Optimal Proximity Proofs

Ioana Boureanu[1] and Serge Vaudenay[2]

[1] Akamai Technologies Limited
EMEA HQ, UK
http://people.itcarlson.com/ioana
[2] EPFL
Lausanne, Switzerland
http://lasec.epfl.ch

Abstract. Provably secure distance-bounding is a rising subject, yet an unsettled one; indeed, very few distance-bounding protocols, with formal security proofs, have been proposed. In fact, so far only two protocols, namely SKI (by Boureanu *et al.*) and FO (by Fischlin and Onete), offer all-encompassing security guaranties, i.e., resistance to distance-fraud, mafia-fraud, and terrorist-fraud. Matters like security, alongside with soundness, or added tolerance to noise do not always coexist in the (new) distance-bounding designs. Moreover, as we will show in this paper, *efficiency* and *simultaneous* protection against all frauds seem to be rather conflicting matters, leading to proposed solutions which were/are sub-optimal. In fact, in this recent quest for provable security, efficiency has been left in the shadow. Notably, the tradeoffs between the security and efficiency have not been studied. In this paper, we will address these limitations, setting the "security vs. efficiency" record straight.
Concretely, by combining ideas from SKI and FO, we propose symmetric protocols that are efficient, noise-tolerant and—at the same time—provably secure against all known frauds. Indeed, our new distance-bounding solutions outperform the two aforementioned provably secure distance-bounding protocols. For instance, with a noise level of 5%, we obtain the same level of security as those of the pre-existent protocols, but we reduce the number of rounds needed from 181 to 54.

1 Introduction

As wireless technologies become more and more pervasive, being used daily in access control, remote unlocking credit-card payments and beyond, relay attacks also become a growing threat to the social acceptance of these techniques. It seems likely that nearly all wireless devices will eventually have to implement solutions to thwart these types of fraud. To defeat relay attacks, Brands and Chaum [9] introduced the notion of *distance-bounding protocols*. This relies on information being local and incapable of travelling faster than the speed of light. So, in distance-bounding, an RFID reader can assess when participants are close enough because the round-trip communication time must have been short enough. The whole idea of distance-bounding is that a *prover*, holding a key x, demonstrates that he is close to a *verifier* (who also knows this key x). The literature on distance-bounding considers several threat models.

- *Distance fraud* (DF): a far-away malicious prover tries to illicitly pass the protocol.
- *Mafia fraud* [12] (MF): a man-in-the-middle (MiM) adversary between a far-away honest prover and a verifier tries to exploit the prover's insights to make the verifier accept. (This generalizes relay attacks as not only does this adversary relay, but he may also modify the messages involved.)
- *Terrorist fraud* [12] (TF): a far-away malicious prover colludes with an adversary to make the verifier accept the adversary's rounds on behalf of this far-away prover, in such a way that the adversary gains no advantage to later pass the protocol on his own.
- *Impersonation fraud* [2]: An adversary tries to impersonate the prover to the verifier.
- *Distance hijacking* [11]: A far-away prover takes advantage of some honest, active provers (of which one is close) to make the verifier grant privileges for the far-away prover.

There are several variants and generalizations of these threats. These are briefly discussed in Appendix A.

There exist many distance-bounding protocols, but so far only the SKI protocol [5,6,7,8] and the Fischlin-Onete (FO) protocol [13] provide an all-encompassing proven security, i.e., they protect against all the above threats. (See [5, Section 2] for a recent obituary review of broken protocols.)

Organization. In Section 2, we advance revised security definitions for distance-bounding, rendering a more intuitive model, whilst maintaining backward compatibility; we also prove the latter preservation of results. In Section 3, we propose new, secure DB protocols DB1, DB2, and DB3. Section 3.2 considers the tradeoffs between security and efficiency, and presents the comparisons made in this sense.

