N$, we let δ_d be the fraction of rows with exactly d 1-entries. Note that if A's first query to H lies in a row with at least $k+1$ entries with the value 1, it succeeds in the game with probability 1. Thus,

$$\Pr\left[\mathsf{CollGame}_k(A, H) = 1\right] \ge \sum_{d=k+1}^{R} \delta_d \cdot \frac{d}{N} \ .$$

In the full version, we show that the above term is bounded by $\approx \epsilon$.

2.6 Expected Time Hardness Framework

In this subsection, we present our framework for analyzing the expected-time hardness of cryptographic problems in generic models. Our framework allows bounding the success probability of query-algorithms in experiments that involve access to an oracle (e.g., solving discrete-logarithm in the generic group model). Here, we consider the number of queries performed by the algorithm and ignore its actual runtime.

Our overall goal is to prove statements of the form: if any algorithm that performs t queries (as a strict parameter) has success probability $\epsilon(t)$ in a particular experiment, then any algorithm A has success probability $\mathbb{E}\left[\epsilon(T_A)\right]$, where T_A is

a *random variable* for the number of queries performed by A. Such a statement would allow us to derive the desired first-moment and second-moment hardness that we need for discrete-logarithm and other problems.

Perhaps surprisingly, such a general statement is *incorrect*, which we demonstrate via the multiple discrete-logarithm problem. Yun [34] showed that any generic t-time algorithm given k instances of the discrete-logarithm problem solves all of them with probability at most $\epsilon(t) \leq (k \cdot t^2/p)^k$ (which is tight). However, this bound does not translate to $\mathbb{E}[\epsilon(T_A)] = k^k \cdot \mathbb{E}\left[T_A^{2k}\right]/p^k$. To illustrate this, consider the following generic algorithm A for the case where $k = 2$:

1. Perform $p^{1/4}$ distinct queries to the group generation oracle and store the query-answer list μ.
2. If there does not exist $(x, y), (x', y') \in \mu$, such that $x \neq x'$ and $y = y'$, abort.
3. Otherwise, perform another $p^{1/2}$ queries to the group generation oracle.

A careful analysis shows that the success probability of this algorithm is $\approx 1/\sqrt{p}$ and the 4-moment of the expected number of queries is $\approx p$, which does not satisfy the bound of $\epsilon \leq 4 \cdot \mathbb{E}\left[T_A^4\right]/p^2$.

This raises the question of when can we derive bounds for expected algorithms. What distinguishes the multiple discrete-logarithm (for which we have no non-trivial bounds for expected algorithms) compared to the single discrete-logarithm (for which we derive tight bounds for expected algorithms)? We define a (relatively natural) property of the experiment, called *history oblivious*, that can precisely distinguish the two cases and allows us to derive our bounds. Roughly speaking, history oblivious experiment is defined via the existence of a predicate on the sequence of query/answer pairs (the trace). When the predicate of the trace is true, then the algorithm is able to solve its task with no additional queries. When the predicate is false, the trace has a limited effect on its success probability (only the size of the trace affects the probability and not its contents).

For example, in the discrete-logarithm problem, the trace to the generic group would be true if it contains a collision. When the predicate is true, one can easily deduce a solution. Otherwise, the trace gives almost no helpful information to the algorithm except for specific elements which are not the discrete-logarithm. That is, in this case, the advantage only depends on the size of the trace. Any two traces of the same size for which the predicate is false yield equal success probability for the algorithm. Observe that this is not the case for multiple discrete-logarithm. Here, we have three types of interesting traces (rather than two). A trace can contain no collisions, or a single collision (from which one can deduce one discrete-logarithm but not the other), or two collisions (from which one can derive both discrete-logarithms). The predicate in this case would identify a trace with two collisions. Thus, two traces of the same size, one from the first type and one from the second type would have drastic different effect on the success probability, as in the latter it needs to solve only a single discrete-logarithm.

In summary, for any history oblivious experiment we show that:

$$\Pr[\textbf{strict} \text{ algorithms succeeds}] \leq \epsilon(t) \implies \Pr[\textbf{expected-time} \text{ algorithms succeeds}] \leq \mathbb{E}[\epsilon(t)] \ .$$

We formalize the above statement in Theorem 2. This allows us to prove first and second-moment hardness of discrete-logarithm Eqs. 1 and 2, which are the basis for our results. It also allows us to derive our bounds for the Micali SNARK construction given in Theorem 5. Our framework is inspired by the work of Jaeger and Tessaro [19], however, their tools do not allow us to prove the second-moment hardness assumptions in generic models. Furthermore, our approach is arguably simpler to use and provides tighter security bounds even for first-moment assumptions. We show that our framework recovers the bounds of [19] in Corollary 3.

3 Preliminaries

For any $n \in \mathbb{N}$, we denote the set of all positive integers up to n as $[n] := \{1, \ldots, n\}$. For any finite set S, $x \leftarrow S$ denotes a uniformly random element x from the set S. Similarly, for any distribution $\mathcal{D}$, $x \leftarrow \mathcal{D}$ denotes an element x drawn from distribution $\mathcal{D}$.

3.1 High-Moment Hardness

A relation $\mathcal{R}$ is a set $\mathcal{R} = \{\mathcal{R}_\lambda\}_{\lambda\in\mathbb{N}}$, where $\mathcal{R}_\lambda \subseteq \mathcal{P}_\lambda \times \mathcal{X}_\lambda \times \mathcal{W}_\lambda$ for any $\lambda \in \mathbb{N}$, for sets $\mathcal{X} = \{\mathcal{X}_\lambda\}_{\lambda\in\mathbb{N}}$, $\mathcal{W} = \{\mathcal{W}_\lambda\}_{\lambda\in\mathbb{N}}$ and $\mathcal{P} = \{\mathcal{P}_\lambda\}_{\lambda\in\mathbb{N}}$. The corresponding language $\mathcal{L}(\mathcal{R}_\lambda)$ is the set of public parameters pp and instances $\mathbb{x}$ for which there exists a witness $\mathbb{w}$ such that $(\mathsf{pp}, \mathbb{x}, \mathbb{w}) \in \mathcal{R}_\lambda$.

We consider distributions $\mathcal{D} = \{\mathcal{D}_\lambda\}_{\lambda\in\mathbb{N}}$ over the relation where each $\mathcal{D}_\lambda$ produces $(\mathsf{pp}, \mathbb{x}, \mathbb{w}) \in \mathcal{R}_\lambda$. We note by $\mathcal{D}_\lambda(\mathsf{pp})$ the distribution that produces $(\mathbb{x}, \mathbb{w})$ such that $(\mathsf{pp}, \mathbb{x}, \mathbb{w}) \in \mathcal{R}_\lambda$.

For any such distribution $\mathcal{D}_\lambda(\mathsf{pp})$ and an algorithm A, we denote by $T_{\mathcal{A},\mathcal{D}_\lambda}$ the random variable indicating the running time of A on input $\mathbb{x}$ where $(\mathbb{x}, \mathbb{w}) \leftarrow \mathcal{D}_\lambda(\mathsf{pp})$.

Definition 1 (First-moment hard relation). *Let $\Delta = \Delta(\lambda), \omega = \omega(\lambda)$ be functions of the security parameter, and let $\mathcal{R} = \{\mathcal{R}_\lambda\}_{\lambda\in\mathbb{N}}$ be a relation where $\mathcal{R}_\lambda \subseteq \mathcal{P}_\lambda \times \mathcal{X}_\lambda \times \mathcal{W}_\lambda$. Let Setup be a setup algorithm that on input 1^λ, outputs $\mathsf{pp} \in \mathcal{P}_\lambda$. We say that $\mathcal{R}$ is first-moment hard (with respect to a distribution $\mathcal{D} = \{\mathcal{D}_\lambda\}_{\lambda\in\mathbb{N}}$ and a setup algorithm Setup) if for every algorithm A and for every $\lambda \in \mathbb{N}$ it holds that*

$$\Pr\left[(\mathsf{pp}, \mathbb{x}, \tilde{\mathbb{w}}) \in \mathcal{R}_\lambda \;\middle|\; \begin{array}{l} \mathsf{pp} \leftarrow \mathsf{Setup}(1^\lambda) \\ (\mathbb{x}, \mathbb{w}) \leftarrow \mathcal{D}_\lambda(\mathsf{pp}) \\ \tilde{\mathbb{w}} \leftarrow A(\mathsf{pp}, \mathbb{x}) \end{array}\right] \leq \frac{\Delta \cdot \mathbb{E}\left[T_{A,\mathcal{D}_\lambda}\right]}{|\mathcal{W}_\lambda|^\omega} \quad .$$

Definition 2 (Second-moment hard relation). *Let $\Delta = \Delta(\lambda), \omega = \omega(\lambda)$ be functions of the security parameter, and let $\mathcal{R} = \{\mathcal{R}_\lambda\}_{\lambda\in\mathbb{N}}$ be a relation where $\mathcal{R}_\lambda \subseteq \mathcal{P}_\lambda \times \mathcal{X}_\lambda \times \mathcal{W}_\lambda$. Let Setup be a setup algorithm that on input 1^λ, outputs $\mathsf{pp} \in \mathcal{P}_\lambda$. We say that $\mathcal{R}$ is second-moment hard (with respect to a distribution*

$\mathcal{D} = \{\mathcal{D}_\lambda\}_{\lambda\in\mathbb{N}}$ *and a setup algorithm* Setup*) if for every algorithm* A *and for every* $\lambda \in \mathbb{N}$ *it holds that*

$$\Pr\left[(\mathsf{pp},\mathbb{x},\tilde{\mathbb{w}}) \in \mathcal{R}_\lambda \;\middle|\; \begin{array}{l} \mathsf{pp} \leftarrow \mathsf{Setup}(1^\lambda) \\ (\mathbb{x},\mathbb{w}) \leftarrow \mathcal{D}_\lambda(\mathsf{pp}) \\ \tilde{\mathbb{w}} \leftarrow A(\mathsf{pp},\mathbb{x}) \end{array}\right] \leq \frac{\Delta \cdot \mathbb{E}\left[T^2_{A,\mathcal{D}_\lambda}\right]}{|\mathcal{W}_\lambda|^\omega} \ .$$

3.2 Sigma Protocols

Definition 3 (Σ-Protocol). *Let* $\mathcal{R} = \{\mathcal{R}_\lambda\}_{\lambda\in\mathbb{N}}$ *be a relation, where* $\mathcal{R}_\lambda \subseteq \mathcal{P}_\lambda \times \mathcal{X}_\lambda \times \mathcal{W}_\lambda$ *for any* $\lambda \in \mathbb{N}$*. A* Σ*-protocol* Π *for relation* $\mathcal{R}$ *is a* 5*-tuple* $(\mathsf{Setup}, \boldsymbol{P}_1, \boldsymbol{P}_2, \boldsymbol{V}, \mathcal{C})$ *where* Setup *and* $\boldsymbol{P}_1$ *are probabilistic polynomial-time algorithms,* $\boldsymbol{P}_2$ *and* $\boldsymbol{V}$ *are deterministic polynomial-time algorithms, and* $\mathcal{C} = \{\mathcal{C}_{\mathsf{pp}}\}_{\mathsf{pp}\in\mathcal{P}}$ *is an ensemble of efficiently sampleable sets. The protocol* Π *is defined as follows:*

1. *The algorithm* $\mathsf{Setup}(1^\lambda)$ *produces public parameters* pp*.*
2. *The algorithm* $\boldsymbol{P}_1(\mathsf{pp},\mathbb{x},\mathbb{w})$ *produces a message* α *and a state* st*.*
3. *A challenge* β *is sampled uniformly at random from the challenge set* $\mathcal{C}_{\mathsf{pp}}$*.*
4. *The algorithm* $\boldsymbol{P}_2(\mathsf{st},\beta)$ *produces a message* γ*.*
5. *The algorithm* $\boldsymbol{V}(\mathsf{pp},\mathbb{x},\alpha,\beta,\gamma)$ *determines the output of the protocol by outputting* 0 *or* 1*.*

We require that for every $\lambda \in \mathbb{N}$ *and* $(\mathbb{x},\mathbb{w}) \in \mathcal{R}_\lambda$ *it holds that*

$$\Pr\left[\boldsymbol{V}(\mathsf{pp},\mathbb{x},\alpha,\beta,\gamma) = 1 \;\middle|\; \begin{array}{l} \mathsf{pp} \leftarrow \mathsf{Setup}(1^\lambda) \\ (\alpha,\mathsf{st}) \leftarrow \boldsymbol{P}_1(\mathsf{pp},\mathbb{x},\mathbb{w}) \\ \beta \leftarrow \mathcal{C}_{\mathsf{pp}} \\ \gamma \leftarrow \boldsymbol{P}_2(\mathsf{st},\beta) \end{array}\right] = 1 \ .$$

Definition 4 (Special soundness). *Let* $\Pi = (\mathsf{Setup}, \boldsymbol{P}_1, \boldsymbol{P}_2, \boldsymbol{V}, \mathcal{C})$ *be a* Σ*-protocol for a relation* $\mathcal{R}$*, and let* $t = t(\lambda)$ *be a function of the security parameter* $\lambda \in \mathbb{N}$*. Then,* Π *has* t*-time special soundness if there exists a deterministic* t*-time algorithm* E *that on any public parameters* $\mathsf{pp} \in \mathcal{P}$*, any input statement* $\mathbb{x} \in \mathcal{X}_\lambda$ *and any two accepting transcripts with a common first message and distinct challenges, outputs a witness* $\mathbb{w}$ *such that* $(\mathsf{pp},\mathbb{x},\mathbb{w}) \in \mathcal{R}$*.*

Definition 5 (Zero knowledge Σ-protocol). *Let* $\Pi = (\mathsf{Setup}, \boldsymbol{P}_1, \boldsymbol{P}_2, \boldsymbol{V}, \mathcal{C})$ *be a* Σ*-protocol for a relation* $\mathcal{R}$*, and let* $t = t(\lambda)$ *be a function of the security parameter* $\lambda \in \mathbb{N}$*. Then,* Π *is* t*-time zero-knowledge if there exists a probabilistic* t*-time algorithm* Sim *such that for every* $\lambda \in \mathbb{N}$ *and public parameters-instance-witness tuple* $(\mathsf{pp},\mathbb{x},\mathbb{w}) \in \mathcal{R}_\lambda$ *the distributions*

$$\left\{(\mathsf{pp},\mathbb{x},\alpha,\beta,\gamma) \;\middle|\; \begin{array}{l} (\alpha,\mathsf{st}) \leftarrow \boldsymbol{P}_1(\mathsf{pp},\mathbb{x},\mathbb{w}) \\ \beta \leftarrow \mathcal{C}_{\mathsf{pp}} \\ \gamma \leftarrow \boldsymbol{P}_2(\mathsf{st},\beta) \end{array}\right\} \qquad \textit{and} \qquad \{\mathsf{Sim}(\mathsf{pp},\mathbb{x})\}$$

are identical.

3.3 Batch Sigma Protocols

Definition 6 (Batch Σ-protocol). *Let $\mathcal{R} = \{\mathcal{R}_\lambda\}_{\lambda\in\mathbb{N}}$ be a relation, where $\mathcal{R}_\lambda \subseteq \mathcal{P}_\lambda \times \mathcal{X}_\lambda \times \mathcal{W}_\lambda$ for any $\lambda \in \mathbb{N}$ and let $\boldsymbol{K} \in \mathbb{N}$ be a bound on the number of instances. A batch Σ-protocol Π for relation $\mathcal{R}$ is a 5-tuple $(\mathsf{Setup}, \boldsymbol{P}_1, \boldsymbol{P}_2, \boldsymbol{V}, \mathcal{C})$ where Setup and $\boldsymbol{P}_1$ are probabilistic polynomial-time algorithms, $\boldsymbol{P}_2$ and $\boldsymbol{V}$ are deterministic polynomial-time algorithms, and $\mathcal{C} = \{\mathcal{C}_{\mathsf{pp}}\}_{\mathsf{pp}\in\mathcal{P}}$ is an ensemble of efficiently sampleable sets. For any $k \leq \boldsymbol{K}$, the protocol Π is defined as follows:*

1. *The algorithm $\boldsymbol{P}_1(\mathsf{pp}, (\mathbb{x}_1, \mathbb{w}_1), \ldots, (\mathbb{x}_k, \mathbb{w}_k))$ produces a message α and a state st.*
2. *A challenge β is sampled uniformly at random from the challenge set $\mathcal{C}_{\mathsf{pp}}$.*
3. *The algorithm $\boldsymbol{P}_2(\mathsf{st}, \beta)$ produces a message γ.*
4. *The algorithm $\boldsymbol{V}(\mathsf{pp}, \mathbb{x}_1, \ldots, \mathbb{x}_k, \alpha, \beta, \gamma)$ determines the output of the protocol by outputting 0 or 1.*

We require that for every $\lambda, k \in \mathbb{N}$ such that $k \leq \boldsymbol{K}$, for any $(\mathbb{x}_1, \mathbb{w}_1), \ldots, (\mathbb{x}_k, \mathbb{w}_k) \in \mathcal{R}_\lambda$ it holds that

$$\Pr\left[\boldsymbol{V}(\mathsf{pp}, \mathbb{x}_1, \ldots, \mathbb{x}_k, \alpha, \beta, \gamma) = 1 \;\middle|\; \begin{array}{l} \mathsf{pp} \leftarrow \mathsf{Setup}(1^\lambda, \boldsymbol{K}) \\ (\alpha, \mathsf{st}) \leftarrow \boldsymbol{P}_1(\mathsf{pp}, (\mathbb{x}_1, \mathbb{w}_1), \ldots, (\mathbb{x}_k, \mathbb{w}_k)) \\ \beta \leftarrow \mathcal{C}_{\mathsf{pp}} \\ \gamma \leftarrow \boldsymbol{P}_2(\mathsf{st}, \beta) \end{array}\right] = 1 \ .$$

Definition 7 (Plus-one special soundness). *Let $\Pi = (\mathsf{Setup}, \boldsymbol{P}_1, \boldsymbol{P}_2, \boldsymbol{V}, \mathcal{C})$ be a batch Σ-protocol for a relation $\mathcal{R}$ with a bound $\boldsymbol{K}$ on the number of instances, and let $t = t(\lambda, \boldsymbol{K})$ be a function of $\boldsymbol{K}$ and the security parameter $\lambda \in \mathbb{N}$. Then, Π has t-time plus-one special soundness if there exists a deterministic t-time algorithm E that for every $\lambda \in \mathbb{N}$ and $k \leq \boldsymbol{K}$, on any public parameters pp, any k inputs statements $\mathbb{x}_1, \ldots, \mathbb{x}_k \in \mathcal{X}_\lambda$ and any $k+1$ accepting transcripts with a common first message and pairwise distinct challenges, outputs k witnesses $\mathbb{w}_1, \ldots, \mathbb{w}_k$ such that for every $i \in [k]$ it holds that $(\mathsf{pp}, \mathbb{x}_i, \mathbb{w}_i) \in \mathcal{R}_\lambda$.*

Definition 8 (Zero knowledge batch Σ-protocol). *Let $\Pi = (\mathsf{Setup}, \boldsymbol{P}_1, \boldsymbol{P}_2, \boldsymbol{V}, \mathcal{C})$ be a batch Σ-protocol for a relation $\mathcal{R}$ with a bound $\boldsymbol{K}$ on the number of instances, and let $t = t(\lambda, \boldsymbol{K})$ be a function of $\boldsymbol{K}$ and the security parameter $\lambda \in \mathbb{N}$. Then, Π is t-time zero-knowledge if there exists a probabilistic t-time algorithm Sim such that for any $k \leq \boldsymbol{K}$, for every $\lambda \in \mathbb{N}$ and $(\mathsf{pp}, \mathbb{x}_1, \mathbb{w}_1), \ldots, (\mathsf{pp}, \mathbb{x}_k, \mathbb{w}_k) \in \mathcal{R}_\lambda$ the distributions*

$$\left\{(\mathsf{pp}, \mathbb{x}_1, \ldots, \mathbb{x}_k, \alpha, \beta, \gamma) \;\middle|\; \begin{array}{l} (\alpha, \mathsf{st}) \leftarrow \boldsymbol{P}_1(\mathsf{pp}, (\mathbb{x}_1, \mathbb{w}_1), \ldots, (\mathbb{x}_k, \mathbb{w}_k)) \\ \beta \leftarrow \mathcal{C}_{k,\lambda} \\ \gamma \leftarrow \boldsymbol{P}_2(\mathsf{st}, \beta) \end{array}\right\} \quad \textit{and} \quad \{\mathsf{Sim}(\mathsf{pp}, \mathbb{x}_1, \ldots, \mathbb{x}_k)\}$$

are identical.

4 Rogue-Instance Security

In this section, we give our definition of rogue-instance security notion for batch protocols and non-interactive batch arguments, which is inspired by the rogue-key security notion for multi-signatures.

4.1 Batch Sigma Protocols

In a batch Σ-protocol, we are given k instance-witness pairs $(\mathbb{x}_1, \mathbb{w}_1), \ldots, (\mathbb{x}_k, \mathbb{w}_k)$. The standard adaptive soundness requirement considers the case where a malicious prover wishes to convince the verifier on k instances of its choice. However, we consider batch Σ-protocols with rogue-instance security, where one instance $\mathbb{x}_1$ is sampled according to a given hard distribution, and the rest of the instances $\mathbb{x}_2, \ldots, \mathbb{x}_k$ are chosen adaptively as a function of $\mathbb{x}_1$. Formally,

Definition 9 (Rogue soundness). *Let $\Pi = (\mathsf{Setup}, \boldsymbol{P}_1, \boldsymbol{P}_2, \boldsymbol{V}, \mathcal{C})$ be a batch Σ-protocol for a relation $\mathcal{R}$ with a bound $\boldsymbol{K}$ on the number of instances. Then, Π has $(t, \epsilon_{\mathcal{D}})$-rogue soundness (with respect to a distribution $\mathcal{D} = \{\mathcal{D}_\lambda\}_{\lambda \in \mathbb{N}}$ and the setup algorithm Setup) if for every $\lambda, k \in \mathbb{N}$ such that $k \leq \boldsymbol{K}$ and for any t-time malicious prover $\tilde{\boldsymbol{P}} = (\tilde{\boldsymbol{P}}_1, \tilde{\boldsymbol{P}}_2)$:*

$$\Pr\left[\boldsymbol{V}(\mathsf{pp}, \mathbb{x}_1, \tilde{\mathbb{x}}_2, \ldots, \tilde{\mathbb{x}}_k, \alpha, \beta, \gamma) = 1 \;\middle|\; \begin{array}{l} \mathsf{pp} \leftarrow \mathsf{Setup}(1^\lambda, \boldsymbol{K}) \\ (\mathbb{x}_1, \mathbb{w}_1) \leftarrow \mathcal{D}_\lambda(\mathsf{pp}) \\ ((\tilde{\mathbb{x}}_2, \ldots, \tilde{\mathbb{x}}_k), \alpha, \mathsf{st}) \leftarrow \tilde{\boldsymbol{P}}_1(\mathsf{pp}, \mathbb{x}_1) \\ \beta \leftarrow \mathcal{C}_{\mathsf{pp}} \\ \gamma \leftarrow \tilde{\boldsymbol{P}}_2(\mathsf{st}, \beta) \end{array}\right] \leq \epsilon_{\mathcal{D}}(\lambda, t, \boldsymbol{K}) \ .$$

In the full version of the paper, we provide an analogous non-interactive definition.