Contribution. The contribution of this paper is threefold:

- We build up on SKI [5,6,7,8] and FO [13,17] to propose DB1, DB2, and DB3, three new distance-bounding protocols which outperform both the SKI and the FO protocols.
  For instance, to offer a false acceptance rate of under 1‰ a false rejection rate of under 1%, at a noise level of 5% during the rapid bit-exchange, DB1 (with parameter $q = 3$) requires 14/14/54 rounds for resistance to distance fraud / mafia fraud / terrorist fraud, respectively. For the same performance, SKI and FO require 84/48/181 and 84/84/? rounds[3], respectively. So, DB1 represents a *substantial improvement* in terms of efficiency, whilst maintaining provable security.
- When considering optimality amongst protocols requiring at least τ out of n correct rounds, with a challenge/response set of size q, we show security as follows:

	DF-resistance	MF-resistance	TF-resistance
DB1 ($q > 2$)	secure, optimal	secure, optimal	secure
DB2 ($q = 2$)	secure, suboptimal	secure, optimal	secure
DB3 ($q = 2$)	secure, optimal	secure, optimal	insecure

- For our security proofs, we build on recent models [6,8,17]. In doing so, we revisit the definition of mafia fraud / man-in-the-middle and the definition of terrorist fraud / collusion fraud. Thus, we provide a complete set of security definitions for distance-bounding, capturing the previous notions, but being in line with the established theory behind interactive proofs.

Useful bounds for noisy communications. Following [6,8], to assert security in noisy communications, we will make use of the tail of the binomial distribution:

$$\mathsf{Tail}(n,\tau,\rho) = \sum_{i=\tau}^{n} \binom{n}{i} \rho^i (1-\rho)^{n-i},$$

We recall that for any $\varepsilon, n, \tau, \rho$ such that $\frac{\tau}{n} < \rho - \varepsilon$, we have $\mathsf{Tail}(n,\tau,\rho) > 1 - e^{-2\varepsilon^2 n}$. For $\frac{\tau}{n} > \rho + \varepsilon$, we have $\mathsf{Tail}(n,\tau,\rho) < e^{-2\varepsilon^2 n}$. This comes from the Chernoff-Hoeffding bound [10,15].

2 Revised DB Security Model and Proofs

We now refine the security definitions and other tools from the security models in [6,8,17]. Throughout this section, we also discuss the links with the original notions.

In this paper, we concentrate on distance-bounding protocols based on symmetric cryptography (which is the overwhelmingly prevalent approach in DB).

[3] As discussed herein, FO has an incomparable approach for TF-resistance in which the number of rounds is not relevant.

Definition 1. *A (symmetric)* distance-bounding protocol *is a tuple* $(\mathcal{K}, P, V, B)$*, constructed of the following: a key domain* $\mathcal{K}$*; a two-party probabilistic polynomial-time (PPT) protocol* $(P(x), V(x))$*, where* P *is the* proving algorithm*,* V *is the* verifying algorithm*, and* x *is taken from* $\mathcal{K}$*; a distance bound* B*. At the end of the protocol, the verifier* $V(x)$ *sends a final message* Out_V*. This output denotes that the verifier* accepts *(*$\mathsf{Out}_V = 1$*) or* rejects *(*$\mathsf{Out}_V = 0$*).*

Informally, a distance-bounding protocol is complete if executing $P(x) \leftrightarrow V(x)$ on locations within a distance bounded by B makes $V(x)$ accept with overwhelming probability. The formalism is straightforward with the settings below.

We compare our protocols to *any* DB protocol that follows what we call the *common structure*.

Definition 2 (Common structure). *A DB protocol with the* common structure *based on parameters* $(n, \tau, \mathsf{num}_c, \mathsf{num}_r)$ *has some initialization and verification phases which do not depend on communication times.*[4] *These phases are separated by n rounds of timed challenge/response exchanges. This is called* the distance bounding phase*. A response is* on time *if the elapsed time between sending the challenge and receiving the response is at most* $2B$*. Provers don't measure time.*[5] *Challenges and responses are in sets of cardinality* num_c *and* num_r*, respectively.*

When the protocol follows the specified algorithms but messages during the distance bounding phase can be corrupted during transmission, we say that the protocol is τ-complete *if the verifier accepts if and only if at least* τ *rounds have a correct and on-time response.*

One can easily see that nearly *all* distance-bounding protocols in the literature fit this definition.

In practice, when the timed phase is subject to *noise*, we assume that there is a probability of p_{noise} that one round of challenge/response is corrupted. The probability that an honest prover, close to the verifier, passes the protocol is thus $\mathsf{Tail}(n, \tau, 1 - p_{\mathsf{noise}})$. So, with $\frac{\tau}{n} < 1 - p_{\mathsf{noise}}$ with a constant gap, the probability to fail is negligible, due to the Chernoff-Hoeffding bound [10,15].