5 Batching Algebraic Sigma Protocols

In this section, we define algebraic Σ-protocols and construct their batch version. Then, we bound the rogue-soundness error of such batch Σ-protocols using the second-moment assumption (Definition 2).

In Sect. 5.1 we define algebraic one-way functions and construct batch Σ-protocols from algebraic Σ-protocols. Then, in Sect. 5.2 we generalize the "collision game" presented in [3,12,18] for multiple instances while referring to the second-moment of the expected running time. Finally, in Sect. 5.3 we prove the rogue-instance security of batch Σ-protocols constructed from algebraic Σ-protocols.

5.1 Algebraic Sigma Protocols

In this section, we refer to Σ-protocols that have a specific structure we call *algebraic Σ-protocols* and then, we define their batch analog.

Our definition of algebraic Σ-protocols relies on algebraic one-way function, presented in [11,13].

Definition 10 (Algebraic one-way function). *A family of m-variate one-way functions consists of two algorithms $(\mathsf{Setup}, \mathsf{F})$ that work as follows. On input 1^λ, the algorithm $\mathsf{Setup}(1^\lambda)$ produces public parameters. Any such public parameters pp, determines the function $\mathsf{F}_{\mathsf{pp}} \colon \mathcal{A}^m_{\mathsf{pp}} \to \mathcal{B}_{\mathsf{pp}}$ such that for every $x \in \mathcal{A}^m_{\mathsf{pp}}$, it is efficient to compute $\mathsf{F}_{\mathsf{pp}}(x)$. A family of one-way functions is algebraic if for every $\lambda \in \mathbb{N}$ and $\mathsf{pp} \leftarrow \mathsf{Setup}(1^\lambda)$ the following holds:*

- ***Algebraic:*** *The sets* $\mathcal{A}_{\mathsf{pp}}, \mathcal{B}_{\mathsf{pp}}$ *are abelian cyclic groups with operators* $(+)$*, and* $(\cdot)$*, respectively.*
- ***Homomorphic****: For any input* $x, x' \in \mathcal{A}_{\mathsf{pp}}^m$ *it holds that* $\mathsf{F}(x + x') = \mathsf{F}(x) \cdot \mathsf{F}(x')$*.*

We now define the notion of algebraic Σ-protocols, which is a generalization of the *preimage protocol* presented in [13].

Definition 11 (Algebraic Σ-protocol). *Let* $\mathcal{R} = \{\mathcal{R}_\lambda\}_{\lambda \in \mathbb{N}}$ *be a relation, where* $\mathcal{R}_\lambda \subseteq \mathcal{P}_\lambda \times \mathcal{X}_\lambda \times \mathcal{W}_\lambda$ *for any* $\lambda \in \mathbb{N}$*. A* Σ*-protocol* $\Pi = (\mathsf{Setup}, \boldsymbol{P}_1, \boldsymbol{P}_2, \boldsymbol{V}, \mathcal{C})$ *for relation* $\mathcal{R}$ *is algebraic if there exists* m*-variate algebraic one-way function* $(\mathsf{Setup}, \mathsf{F})$ *such that for every* pp *produced by* $\mathsf{Setup}(1^\lambda)$ *the following holds:*

- *For every* $\mathbb{x}, \mathbb{w}$ *it holds that* $(\mathsf{pp}, \mathbb{x}, \mathbb{w}) \in \mathcal{R}_\lambda$ *if and only if* $\mathsf{F}_{\mathsf{pp}}(\mathbb{w}) = \mathbb{x}$*.*
- *The challenge space* $\mathcal{C}_{\mathsf{pp}} \subseteq \mathbb{Z}_p$ *where* p *is the order of* $\mathcal{A}_{\mathsf{pp}}$*.*
- *The protocol* Π *is defined as follows:*
  1. *The algorithm* $\boldsymbol{P}_1(\mathbb{x}, \mathbb{w})$ *produces a message* $\alpha = \mathsf{F}(r)$ *for some* $r \in \mathcal{A}_{\mathsf{pp}}$ *and a state* st*.*
  2. *A challenge* β *is sampled uniformly at random from the challenge set* $\mathcal{C}_{\mathsf{pp}}$*.*
  3. *The algorithm* $\boldsymbol{P}_2(\mathsf{st}, \beta)$ *produces a message* $\gamma = r + \beta \cdot \mathbb{w}$*.*
  4. *The algorithm* $\boldsymbol{V}(\mathbb{x}, \alpha, \beta, \gamma)$ *determines the output of the protocol by checking whether* $\mathsf{F}(\gamma) \stackrel{?}{=} \alpha \cdot \mathbb{x}^\beta$*.*

Note that the setup algorithm of the function is the setup algorithm of the protocol. In fact, the prover holds a public parameters-instance-witness tuple such that $\mathbb{x} = \mathsf{F}_{\mathsf{pp}}(\mathbb{w})$. Thus, the prover convinces the verifier that it knows the preimage of $\mathbb{x}$. Note that the verifier's computation can be performed using exponentiation by squaring, however there may exist more efficient algorithms.

Next, we construct a batch version of any algebraic Σ-protocol as follows.

Construction 1 (Batch Σ-protocol). *Let* $\mathcal{R} = \{\mathcal{R}_\lambda\}_{\lambda \in \mathbb{N}}$ *be a relation, where* $\mathcal{R}_\lambda \subseteq \mathcal{P}_\lambda \times \mathcal{X}_\lambda \times \mathcal{W}_\lambda$ *for any* $\lambda \in \mathbb{N}$ *and let* $\boldsymbol{K} \in \mathbb{N}$ *be a bound on the number of instances. Let* $\Pi = (\mathsf{Setup}, \boldsymbol{P}_1, \boldsymbol{P}_2, \boldsymbol{V}, \mathcal{C})$ *be an algebraic* Σ*-protocol with an algebaric one-way function* $(\mathsf{Setup}, \mathsf{F})$*. We define* $\Pi^* = (\mathsf{Setup}^*, \mathbf{P}_1^*, \mathbf{P}_2^*, \mathbf{V}^*, \mathcal{C})$ *to be a batch* Σ*-protocol for relation* $\mathcal{R}$ *as follows. The algorithms* Setup^* *and* $\mathbf{P}_1^*$ *are probabilistic polynomial-time algorithms,* $\mathbf{P}_2^*$ *and* $\mathbf{V}^*$ *are deterministic polynomial-time algorithms, and* $\mathcal{C} = \{\mathcal{C}_{\mathsf{pp}}\}_{\mathsf{pp} \in \mathcal{P}}$ *is an ensemble of efficiently sampleable sets. For every* $k \leq \boldsymbol{K}$ *the protocol is defined as follows:*

1. *The algorithm* $\mathsf{Setup}^*(1^\lambda, \boldsymbol{K})$ *is the same algorithm as* $\mathsf{Setup}(1^\lambda)$*.*
2. *The algorithm* $\mathbf{P}_1^*(\mathsf{pp}, (\mathbb{x}_1, \mathbb{w}_1), \ldots, (\mathbb{x}_k, \mathbb{w}_k))$ *invokes* $(R_i, \mathsf{st}_i) \leftarrow \boldsymbol{P}_1(\mathsf{pp}, \mathbb{x}_i, \mathbb{w}_i)$ *for every* $i \in [k]$ *and produces a message* $\alpha = \Pi_{i=1}^k R_i$ *and a state* $\mathsf{st} = (\mathsf{st}_1 \| \ldots \| \mathsf{st}_k)$*.*
3. k *different challenges* $\beta_1, \ldots, \beta_k$ *are sampled uniformly at random from the challenge set* $\mathcal{C}_{\mathsf{pp}}$*.*
4. *The algorithm* $\mathbf{P}_2^*(\mathsf{st}, \beta_1, \ldots, \beta_k)$ *parses* $\mathsf{st} = (\mathsf{st}_1 \| \ldots \| \mathsf{st}_k)$*, invokes* $\gamma_i \leftarrow \boldsymbol{P}_2(\mathsf{st}_i, \beta_i)$ *and produces a message* $\gamma = \sum_{i=1}^k \gamma_i$*.*

5. *The algorithm* $\boldsymbol{V}(\mathsf{pp}, \mathbb{x}_1, \ldots, \mathbb{x}_k, \alpha, \beta, \gamma)$ *determines the output of the protocol checking whether* $\mathsf{F}(\gamma) \stackrel{?}{=} \alpha \cdot \Pi_{i=1}^{k} \mathbb{x}_i^{\beta_i}$.

Note that the completeness of the protocol above follows from the homomorphic property of F and that the prover-to-verifier communication is two-group elements. The verifier sends k elements, but since they are all uniformly random strings, they can be easily compressed to a single group element using any pseudo-random generator (e.g., using a random oracle).

Definition 12 (Local special soundness). *Let* $\Pi = (\mathsf{Setup}, \boldsymbol{P}_1, \boldsymbol{P}_2, \boldsymbol{V}, \mathcal{C})$ *be an algebraic* Σ*-protocol for a relation* $\mathcal{R}$ *and let* Π^* *be the batch* Σ*-protocol defined in Lemma 1 with a bound* $\boldsymbol{K}$ *on the number of instances. Then,* Π^* *has local special soundness if there exists a deterministic polynomial time algorithm* E *that for every* $\lambda \in \mathbb{N}$ *and* $k \leq \boldsymbol{K}$*, given public parameters* pp*, any* k *inputs statements* $\mathbb{x}_1, \ldots, \mathbb{x}_k \in \mathcal{X}_\lambda$ *and any pair of accepting transcripts* $(\alpha, \beta_1, \ldots, \beta_k, \gamma), (\alpha, \beta'_1, \ldots, \beta'_k, \gamma')$ *such that there exists only one index* j *on which* $\beta_j \neq \beta'_j$*, outputs a witness* $\mathbb{w}_j$ *such that* $(\mathbb{x}_j, \mathbb{w}_j) \in \mathcal{R}_\lambda$.

We now show that every batch Σ-protocol defined in Lemma 1 has local special soundness.

Claim 2. *Let* $\Pi = (\mathsf{Setup}, \boldsymbol{P}_1, \boldsymbol{P}_2, \boldsymbol{V}, \mathcal{C})$ *be an algebraic* Σ*-protocol for a relation* $\mathcal{R}$ *and let* Π^* *be the batch* Σ*-protocol constructed from* Π *as defined in Lemma 1 with a bound* $\boldsymbol{K}$ *on the number of instances. Then,* Π^* *has local special soundness.*

Proof. Consider the algorithm E which takes as input public parameters pp, instances $\mathbb{x}_1, \ldots, \mathbb{x}_k$ and a pair of accepting transcripts $(\alpha, \beta_1, \ldots, \beta_k, \gamma)$, $(\alpha, \beta'_1, \ldots, \beta'_k, \gamma')$ such that there exists only one index j on which $\beta_j \neq \beta'_j$, defined as follows:

1. Let i^* be the index on which $\beta_{i^*} \neq \beta'_{i^*}$.
2. Output $(\gamma - \gamma')/(\beta_{i^*} - \beta'_{i^*})$ on the group $\mathbb{Z}_p$ where p is the order of $\mathcal{A}_{\mathsf{pp}}$.