Participants, Instances, Setup and Locations.

- In a DB protocol, participants can be a *prover*, a *verifier*, or *adversaries*. The prover and the verifier receive a key x which is randomly selected from the key space. We adopt a *static* adversarial model: i.e., at the beginning of the experiment, it is decided whether the prover is malicious or not. Participants have several *instances*. An instance has a *location*. It corresponds to the execution of a protocol during one session.
- A honest prover runs instances of the algorithm P denoted by $P(x)$. An instance of a malicious prover runs an arbitrary algorithm denoted by $P^*(x)$. $\mathbf{P}$ denotes the set of instances of the prover.
- The verifier is honest without loss of generality.[6] He runs instances of the algorithm V denoted by $V(x)$. $\mathbf{V}$ denotes the set of instances of the verifier.
- Other participants are (without loss of generality) malicious and may run whatever algorithm, but with no initialized key. The set of such malicious participants is denoted $\mathbf{A}$. By contrast, a designated, one such instance is denoted $\mathcal{A}$.
- Locations are elements of a metric space.

Why a Single Identity? Our definition uses a single identity, without loss of generality. This is because provers or verifiers running the protocol with other identities (and keys independent of x) could be considered as elements of $\mathbf{A}$.

[4] The verification phase can be interactive or not.

[5] Provers are in a waiting state to receive the challenge and loose the notion of time while waiting.

[6] A "malicious verifier" running an algorithm $V^*(x)$ can be seen as a malicious prover running $V^*(x)$.

Definition 3 (DB Experiment). *An* experiment exp *for a distance-bounding protocol* $(\mathcal{K}, P, V, B)$ *is a setting* $(\mathbf{P}, \mathbf{V}, \mathbf{A})$ *with several instances of* participants*, at some* locations*, set up as above, and running an overall PPT sequence.*

In the above definition, the notion of experiment implies simultaneously several *different* entities: participants, physical locations, algorithms to be run by these participants and corruption states. As such, when used inside further definitions, the notion of experiment will implicitly or explicitly, upon the case, quantify over these entities.

We further assume that communicating from a location to another takes time equal to the distance. Indeed, no one can violate the fact that communication is limited by the speed of light. Adversaries can intercept some messages and replace them by others, but must adhere to the fact that computation is local.

Ideally, one should develop a formal model to define all these. This has actually been done in [6,8]. In this paper, we keep the notions at the intuitive level, mainly due to space limitations, and since such a formal model would only be needed to prove the fundamental Lemma 4 below (which is proven in and adapted from [8, Lemma 1] and, herein, taken axiomatically).

Lemma 4 (Fundamental Lemma). *Assume an experiment in which at some point a participant* $\mathcal{V}$ *broadcasts a message c, then waits for a response r. We let E be the event that the elapsed time between sending c and receiving r is at most* $2B$*. In the experiment,* Close *is the set of all participants (except* $\mathcal{V}$*) which are within a distance of up to B from* $\mathcal{V}$*, and* Far *is the set of all participants at a larger distance. For each user U, we consider his view* $View_U$ *just before the time when U can see the broadcast message c.*

We say that a message by U is independent[7] *from c if it is the result of applying algorithm U on* $View_U$*, or on a prefix of it.*

There exists an algorithm Algo *with the following property. If E holds and r was sent from a participant in* Close*, we have* $r = \mathsf{Algo}((View_U)_{U \in \mathsf{Close}}, c, w)$*, where w is the list of all messages independent from c which are seen*[8] *by any* $U \in \mathsf{Close}$ *and not already in any* $View_U$*. If E holds and r was sent from a participant in* Far*, then the message r is independent from c.*

This lemma can be summarized as follows: a close-by participant cannot get online help from far away to answer correctly and in time to the challenge c.

Definition 5 (Distinguished Experiment). *We denote by* $\mathsf{exp}(\mathcal{V})$ *an experiment in which we fix a verifier instance* $\mathcal{V} = V(x)$ *from* $\mathbf{V}$*, which we call* distinguished verifier*. Participants which are within a distance of at most B from a distinguished verifier* $\mathcal{V}$ *are called* close-by participants*. Others are called* far-away participants*.*

Participants can move during the experiment, but not faster than the transmission of information. For simplicity, we assume that far-away participants remain far away during the experiment.