Observe that since the two transcripts are accepting it holds that

$$\mathsf{F}_{\mathsf{pp}}(\gamma) = \alpha \cdot \Pi_{i=1}^{k} \mathbb{x}_i^{\beta_i} \qquad \text{and} \qquad \mathsf{F}_{\mathsf{pp}}(\gamma') = \alpha \cdot \Pi_{i=1}^{k} \mathbb{x}_i^{\beta'_i} \ .$$

Since $\beta_i = \beta'_i$ for every $i \neq i^*$, it holds that

$$\mathbb{x}_{i^*}^{\beta_{i^*}} \cdot \mathsf{F}_{\mathsf{pp}}(\gamma') = \mathbb{x}_{i^*}^{\beta'_{i^*}} \cdot \mathsf{F}_{\mathsf{pp}}(\gamma) \ .$$

Note that $\mathbb{x}_{i^*} = \mathsf{F}_{\mathsf{pp}}(\mathbb{w}_{i^*})$, therefore, by the homomorphic property, it holds that

$$\mathsf{F}_{\mathsf{pp}}((\beta_{i^*} - \beta'_{i^*})\mathbb{w}_{i^*}) = \mathsf{F}_{\mathsf{pp}}(\gamma - \gamma') \ .$$

Thus, $(\gamma - \gamma')/(\beta_{i^*} - \beta'_{i^*})$ is a preimage of $\mathbb{x}_{i^*}$, i.e., a valid witness for $\mathbb{x}_{i^*}$. The extractor E performs only three group operations, therefore, Π^* has local special soundness.

In Sect. 5.3, we show a concrete bound on the rogue soundness error of batch Σ-protocols defined in Lemma 1. Formally, we prove the following.

Theorem 1. *Let $\Delta = \Delta(\lambda), \omega = \omega(\lambda), t_{\tilde{\boldsymbol{P}}} = t_{\tilde{\boldsymbol{P}}}(\lambda, \boldsymbol{K}), t_{\boldsymbol{V}} = t_{\boldsymbol{V}}(\lambda, \boldsymbol{K}), t_W = t_W(\lambda, \boldsymbol{K})$ be functions of the security parameter $\lambda \in \mathbb{N}$ and the bound on the number of instances $\boldsymbol{K} \in \mathbb{N}$. Let Π be an algebraic Σ-protocol for a relation $\mathcal{R}$ and let $\Pi^* = (\mathsf{Setup}, \boldsymbol{P}_1, \boldsymbol{P}_2, \boldsymbol{V}, \mathcal{C})$ be the batch Σ-protocol constructed from Π as defined in Lemma 1. If $\mathcal{R}$ is second-moment hard with respect to a distribution $\mathcal{D}$ and the setup algorithm Setup, then Π^* has $(t_{\tilde{\boldsymbol{P}}}, \epsilon)$-rogue soundness error such that*

$$\epsilon_{\mathcal{D}}(\lambda, t_{\tilde{\boldsymbol{P}}}, t_{\boldsymbol{V}}, t_W, \boldsymbol{K}) \leq \left(\frac{\Delta \cdot 32 \cdot (t_{\tilde{\boldsymbol{P}}} + t_{\boldsymbol{V}} + t_W)^2}{|\mathcal{W}_\lambda|^\omega} \right)^{2/3} + \frac{4}{|\mathcal{C}_{\mathsf{pp}}|} \ ,$$

where $t_{\boldsymbol{V}}$ denotes the running time of the verifier $\boldsymbol{V}$ and t_W denotes the running time of the witness extractor.

5.2 The Collision Game

Similar to the collision game presented in [3,12,18], we consider a binary matrix $H \in \{0,1\}^{R \times N}$. However, instead of looking for two 1-entries in the same row, the generalized algorithm A is given as input a number $k \in \mathbb{N}$ and oracle access to the matrix and its goal is to find $k+1$ entries with the value 1 in the same row in H. Formally, the game is constructed as follows:

$\mathsf{CollGame}_k(A, H)$

1. The algorithm $A(k)$ is given oracle access to H and outputs ρ and $\beta_1, \ldots, \beta_{k+1}$.
2. The output of the game is 1 if and only if $H[\rho, \beta_1] = \ldots = H[\rho, \beta_{k+1}] = 1$ and $\beta_1, \ldots, \beta_{k+1}$ are distinct.

In particular, in this section, we refer to the collision game when $k = 1$. We construct an algorithm that finds two 1-entries in the same row in H with probability at least $\approx \epsilon^{3/2}$ and performs ≈ 2 queries to H where ϵ is the fraction of 1-entries in H. Formally, we prove the following.

Lemma 3. *Let $H \in \{0,1\}^{R \times N}$ be a binary matrix and let ϵ be the fraction of 1-entries in H. Let $\mathcal{Q}_A$ be a random variable indicating the number of queries performed by A to H. Then, there exists an algorithm A with oracle access to H such that on input $k = 1$ the following holds:*

1. $\mathbb{E}[\mathcal{Q}_A] \leq 2$.
2. $\mathbb{E}\left[\mathcal{Q}_A^2\right] \leq 4$.
3. *Either $\epsilon < \frac{4}{N}$ or $\Pr[\mathsf{CollGame}(A, H) = 1] \geq \frac{\epsilon^{1.5}}{8}$.*

Proof. Let $B = \left\lceil \frac{1}{\sqrt{\epsilon}} - 1 \right\rceil$ and consider the following algorithm A:

$A^H(1)$

1. Sample $\rho \leftarrow R$ and $\beta \leftarrow N$. If $H[\rho, \beta] = 0$ abort.
2. Let $S = \emptyset$. For every $i \in [B]$, sample $\beta_i \leftarrow N \setminus S$ and set $S = S \cup \{\beta_i\}$. If for every $i \in [B]$ it holds that $H[\rho, \beta_i] = 0$, abort.
3. Choose uniformly at random an index i for which $H[\rho, \beta_i] = 1$.
4. Return ρ, β and β_i.

We now prove each claim separately.

Claim 4. *It holds that* $\mathbb{E}[\mathcal{Q}_A] \leq 2$.

Proof. By the description of A, it performs a single query to H, and then only with probability ϵ it performs B queries. Thus, we can bound the expectation by

$$\mathbb{E}[\mathcal{Q}_A] = 1 + \epsilon \cdot B \leq 1 + \frac{1}{\sqrt{\epsilon}} \cdot \epsilon \leq 2 \ .$$

Claim 5. *It holds that* $\mathbb{E}\left[\mathcal{Q}_A^2\right] \leq 4$.

Proof. By the description of A, with probability $1-\epsilon$, it performs a single query, and with probability ϵ it performs $(1 + B)$ queries. Thus, we can bound the expectation squared by

$$\begin{aligned} \mathbb{E}\left[\mathcal{Q}_A^2\right] &= (1-\epsilon) \cdot 1^2 + \epsilon \cdot (1+B)^2 = 1 - \epsilon + \epsilon(1 + 2B + B^2) \\ &= 1 + 2\epsilon B + \epsilon B^2 \leq 1 + 2\sqrt{\epsilon} + 1 \leq 4 \ . \end{aligned}$$

Claim 6 (Success probability). *Either* $\epsilon < \frac{4}{N}$ *or* $\Pr[\mathsf{CollGame}(A, H) = 1] \geq \frac{\epsilon^{1.5}}{8}$.

In order to bound A's success probability, we first show a lower bound on the probability that A does not abort in Item 2.

Claim 7. *Let* X_d *be a random variable indicating the number of* 1*-entries found in* B *draws in a row with exactly* d 1*-entries. For every* $d > 1$*, it holds that* $\Pr[X_d \geq 1] \geq \min\{0.5, \frac{d \cdot B}{2N}\}$.

The proof of Claim 7 appears in the full version.

Proof (Proof of Claim 6). Assuming the first query to the matrix was 1-entry, A continues to sample entries from the same row. Note that for each row, the number of 1-entries may be different which affects the success probability of the algorithm. Therefore, we "divide" the rows into "buckets" by the number of 1-entries in it. Formally, for every $0 \leq d \leq N$, we define δ_d be the fraction of rows with exactly d 1-entries.

When $d \leq 1$, we know that the success probability is 0. Thus, we consider only δ_d for $d \geq 2$. This lets us derive the following:

$$\Pr[\mathsf{CollGame}(A,H)=1] \geq \sum_{d=2}^{N} \delta_d \frac{d}{N} \cdot \Pr\left[X_d \geq 1\right] \geq \sum_{d=2}^{N} \delta_d \frac{d}{N} \cdot \left(\min\left\{\frac{1}{2}, \frac{(d-1)\cdot B}{2(N-1)}\right\}\right)$$

Let $n := \left\lfloor 1 + \frac{N-1}{B} \right\rfloor$, then,

$$\begin{aligned}
\Pr[\mathsf{CollGame}(A,H)=1] &\geq \sum_{d=2}^{n} \delta_d \frac{d}{N} \cdot \frac{(d-1)\cdot B}{2(N-1)} + \sum_{d=n+1}^{N} \delta_d \frac{d}{N} \cdot \frac{1}{2} \\
&= \frac{B}{2} \sum_{d=2}^{n} \delta_d \frac{d(d-1)}{N(N-1)} + \frac{1}{2} \cdot \sum_{d=n+1}^{N} \delta_d \frac{d}{N} \\
&= \frac{B}{2N(N-1)} \sum_{d=0}^{n} \delta_d \cdot d(d-1) + \frac{1}{2} \cdot \sum_{d=n+1}^{N} \delta_d \frac{d}{N}
\end{aligned}$$

Let $\epsilon_1 := \sum_{d=0}^{n} \delta_d \frac{d}{N}$ and $\epsilon_2 := \sum_{d=n+1}^{N} \delta_d \frac{d}{N}$. By Jensen's inequality we get that

$$\frac{1}{N(N-1)} \sum_{d=0}^{n} \delta_d \cdot d(d-1) \geq \frac{1}{N(N-1)} \cdot \epsilon_1 N \left(\epsilon_1 N - 1\right) \geq \frac{\epsilon_1^2 \cdot N - \epsilon_1}{N} = \epsilon_1^2 - \frac{\epsilon_1}{N} \ .$$

Therefore we get, $\Pr[\mathsf{CollGame}(A,H)=1] \geq \frac{B}{2}\left(\epsilon_1^2 - \frac{\epsilon_1}{N}\right) + \frac{1}{2}\epsilon_2$. Since $\epsilon_1 + \epsilon_1 = \epsilon$, the minimum of the above expression is where $\epsilon_1 = \epsilon$. Thus, we can write

$$\Pr[\mathsf{CollGame}(A,H)=1] \geq \frac{B}{2}\left(\epsilon^2 - \frac{\epsilon}{N}\right) \geq \frac{1}{2 \cdot 2\sqrt{\epsilon}} \cdot \epsilon^2 - \frac{\epsilon}{2 \cdot \sqrt{\epsilon} N} = \frac{\epsilon^{1.5}}{4} - \frac{\sqrt{\epsilon}}{2N} \ .$$

Since $\epsilon \geq \frac{4}{N}$, it holds that,

$$\frac{\sqrt{\epsilon}}{2N} \leq \frac{\sqrt{\epsilon}}{2\left(\frac{4}{\epsilon}\right)} = \frac{\epsilon^{1.5}}{8} \ .$$

This leads to,

$$\Pr[\mathsf{CollGame}(A,H)=1] \geq \frac{\epsilon^{1.5}}{8} \ ,$$

which completes the proof.

5.3 Rogue Soundness Error Bound from the Collision Game

We now use the algorithm for the collision game in order to construct an algorithm that extracts a witness $\mathbb{w}$ for an instance $\mathbb{x}$. Then, combined with the second-moment assumption we prove Theorem 1.

First, we prove the following lemma (which is interesting on its own):

Lemma 8. *Let* $t_{\tilde{\boldsymbol{P}}} = t_{\tilde{\boldsymbol{P}}}(\lambda, \boldsymbol{K}), t_{\boldsymbol{V}} = t_{\boldsymbol{V}}(\lambda, \boldsymbol{K}), t_W = t_W(\lambda, \boldsymbol{K})$ *be functions of the security parameter* $\lambda \in \mathbb{N}$ *and the bound on the number of instances* $\boldsymbol{K} \in \mathbb{N}$*. Let* Π *be an algebraic batch* Σ*-protocol for a relation* $\mathcal{R}$ *and let* $\Pi^* = (\mathsf{Setup}, \boldsymbol{P}_1, \boldsymbol{P}_2, \boldsymbol{V}, \mathcal{C})$ *be the batch* Σ*-protocol constructed from* Π *as defined in Lemma 1. Let* $t_{\boldsymbol{V}}$ *denote the running time of the verifier* $\boldsymbol{V}$ *and let* t_W *denote the running time of the witness extractor. Let* $\mathcal{D} = \{\mathcal{D}_\lambda\}_{\lambda \in \mathbb{N}}$ *be a distribution over the relation where each* $\mathcal{D}_\lambda$ *produces* $(\mathsf{pp}, \mathbb{x}, \mathbb{w}) \in \mathcal{R}_\lambda$*. For every prover* $\tilde{\boldsymbol{P}} = (\tilde{\boldsymbol{P}}_1, \tilde{\boldsymbol{P}}_2)$ *that runs in time* $t_{\tilde{\boldsymbol{P}}}$*, there exists an algorithm* A^* *such that:*

1. $\mathbb{E}\left[T_{A^*, \mathcal{D}_\lambda}\right] \leq 2 \cdot (t_{\tilde{\boldsymbol{P}}} + t_{\boldsymbol{V}} + t_W)$.
2. $\mathbb{E}\left[T^2_{A^*, \mathcal{D}_\lambda}\right] \leq 4 \cdot (t_{\tilde{\boldsymbol{P}}} + t_{\boldsymbol{V}} + t_W)^2$.
3. *Either* $\epsilon < \frac{4}{|\mathcal{C}_{\mathsf{pp}}|}$ *or* $\Pr\left[(\mathsf{pp}, \mathbb{x}_1, \tilde{\mathbb{w}}_1) \in \mathcal{R} \;\middle|\; \begin{array}{l} \mathsf{pp} \leftarrow \mathsf{Setup}(1^\lambda, \boldsymbol{K}) \\ (\mathbb{x}_1, \mathbb{w}_1) \leftarrow \mathcal{D}_\lambda(\mathsf{pp}) \\ \tilde{\mathbb{w}}_1 \leftarrow A^*(\mathsf{pp}, \mathbb{x}_1) \end{array}\right] \geq \frac{\epsilon^{1.5}}{8}$ *where* ϵ *is the rogue-soundness error of* Π^* *with respect to a distribution* $\mathcal{D}$ *and the setup algorithm* Setup.

Proof. We denote by aux the variable for tuples of $(\mathsf{pp}, \mathbb{x}, \boldsymbol{\beta})$ where $\boldsymbol{\beta} = (\beta_2, \ldots, \beta_k)$ and $\beta_i \in \{0,1\}^{\mathsf{r}}$. We consider binary matrices $H = \{H_{\mathsf{aux}}\}_{\mathsf{pp}, \mathbb{x}, \boldsymbol{\beta}} \in \{0,1\}^{R \times N}$, where the R rows correspond to $\tilde{\mathbf{P}}$'s randomness and the N columns correspond to $\mathbf{V}$'s randomness for one instance. Note that although $\tilde{\mathbf{P}}$'s and $\mathbf{V}$'s randomness depends on the number of instances that the prover outputs, we can always bound it by the randomness size when $\tilde{\mathbf{P}}$ outputs $\mathbf{K}$ instances.

An entry of H_{aux} equals 1 if and only if the corresponding transcript (between $\tilde{\mathbf{P}}$ and $\mathbf{V}$) is accepting. Recall that every algorithm A for the collision game aims to find $k+1$ entries with the value 1 in the same row. As $\tilde{\mathbf{P}}$'s randomness is fixed along one row, finding two 1-entries in the same row correspond to finding two accepting transcripts $(\alpha, \beta_1, \boldsymbol{\beta}, \gamma), (\alpha, \beta_1', \boldsymbol{\beta}, \gamma')$. Given Claim 2, Π^* has local special soundness, i.e., there exists an algorithm E that runs in time t_W which given two accepting transcripts as considered above, extracts a witness for the instance $\mathbb{x}_1$.

Let A be the algorithm for the collision game constructed in Lemma 3, we construct the algorithm A^* as follows:

$A^*(\mathsf{pp}, \mathbb{x}_1)$

1. Initialize an empty mapping M between the randomness used by $\tilde{\mathbf{P}}$ and $\mathbf{V}$ and the transcript between them.
2. Let r be $\mathbf{V}$'s randomness size for each instance. For $2 \leq i \leq \mathbf{K}$, sample $\beta_i \leftarrow \{0,1\}^{\mathsf{r}}$.
3. Invoke $A(1)$. When A performs a query on (ρ, β) answer as follows:
   (a) Invoke $((\tilde{\mathbb{x}}_2, \ldots, \tilde{\mathbb{x}}_k), \alpha, \mathsf{st}) \leftarrow \tilde{\mathbf{P}}_1(\mathsf{pp}, \mathbb{x}_1; \rho)$.
   (b) Invoke $\gamma \leftarrow \tilde{\mathbf{P}}_2(\beta, \beta_2, \ldots, \beta_k, \mathsf{st})$.
   (c) Set $M[(\rho, \beta)] \leftarrow (\tilde{\mathbb{x}}_2, \ldots, \tilde{\mathbb{x}}_k, \alpha, \beta, \beta_2, \ldots, \beta_k, \gamma)$.
   (d) Return $\mathbf{V}(\mathsf{pp}, \mathbb{x}_1, \tilde{\mathbb{x}}_2, \ldots, \tilde{\mathbb{x}}_k, \alpha, \beta, \beta_2, \ldots, \beta_k, \gamma)$ as the answer to the query.
4. When A outputs ρ, β_1, β_2: set $(\tilde{\mathbb{x}}_2, \ldots, \tilde{\mathbb{x}}_k, \alpha_1^*, \beta_1^*, \beta_{1,2}^* \ldots, \beta_{1,k}^*, \gamma_1^*) \leftarrow M[\rho, \beta_1]$ and $(\tilde{\mathbb{x}}_2, \ldots, \tilde{\mathbb{x}}_k, \alpha_2^*, \beta_2^*, \beta_{2,2}^* \ldots, \beta_{2,k}^*, \gamma_2^*) \leftarrow M[\rho, \beta_2]$.
5. Run $\tilde{\mathbb{w}}_1 \leftarrow E(\tilde{\mathbb{x}}_2, \ldots, \tilde{\mathbb{x}}_k, \alpha_1^*, \beta_{1,2}^*, \ldots, \beta_{1,k}^*, (\beta_{1,1}^*, \gamma_{1,1}^*), (\beta_{2,1}^*, \gamma_{2,1}^*))$.
6. Output $\tilde{w}_1$.

We prove each claim separately.

Claim 9 (Expected running time). *It holds that* $\mathbb{E}\left[T_{A^*,\mathcal{D}_\lambda}\right] \leq 2 \cdot (t_{\tilde{\boldsymbol{P}}} + \boldsymbol{V} + t_W)$.

Proof. Observe that whenever A query H, the algorithm A^* invokes $\tilde{\mathbf{P}}$ and $\mathbf{V}$. Thus, the expected number of invocations that A^* performs to $\tilde{\mathbf{P}}$ and $\mathbf{V}$ is the expected number of queries performed by A. Thus, $\mathbb{E}\left[T_{A^*,\mathcal{D}_\lambda}\right] \leq \mathbb{E}\left[\mathcal{Q}_A\right] \cdot (t_{\tilde{\mathbf{P}}} + \mathbf{V}) + t_W \leq 2 \cdot (t_{\tilde{\mathbf{P}}} + t_{\mathbf{V}} + t_W)$.

Claim 10 (Second-moment of expected running time). *It holds that* $\mathbb{E}\left[T_{A^*,\mathcal{D}_\lambda}^2\right] \leq 4 \cdot (t_{\tilde{\boldsymbol{P}}} + t_{\boldsymbol{V}} + t_W)^2$.

Proof. Following the same observation as in Claim 9 we obtain that

$$\mathbb{E}\left[T_{A^*,\mathcal{D}_\lambda}^2\right] \leq \left(\mathbb{E}\left[\mathcal{Q}_A\right] \cdot (t_{\tilde{\mathbf{P}}} + t_{\mathbf{V}})\right)^2 + t_W^2 \leq \left(\mathbb{E}\left[\mathcal{Q}_A\right] \cdot (t_{\tilde{\mathbf{P}}} + t_{\mathbf{V}} + t_W)\right)^2 = \mathbb{E}\left[\mathcal{Q}_A\right]^2 \cdot (t_{\tilde{\mathbf{P}}} + t_{\mathbf{V}} + t_W)^2 \ .$$

Jensen's inequality leads to $\mathbb{E}\left[T_{A^*,\mathcal{D}_\lambda}^2\right] \leq \mathbb{E}\left[\mathcal{Q}_A^2\right] \cdot (t_{\tilde{\mathbf{P}}} + t_{\mathbf{V}} + t_W)^2 \leq 4(t_{\tilde{\mathbf{P}}} + t_{\mathbf{V}} + t_W)^2$.

Claim 11 (Success probability). *Either* $\epsilon < \frac{4}{|\mathcal{C}_{\mathsf{pp}}|}$ *or* $\Pr\left[(\mathsf{pp}, \mathbb{x}_1, \tilde{\mathbb{w}}_1) \in \mathcal{R} \;\middle|\; \begin{array}{l} \mathsf{pp} \leftarrow \mathsf{Setup}(1^\lambda, \boldsymbol{K}) \\ (\mathbb{x}_1, \mathbb{w}_1) \leftarrow \mathcal{D}_\lambda(\mathsf{pp}) \\ \tilde{\mathbb{w}}_1 \leftarrow A^*(\mathsf{pp}, \mathbb{x}_1) \end{array}\right] \geq \frac{\epsilon^{1.5}}{8}$ *where* ϵ *is the rogue-soundness error of* Π^* *with respect to a distribution* $\mathcal{D}$ *and the setup algorithm* Setup.

Proof. Whenever A succeeds in the collision game with H_{aux}, the algorithm A^* outputs a witness for $\mathbb{x}_1$. Thus,

$$\Pr\left[(\mathsf{pp},\mathbb{x}_1,\tilde{\mathbb{w}}_1)\in\mathcal{R}\;\middle|\;\begin{array}{l}\mathsf{pp}\leftarrow\mathsf{Setup}(1^\lambda,\mathbf{K})\\(\mathbb{x}_1,\mathbb{w}_1)\leftarrow\mathcal{D}_\lambda(\mathsf{pp})\\\tilde{\mathbb{w}}_1\leftarrow A^*(\mathsf{pp},\mathbb{x}_1)\end{array}\right]=\sum_{\mathsf{aux}}\Pr[\mathsf{aux}]\cdot\Pr\left[\mathsf{CollGame}(A,H_{\mathsf{aux}})=1\right]\ .$$

For every $\mathsf{aux}=(\mathsf{pp},\mathbb{x},\boldsymbol{\beta})$, we let

$$\epsilon_{\mathsf{aux}}=\Pr\left[\begin{array}{l}\mathbf{V}(\mathsf{pp},\mathbb{x},\tilde{\mathbb{x}}_2,\ldots,\tilde{\mathbb{x}}_k,\alpha,\beta,\beta_2,\ldots,\beta_k,\gamma)=1\\\text{conditioned on }\mathsf{pp}\leftarrow\mathsf{Setup}(1^\lambda,\mathbf{K})\\\wedge\ (\mathbb{x}_1,\mathbb{w}_1)\leftarrow\mathcal{D}_\lambda(\mathsf{pp})\\\wedge\ \beta_2,\ldots,\beta_k\leftarrow\mathcal{C}_{\mathsf{pp}}\end{array}\;\middle|\;\begin{array}{l}((\tilde{\mathbb{x}}_2,\ldots,\tilde{\mathbb{x}}_k),\alpha,\mathsf{st})\leftarrow\tilde{\mathbf{P}}_1(1^\lambda,\mathsf{pp},\mathbb{x})\\\beta_2,\ldots,\beta_k\leftarrow\mathcal{C}_{\mathsf{pp}}\\\gamma\leftarrow\tilde{\mathbf{P}}_2(\mathsf{st},\beta_2,\ldots,\beta_k)\end{array}\right]\ .$$

The collision game matrix H_{aux} has ϵ_{aux} fraction of 1-entries. Thus, conditioned on aux, the probability that A succeeds in the collision game is $\frac{\epsilon_{\mathsf{aux}}^{1.5}}{8}$. Therefore,

$$\begin{aligned}\Pr\left[(\mathsf{pp},\mathbb{x}_1,\tilde{\mathbb{w}}_1)\in\mathcal{R}\;\middle|\;\begin{array}{l}\mathsf{pp}\leftarrow\mathsf{Setup}(1^\lambda,\mathbf{K})\\(\mathbb{x}_1,\mathbb{w}_1)\leftarrow\mathcal{D}_\lambda(\mathsf{pp})\\\tilde{\mathbb{w}}_1\leftarrow A^*(\mathsf{pp},\mathbb{x}_1)\end{array}\right]&=\sum_{\mathsf{aux}}\Pr[\mathsf{aux}]\cdot\frac{\epsilon_{\mathsf{aux}}^{1.5}}{8}=\mathop{\mathbb{E}}_{\mathsf{aux}}\left[\frac{\epsilon_{\mathsf{aux}}^{1.5}}{8}\right]\\&\geq\frac{\mathop{\mathbb{E}}_{\mathsf{aux}}\left[\epsilon_{\mathsf{aux}}\right]^{1.5}}{8}\geq\frac{\epsilon^{1.5}}{8}\ ,\end{aligned}$$

where first inequality follows from Jensen's inequality and the last inequality follows from the fact that $\mathop{\mathbb{E}}_{\mathsf{aux}}[\epsilon_{\mathsf{aux}}]=\epsilon$.

We are now ready to show a bound on the rogue soundness error of batch Σ-protocol defined in Lemma 1.

Proof (Proof of Theorem 1). Let $\tilde{\mathbf{P}}$ be a cheating prover and let $\epsilon_{\mathcal{D}}$ be the rogue soundness error with respect to $\mathcal{D}$ and Setup. Given Lemma 8 and the assumption that $\mathcal{R}$ is second-moment hard with respect to the distribution $\mathcal{D}$ and the setup algorithm Setup, it holds that either $\epsilon_{\mathcal{D}}<\frac{4}{|\mathcal{C}_{\mathsf{pp}}|}$ or,

$$\frac{\epsilon_{\mathcal{D}}^{1.5}}{8}\leq\Pr\left[(\mathsf{pp},\mathbb{x}_1,\tilde{\mathbb{w}}_1)\in\mathcal{R}\;\middle|\;\begin{array}{l}\mathsf{pp}\leftarrow\mathsf{Setup}(1^\lambda,\mathbf{K})\\(\mathbb{x}_1,\mathbb{w}_1)\leftarrow\mathcal{D}_\lambda(\mathsf{pp})\\\tilde{\mathbb{w}}_1\leftarrow A^*(\mathsf{pp},\mathbb{x}_1)\end{array}\right]\leq\frac{\Delta\cdot\mathbb{E}\left[T_{A^*,\mathcal{D}}^2\right]}{|\mathcal{W}_\lambda|^\omega}\leq\frac{\Delta\cdot4\cdot(t_{\tilde{\mathbf{P}}}+t_{\mathbf{V}}+t_W)^2}{|\mathcal{W}_\lambda|^\omega}\ .$$

This leads to

$$\epsilon_{\mathcal{D}}\leq\left(\frac{\Delta\cdot32\cdot(t_{\tilde{\mathbf{P}}}+t_{\mathbf{V}}+t_W)}{|\mathcal{W}_\lambda|^\omega}\right)^{2/3}\ .$$

Overall we derive the following bound

$$\epsilon_{\mathcal{D}}\leq\max\left\{\left(\frac{\Delta\cdot32\cdot(t_{\tilde{\mathbf{P}}}+t_{\mathbf{V}}+t_W)}{|\mathcal{W}_\lambda|^\omega}\right)^{2/3},\frac{4}{|\mathcal{C}_{\mathsf{pp}}|}\right\}\leq\left(\frac{\Delta\cdot32\cdot(t_{\tilde{\mathbf{P}}}+t_{\mathbf{V}}+t_W)}{|\mathcal{W}_\lambda|^\omega}\right)^{2/3}+\frac{4}{|\mathcal{C}_{\mathsf{pp}}|}$$

5.4 Algebraic Batch Identification Schemes

An identification scheme consists of a Σ-protocol for relation $\mathcal{R}$ and an algorithm Gen that produces a distribution over $(\mathbb{x},\mathbb{w})\in\mathcal{R}$ where the public key is the instance $\mathbb{x}$ and the secret key is the witness $\mathbb{w}$. Similarly, we construct a batch

identification scheme that consists of batch Σ-protocol defined in Lemma 1 and an algorithm Gen that given public parameters pp, produces a distribution over $(\mathbb{x}, \mathbb{w}) \in \mathcal{R}(\mathsf{pp})$.

Note that the execution of ID is as the execution of the batch Σ-protocol where each public key pk corresponds to an instance, and a secret key sk corresponds to a witness.

We consider the rogue-security notion of batch identification scheme, asking a cheating prover $\tilde{\mathbf{P}}$ given as input an instance $\mathbb{x}$ produced by Gen, to convince the verifier $\mathbf{V}$ on $(\mathbb{x}, \tilde{\mathbb{x}}_2, \ldots, \tilde{\mathbb{x}}_k)$ where $\tilde{\mathbb{x}}_2, \ldots, \tilde{\mathbb{x}}_k$ are adaptively chosen by $\tilde{\mathbf{P}}$ while given access to an honest transcript-generator for the instance $\mathbb{x}$ and another $(k-1)$ instances by its choice. Formally, we let $\mathsf{Trans}_{\mathsf{pk}_1,\mathsf{sk}_1}(\cdot)$ denote an oracle that when queried with input $(\mathsf{pk}_2, \mathsf{sk}_2), \ldots (\mathsf{pk}_k, \mathsf{sk}_k)$, runs an honest execution of the protocol on input $(\mathsf{pk}_1, \mathsf{sk}_1), \ldots (\mathsf{pk}_k, \mathsf{sk}_k)$ and returns the resulting transcripts (α, β, γ). We define the rogue-security of a batch identification scheme as follows:

Definition 13 (Rogue soundness). *Let* $\mathsf{ID} = (\mathsf{Setup}, \mathsf{Gen}, \boldsymbol{P}_1, \boldsymbol{P}_2, \boldsymbol{V}, \mathcal{C})$ *be a batch identification scheme for a relation* $\mathcal{R}$*. Then,* ID *is* (t, ϵ)*-rogue soundness (with respect to* Gen *and* Setup*) if for every* $\lambda, k \in \mathbb{N}$ *such that* $k \leq \boldsymbol{K}$ *and for any* t*-time malicious prover* $\tilde{\boldsymbol{P}} = (\tilde{\boldsymbol{P}}_1, \tilde{\boldsymbol{P}}_2)$ *that performs* q *queries to the transcript-generation oracle it holds that:*

$$\Pr\left[\mathsf{StrongIdent}_{\mathsf{ID}}(\tilde{\boldsymbol{P}}, \lambda)\right] \leq \epsilon(\lambda, t, \mathsf{q}, \boldsymbol{K}) \ ,$$

where the experiment $\mathsf{StrongIdent}_{\mathsf{ID}}(\tilde{\boldsymbol{P}}, \lambda)$ *defined as follows:*

$\mathsf{StrongIdent}_{\mathsf{ID}}(\tilde{\boldsymbol{P}}, \lambda)$*:*

1. $\mathsf{pp} \leftarrow \mathsf{Setup}(1^\lambda, \boldsymbol{K})$.
2. $(\mathsf{pk}_1, \mathsf{sk}_1) \leftarrow \mathsf{Gen}(\mathsf{pp})$.
3. $((\tilde{\mathsf{pk}}_2, \ldots, \tilde{\mathsf{pk}}_k), \alpha, \mathsf{st}) \leftarrow \tilde{\boldsymbol{P}}_1^{\mathsf{Trans}_{\mathsf{pk}_1,\mathsf{sk}_1}(\cdot)}(\mathsf{pp}, \mathsf{pk}_1)$.
4. $\beta \leftarrow \mathcal{C}_{\mathsf{pp}}$.
5. $\gamma \leftarrow \tilde{\boldsymbol{P}}_2(\mathsf{st}, \beta)$.
6. *Output* $\boldsymbol{V}(\mathsf{pp}, \mathsf{pk}_1, \tilde{\mathsf{pk}}_2, \ldots, \tilde{\mathsf{pk}}_k, \alpha, \beta, \gamma) = 1$.

Recall that batch identification scheme ID consists of a batch Σ-protocol Π^* defined in Lemma 1 such that the execution of ID is as the execution of Π^* where each public key pk corresponds to an instance and a secret key sk corresponds to a witness. Thus, if Π^* is zero-knowledge, we can assume that every malicious prover does not query the transcript-generation oracle, as such queries can be internally simulated given the public keys. Formally, if Π^* is t-time zero-knowledge (Definition 8), for every malicious prover that performs q queries to the transcript-generation oracle $\mathsf{Trans}_{\mathsf{pk}_1,\mathsf{sk}_1}(\cdot)$, we can construct a malicious prover that does not query the transcript-generation oracle and instead runs the simulator q times to generate transcripts. Specifically, if Π^* has t_{Sim}-time zero-knowledge, any malicious prover that runs in time $t_{\tilde{\mathbf{P}}}$ and performs

q queries to $\mathsf{Trans}_{\mathsf{pk}_1,\mathsf{sk}_1}(\cdot)$, can be simulated by a malicious prover that runs in time $t_{\tilde{\boldsymbol{P}}} + \mathsf{q} \cdot t_{\mathsf{Sim}}$.

Recall that every batch Σ-protocol Π^* defined in Lemma 1 is constructed from an algebraic Σ-protocol Π. We now show that if Π is t_{Sim}-time zero-knowledge, then Π^* is $(k \cdot t_{\mathsf{Sim}})$-zero-knowledge. Formally, we prove the following.

Claim 12. *Let $\Pi = (\mathsf{Setup}, \boldsymbol{P}_1, \boldsymbol{P}_2, \boldsymbol{V}, \mathcal{C})$ be an algebraic Σ-protocol for a relation $\mathcal{R}$ and let Π^* be the batch Σ-protocol constructed from Π as defined in Lemma 1 with a bound $\boldsymbol{K}$ on the number of instances. If Π is t_{Sim}-time zero-knowledge, then Π^* is $(\boldsymbol{K} \cdot t_{\mathsf{Sim}})$-time zero-knowledge.*

The proof of Claim 12 appears in the full version. Combined with Theorem 1, we derive the following corollary:

Corollary 2. *Let $\Delta = \Delta(\lambda), \omega = \omega(\lambda), t_{\tilde{\boldsymbol{P}}} = t_{\tilde{\boldsymbol{P}}}(\lambda), t_{\boldsymbol{V}} = t_{\boldsymbol{V}}(\lambda, \boldsymbol{K}), t_W = t_W(\lambda, \boldsymbol{K}), t_{\mathsf{Sim}} = t_{\mathsf{Sim}}(\lambda, \boldsymbol{K}), \mathsf{q} = \mathsf{q}(\lambda)$ be functions of the security parameter $\lambda \in \mathbb{N}$ and the bound on the number of instances $\boldsymbol{K} \in \mathbb{N}$. Let Π be an algebraic Σ-protocol for relation $\mathcal{R}$ with t_{Sim}-time zero-knowledge and let $\Pi^* = (\mathsf{Setup}, \boldsymbol{P}_1, \boldsymbol{P}_2, \boldsymbol{V}, \mathcal{C})$ be the batch Σ-protocol constructed from Π as defined in Lemma 1. Let $\mathsf{ID} = (\mathsf{Setup}, \mathsf{Gen}, \boldsymbol{P}_1, \boldsymbol{P}_2, \boldsymbol{V}, \mathcal{C})$ be the batch identification scheme consists with Π^*. If $\mathcal{R}$ is second-moment hard with respect to Gen, then for any malicious prover $\tilde{\boldsymbol{P}}$ that runs in time $t_{\tilde{\boldsymbol{P}}}$ and issues q transcript-generation queries it holds that*

$$\Pr\left[\mathsf{StrongIdent}_{\mathsf{ID}}(\tilde{\boldsymbol{P}}, \lambda)\right] \leq \left(\frac{\Delta \cdot 32 \cdot (t_{\tilde{\boldsymbol{P}}} + \mathsf{q} \cdot \boldsymbol{K} \cdot t_{\mathsf{Sim}} + t_{\boldsymbol{V}} + t_W)^2}{|\mathcal{W}_\lambda|^\omega}\right)^{2/3} + \frac{4}{|\mathcal{C}_{\mathsf{pp}}|} \ ,$$

where $t_{\boldsymbol{V}}$ is the running time of the verifier $\boldsymbol{V}$ and t_W is the running time of the witness extractor.

6 Proving Expected-Time Hardness in Generic Models

In this section, we present a generic framework for analyzing expected-time hardness of cryptographic problems. In fact, applying our framework proves the second-moment assumption (Definition 2) for the discrete-logarithm problem in the generic group model. Shoup [33] analyzed generic hardness of the discrete-logarithm problem with respect to strict time algorithms. He showed that any generic t-time algorithm that solves the discrete-logarithm problem has success probability at most $\epsilon \leq t^2/p$. Applying our framework yields a bound of $\epsilon \leq \mathbb{E}\left[T_A^2\right]/p$ when considering *unbounded* algorithms where T_A denotes the random variable indicating the algorithm's running time.

Our framework is inspired by [19] which showed a generic framework to prove bounds with respect to expected-time algorithms when considering only the first-moment of the expected running time. Their result proves the first-moment assumption (Definition 1) but cannot be used to derive the second-moment assumption.

In Sect. 6.1 we introduce our framework for proving expected-time hardness.

6.1 Our Framework

Definition 14 (Monotonic predicate). *A predicate P is* monotonic *if for every tr such that $P(\mathsf{tr}) = 1$, it holds that $P(\mathsf{tr}||\mathsf{tr}') = 1$ for every tr'.*

We consider distributions $\mathcal{D}(\lambda)$ which produces an oracle $\mathcal{O}$ and define the hardness of a predicate as follows:

Definition 15 (Hard predicate). *A predicate P is* ϵ-hard *if for every strict time algorithm $\mathcal{A}_t$ it holds that*

$$\Pr\left[P(\mathsf{tr}) = 1 \;\middle|\; \begin{array}{l} \mathcal{O} \leftarrow \mathcal{D}(\lambda) \\ \mathsf{out} \overset{\mathsf{tr}}{\leftarrow} {\mathcal{A}_t}^{\mathcal{O}}(\mathsf{in}) \end{array}\right] \leq \epsilon(t) \ .$$

In addition, we define history-oblivious predicates. Intuitively, this family of predicates includes predicates on which each query is oblivious to the history of the query-answer list (see Sect. 2.6 for further discussion). We define history-oblivious by considering the hardness to set the predicate to output 1 on input $\mathsf{tr}\|(x, y)$ where (x, y) is a fresh query-answer pair and tr is a query-answer list on which the predicate outputs 0.

For any list of query-answer pairs μ we denote by $\mathcal{D}(\lambda, \mu)$ the distribution $\mathcal{D}(\lambda)$ of all oracles such that for every $(x_i, y_i) \in \mu$ it holds that $y_i = \mathcal{O}(x_i)$. We let X, Y be the query and answer spaces.

Definition 16 (History-oblivious predicate). *Let P be an ϵ-hard predicate. We say that P is* history-oblivious *with respect to $\mathcal{O}$ if there is a function $\kappa(\cdot)$, such that for every $t \in \mathbb{N}$ the following holds:*

1. *For every $0 \leq i \leq t$, every trace tr of length i with $P(\mathsf{tr}) = 0$, and any query $x \in X$:*

$$\Pr\left[P(\mathsf{tr}\|(x, y)) = 1 \;\middle|\; \begin{array}{l} \mathcal{O} \leftarrow \mathcal{D}(\lambda, \mathsf{tr}) \\ y = \mathcal{O}(x) \end{array}\right] \leq \kappa(i) \ .$$

2. $\sum_{j=0}^{t} \kappa(j) \leq \epsilon(t)$.

(Above, the length of a trace is the number of query/answer pairs it contains.) We consider experiments relative to an oracle, for which their security relies on the trace between the adversary and the oracle. We capture this using a monotonic predicate on the trace. Formally, we define the following:

Definition 17 (δ-bounded experiment). *Let $\mathsf{Exp}^{\mathcal{O}}$ be an experiment with oracle access $\mathcal{O}$, and let $\delta = \delta(\lambda)$ be a function of the security parameter $\lambda \in \mathbb{N}$. We say that $\mathsf{Exp}^{\mathcal{O}}$ is* δ-bounded *with respect to a monotonic predicate P if for every (bounded and unbounded) algorithm $\mathcal{A}$ it holds that,*

$$\Pr\left[\mathsf{Exp}^{\mathcal{O}}(\mathsf{in}, \mathsf{out}) = 1 \;\middle|\; \begin{array}{l} \mathcal{O} \leftarrow \mathcal{D}(\lambda) \\ \mathsf{out} \overset{\mathsf{tr}}{\leftarrow} \mathcal{A}^{\mathcal{O}}(\mathsf{in}) \end{array}\right] \leq \Pr\left[P(\mathsf{tr}) = 1 \;\middle|\; \begin{array}{l} \mathcal{O} \leftarrow \mathcal{D}(\lambda) \\ \mathsf{out} \overset{\mathsf{tr}}{\leftarrow} \mathcal{A}^{\mathcal{O}}(\mathsf{in}) \end{array}\right] + \delta \ .$$

Given the definitions above, we prove the following theorem.

Theorem 2. *Let* $\mathsf{Exp}^{\mathcal{O}}$ *be a* δ*-bounded experiment with respect to a predicate* P *which is* ϵ*-hard. If* P *is history-oblivious, then, for every* unbounded *algorithm* $\mathcal{A}$ *it holds that,*

$$\Pr\left[\mathsf{Exp}^{\mathcal{O}}(\mathsf{in},\mathsf{out})=1 \;\middle|\; \begin{array}{l}\mathcal{O}\leftarrow\mathcal{D}(\lambda)\\ \mathsf{out}\overset{\mathsf{tr}}{\leftarrow}\mathcal{A}^{\mathcal{O}}(\mathsf{in})\end{array}\right]\leq\mathbb{E}\left[\epsilon(t)\right]+\delta\ .$$

In particular, Theorem 2 allows us to recover the same bounds given in [19], which is captured in the following corollary.

Corollary 3. *Let* $\mathsf{Exp}^{\mathcal{O}}$ *be a* δ*-bounded experiment with respect to a predicate* P *which is* ϵ*-hard where* $\epsilon(t)=\frac{\Delta t^d}{N}$ *for* $\Delta,d,N\geq 1$*. If* P *is history-oblivious, then, for every* unbounded *algorithm* $\mathcal{A}$ *it holds that,*

$$\Pr\left[\mathsf{Exp}^{\mathcal{O}}(\mathsf{in},\mathsf{out})=1 \;\middle|\; \begin{array}{l}\mathcal{O}\leftarrow\mathcal{D}(\lambda)\\ \mathsf{out}\overset{\mathsf{tr}}{\leftarrow}\mathcal{A}^{\mathcal{O}}(\mathsf{in})\end{array}\right]\leq\sqrt[d]{\epsilon(\mathbb{E}\left[T_{\mathcal{A}}\right])}+\delta=\sqrt[d]{\frac{\Delta}{N}}\cdot\mathbb{E}\left[T_{\mathcal{A}}\right]+\delta\ ,$$

where $T_{\mathcal{A}}$ *is a random variable indicating the number of queries performed by* $\mathcal{A}$ *until he stops, when given access to an oracle* $\mathcal{O}$*.*

The proof of Corollary 3 appears in the full version, we now prove Theorem 2.

Proof (Proof of Theorem 2). Let tr_i be the first i pairs in the query-answer list between the algorithm and the oracle $\mathcal{O}$. Let Y_i be an indicator random variable for the event that (i) $|\mathsf{tr}|\geq i$; (ii) $P\left(\mathsf{tr}_i\right)=1$; and (iii) $P\left(\mathsf{tr}_{i-1}\right)=0$. Note that, the events $Y_i=1$ are mutually exclusive, thus:

$$\Pr\left[P\left(\mathsf{tr}\right)=1 \;\middle|\; \begin{array}{l}\mathcal{O}\leftarrow\mathcal{D}(\lambda)\\ \mathsf{out}\overset{\mathsf{tr}}{\leftarrow}\mathcal{A}^{\mathcal{O}}(\mathsf{in})\end{array}\right]=\sum_{i=1}^{\infty}\Pr\left[Y_i=1 \;\middle|\; \begin{array}{l}\mathcal{O}\leftarrow\mathcal{D}(\lambda)\\ \mathsf{out}\overset{\mathsf{tr}}{\leftarrow}\mathcal{A}^{\mathcal{O}}(\mathsf{in})\end{array}\right]\ ,$$

To simplify the notation throughout the proof, we omit the explicit reference to the probability taken over the sampling of the oracle $\mathcal{O}\leftarrow\mathcal{D}(\lambda)$ and the execution of the algorithm.

Let $T_{\mathcal{A}}=T_{\mathcal{A}}(\lambda)$ be a random variable indicating the number of queries performed by $\mathcal{A}$ until he stops, when given access to an oracle $\mathcal{O}$. Note that for every $i\in\mathbb{N}$ it holds that $Y_i=1$ only if the number of queries performed by the algorithm is at least i. Thus,

$$\begin{aligned}&\Pr\left[P\left(\mathsf{tr}\right)=1 \;\middle|\; \begin{array}{l}\mathcal{O}\leftarrow\mathcal{D}(\lambda)\\ \mathsf{out}\overset{\mathsf{tr}}{\leftarrow}\mathcal{A}^{\mathcal{O}}(\mathsf{in})\end{array}\right]=\sum_{i=1}^{\infty}\Pr\left[Y_i=1\;\middle|\;T_{\mathcal{A}}\geq i\right]\cdot\Pr[T_{\mathcal{A}}\geq i]\\ &\leq\sum_{i=1}^{\infty}\Pr\left[Y_i=1\;\middle|\;T_{\mathcal{A}}\geq i\right]\cdot\sum_{t=i}^{\infty}\Pr[T_{\mathcal{A}}=t]\end{aligned}$$

The following claim shows an upper bound on the above term $\Pr\left[Y_i = 1 \;\middle|\; T_{\mathcal{A}} \geq i\right]$. The proof of the claim appears in the full version.

Claim 13. *If P is ϵ-hard and history-oblivious, then for every $i \in \mathbb{N}$, it holds that* $\Pr\left[Y_i = 1 \;\middle|\; T_{\mathcal{A}} \geq i\right] \leq \kappa(i)$.

Given Claim 13 it holds that,

$$\Pr\left[P(\mathsf{tr}) = 1 \;\middle|\; \begin{array}{l} \mathcal{O} \leftarrow \mathcal{D}(\lambda) \\ \mathsf{out} \xleftarrow{\mathsf{tr}} \mathcal{A}^{\mathcal{O}}(\mathsf{in}) \end{array}\right] \leq \sum_{i=1}^{\infty} \kappa(i) \cdot \sum_{t=i}^{\infty} \Pr\left[T_{\mathcal{A}} = t\right]$$

$$= \sum_{t=1}^{\infty} \Pr\left[T_{\mathcal{A}} = t\right] \cdot \sum_{i=1}^{t} \kappa(i) \leq \sum_{t=1}^{\infty} \Pr\left[T_{\mathcal{A}} = t\right] \cdot \epsilon(t) = \mathbb{E}\left[\epsilon(t)\right] \quad ,$$

where the first equality follows from rearranging the summation, and the last inequality follows from the fact that P is ϵ-hard and history-oblivious. Overall, we conclude that,

$$\Pr\left[\mathsf{Exp}^{\mathcal{O}}(\mathsf{out}) = 1 \;\middle|\; \begin{array}{l} \mathcal{O} \leftarrow \mathcal{D}(\lambda) \\ \mathsf{out} \xleftarrow{\mathsf{tr}} \mathcal{A}^{\mathcal{O}}(\mathsf{in}) \end{array}\right] \leq \mathbb{E}\left[\epsilon(t)\right] + \delta \quad .$$

References

1. Abdalla, M., An, J.H., Bellare, M., Namprempre, C.: From identification to signatures via the fiat-shamir transform: minimizing assumptions for security and forward-security. In: Knudsen, L.R. (ed.) EUROCRYPT 2002. LNCS, vol. 2332, pp. 418–433. Springer, Heidelberg (2002). https://doi.org/10.1007/3-540-46035-7_28
2. Agrikola, T., Hofheinz, D., Kastner, J.: On instantiating the algebraic group model from falsifiable assumptions. In: Canteaut, A., Ishai, Y. (eds.) EUROCRYPT 2020. LNCS, vol. 12106, pp. 96–126. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-45724-2_4
3. Attema, T., Cramer, R., Kohl, L.: A compressed Σ-protocol theory for lattices. In: Malkin, T., Peikert, C. (eds.) CRYPTO 2021. LNCS, vol. 12826, pp. 549–579. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-84245-1_19
4. Bauer, B., Fuchsbauer, G., Loss, J.: A classification of computational assumptions in the algebraic group model. In: Micciancio, D., Ristenpart, T. (eds.) CRYPTO 2020. LNCS, vol. 12171, pp. 121–151. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-56880-1_5
5. Bellare, M., Dai, W.: The multi-base discrete logarithm problem: tight reductions and non-rewinding proofs for Schnorr identification and signatures. In: Bhargavan, K., Oswald, E., Prabhakaran, M. (eds.) INDOCRYPT 2020. LNCS, vol. 12578, pp. 529–552. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-65277-7_24
6. Bellare, M., Dai, W.: Chain reductions for multi-signatures and the HBMS scheme. In: Tibouchi, M., Wang, H. (eds.) ASIACRYPT 2021. LNCS, vol. 13093, pp. 650–678. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-92068-5_22

7. Bellare, M., Neven, G.: Multi-signatures in the plain public-key model and a general forking lemma. In: Proceedings of the ACM Conference on Computer and Communications Security, pp. 390–399 (2006)
8. Ben-Sasson, E., Chiesa, A., Spooner, N.: Interactive oracle proofs. In: Hirt, M., Smith, A. (eds.) TCC 2016. LNCS, vol. 9986, pp. 31–60. Springer, Heidelberg (2016). https://doi.org/10.1007/978-3-662-53644-5_2
9. Boneh, D., Drijvers, M., Neven, G.: Compact multi-signatures for smaller blockchains. In: Peyrin, T., Galbraith, S. (eds.) ASIACRYPT 2018. LNCS, vol. 11273, pp. 435–464. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-03329-3_15
10. Bootle, J., Cerulli, A., Chaidos, P., Groth, J., Petit, C.: Efficient zero-knowledge arguments for arithmetic circuits in the discrete log setting. In: Fischlin, M., Coron, J.-S. (eds.) EUROCRYPT 2016. LNCS, vol. 9666, pp. 327–357. Springer, Heidelberg (2016). https://doi.org/10.1007/978-3-662-49896-5_12
11. Catalano, D., Fiore, D., Gennaro, R., Vamvourellis, K.: Algebraic (trapdoor) one-way functions: constructions and applications. Theoret. Comput. Sci. **592**, 143–165 (2015)
12. Cramer, R.: Modular design of secure yet practical cryptographic protocols. Ph.D. thesis, CWI and University of Amsterdam (1996)
13. Cramer, R., Damgård, I.: Zero-knowledge proofs for finite field arithmetic, or: can zero-knowledge be for free? In: Krawczyk, H. (ed.) CRYPTO 1998. LNCS, vol. 1462, pp. 424–441. Springer, Heidelberg (1998). https://doi.org/10.1007/BFb0055745
14. Fuchsbauer, G., Kiltz, E., Loss, J.: The algebraic group model and its applications. In: Shacham, H., Boldyreva, A. (eds.) CRYPTO 2018. LNCS, vol. 10992, pp. 33–62. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-96881-0_2
15. Fuchsbauer, G., Plouviez, A., Seurin, Y.: Blind Schnorr signatures and signed ElGamal encryption in the algebraic group model. In: Canteaut, A., Ishai, Y. (eds.) EUROCRYPT 2020. LNCS, vol. 12106, pp. 63–95. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-45724-2_3
16. Gennaro, R., Leigh, D., Sundaram, R., Yerazunis, W.: Batching Schnorr identification scheme with applications to privacy-preserving authorization and low-bandwidth communication devices. In: Lee, P.J. (ed.) ASIACRYPT 2004. LNCS, vol. 3329, pp. 276–292. Springer, Heidelberg (2004). https://doi.org/10.1007/978-3-540-30539-2_20
17. Guillou, L.C., Quisquater, J.-J.: A "paradoxical" indentity-based signature scheme resulting from zero-knowledge. In: Goldwasser, S. (ed.) CRYPTO 1988. LNCS, vol. 403, pp. 216–231. Springer, New York (1990). https://doi.org/10.1007/0-387-34799-2_16
18. Hazay, C., Lindell, Y.: Efficient Secure Two-Party Protocols - Techniques and Constructions. Information Security and Cryptography, Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-14303-8
19. Jaeger, J., Tessaro, S.: Expected-time cryptography: generic techniques and applications to concrete soundness. In: Pass, R., Pietrzak, K. (eds.) TCC 2020. LNCS, vol. 12552, pp. 414–443. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-64381-2_15
20. Katz, J., Lindell, Y.: Handling expected polynomial-time strategies in simulation-based security proofs. J. Cryptol. **21**(3), 303–349 (2008)

21. Kiltz, E., Masny, D., Pan, J.: Optimal security proofs for signatures from identification schemes. In: Robshaw, M., Katz, J. (eds.) CRYPTO 2016. LNCS, vol. 9815, pp. 33–61. Springer, Heidelberg (2016). https://doi.org/10.1007/978-3-662-53008-5_2
22. Maurer, U.: Abstract models of computation in cryptography. In: Smart, N.P. (ed.) Cryptography and Coding 2005. LNCS, vol. 3796, pp. 1–12. Springer, Heidelberg (2005). https://doi.org/10.1007/11586821_1
23. Maxwell, G., Poelstra, A., Seurin, Y., Wuille, P.: Simple Schnorr multi-signatures with applications to Bitcoin. Des. Codes Crypt. **87**(9), 2139–2164 (2019). https://doi.org/10.1007/s10623-019-00608-x
24. Micali, S.: Computationally sound proofs. SIAM J. Comput. **30**(4), 1253–1298 (2000)
25. Mizuide, T., Takayasu, A., Takagi, T.: Tight reductions for Diffie-Hellman variants in the algebraic group model. In: Matsui, M. (ed.) CT-RSA 2019. LNCS, vol. 11405, pp. 169–188. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-12612-4_9
26. Nick, J., Ruffing, T., Seurin, Y.: MuSig2: simple two-round Schnorr multi-signatures. In: Malkin, T., Peikert, C. (eds.) CRYPTO 2021. LNCS, vol. 12825, pp. 189–221. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-84242-0_8
27. Okamoto, T.: Provably secure and practical identification schemes and corresponding signature schemes. In: Brickell, E.F. (ed.) CRYPTO 1992. LNCS, vol. 740, pp. 31–53. Springer, Heidelberg (1993). https://doi.org/10.1007/3-540-48071-4_3
28. Pointcheval, D., Stern, J.: Security arguments for digital signatures and blind signatures. J. Cryptol. **13**, 361–396 (2000)
29. Rotem, L., Segev, G.: Algebraic distinguishers: from discrete logarithms to decisional uber assumptions. In: Pass, R., Pietrzak, K. (eds.) TCC 2020. LNCS, vol. 12552, pp. 366–389. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-64381-2_13
30. Rotem, L., Segev, G.: Tighter security for Schnorr identification and signatures: a high-moment forking lemma for Σ-protocols. In: Malkin, T., Peikert, C. (eds.) CRYPTO 2021. LNCS, vol. 12825, pp. 222–250. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-84242-0_9
31. Schnorr, C.P.: Efficient identification and signatures for smart cards. In: Brassard, G. (ed.) CRYPTO 1989. LNCS, vol. 435, pp. 239–252. Springer, New York (1990). https://doi.org/10.1007/0-387-34805-0_22
32. Schnorr, C.P.: Efficient signature generation by smart cards. J. Cryptol. **4**(3), 161–174 (1991). https://doi.org/10.1007/BF00196725
33. Shoup, V.: Lower bounds for discrete logarithms and related problems. In: Fumy, W. (ed.) EUROCRYPT 1997. LNCS, vol. 1233, pp. 256–266. Springer, Heidelberg (1997). https://doi.org/10.1007/3-540-69053-0_18
34. Yun, A.: Generic hardness of the multiple discrete logarithm problem. In: Oswald, E., Fischlin, M. (eds.) EUROCRYPT 2015. LNCS, vol. 9057, pp. 817–836. Springer, Heidelberg (2015). https://doi.org/10.1007/978-3-662-46803-6_27