Definition 6 (α-resistance to distance fraud). *We say that a distance-bounding protocol* α-resists to distance fraud *if for any distinguished experiment* $\mathsf{exp}(\mathcal{V})$ *where there is no participant close to* $\mathcal{V}$*, the probability that* $\mathcal{V}$ *accepts is bounded by* α*.*

Compared to [8], this definition is simplified and does not capture the notion of distance hijacking; therein, a far-away malicious $P^*(x)$ can make $\mathcal{V}$ accept by taking advantage of several honest provers

[7] we stress that this is a local definition of independence which is unrelated to statistical independence.

[8] "Seen" means either received as being the destinator or by eavesdropping.

which do not hold x but are close to $\mathcal{V}$. In [8], some close-by honest participants are allowed in the definition of distance fraud resistance. However, distance hijacking could generalize to the presence of any close-by honest participant who is running a protocol (for whatever honest reason) which could match (by some weird coincidence) the response function of the malicious prover. This is not captured by the definition of [8]. Nonetheless, in most of the cases, this bizarre situation can be ignored and we can concentrate on regular distance frauds. So, we simplified on purpose our Def. 6, excluding the more corner-case fraud of distance hijacking, as this simplifies the proofs quite a lot. Nonetheless, distance hijacking and other extensions of classical frauds will be captured by the notion of soundness, which we introduce below. Overall, we will treat all threats.

Theorem 7. *A DB protocol following the common structure with parameters* $(n, \tau, \mathsf{num}_c, \mathsf{num}_r)$ *cannot* α*-resist to distance fraud for* α *lower than* $\mathsf{Tail}\left(n, \tau, \max\left(\frac{1}{\mathsf{num}_c}, \frac{1}{\mathsf{num}_r}\right)\right)$.[9]

Proof. We construct a DF following *the early-reply strategy*: a malicious prover guesses with probability $\frac{1}{\mathsf{num}_c}$ the challenge c_i before it is emitted, and then he already sends the response so that it arrives on time. The rest of the protocol is correctly simulated (with delay) after receiving the challenges. An incorrect guess would look like a round which was the victim of noise. So, the attack succeeds with probability $\mathsf{Tail}\left(n, \tau, \frac{1}{\mathsf{num}_c}\right)$. We can have a similar attack guessing the response r and succeeding with probability $\mathsf{Tail}\left(n, \tau, \frac{1}{\mathsf{num}_r}\right)$. □

While the above definition protects verifiers against malicious provers, we need an extra notion to protect the honest prover against men-in-the-middle. This is as follows.

Definition 8 (β-secure distance-bounding protocol). *We say that a distance-bounding protocol is* β-secure *if for any distinguished experiment* $\mathsf{exp}(\mathcal{V})$ *where the prover is honest, and the prover instances are all far-away from* $\mathcal{V}$*, the probability that* $\mathcal{V}$ *accepts is bounded by* β.

Intuitively, this notion protects honest provers from identity theft. It implies that x cannot be extracted by a malicious participant; this is along the same lines as in zero-knowledge interactive protocols. This notion of security also captures resistance to relay attacks, mafia fraud, and man-in-the-middle attacks. The advantage of Def. 8 over the resistance to man-in-the-middle attacks, as it was defined in [6,8, Def. 4], is that we no longer need to formalize a *learning phase*, although we can easily show we capture these notions as well. Our definition is therefore simpler.

Theorem 9. *A DB protocol following the common structure with parameters* $(n, \tau, \mathsf{num}_c, \mathsf{num}_r)$ *cannot be* β*-secure for* β *lower than* $\mathsf{Tail}\left(n, \tau, \max\left(\frac{1}{\mathsf{num}_c}, \frac{1}{\mathsf{num}_r}\right)\right)$.[10]

Proof. We consider $\mathcal{V}$ and a far-away instance of the prover P, and a close-by MiM $\mathcal{A}$. In the initialization phase and the verification phase, $\mathcal{A}$ passively relays messages between $\mathcal{V}$ and P. During the challenge phase, and in the *pre-ask strategy*, $\mathcal{A}$ guesses the challenge before it is released and asks for the response to P on time so that he can later on answer to $\mathcal{V}$. Clearly, the attack succeeds with probability $\mathsf{Tail}\left(n, \tau, \frac{1}{\mathsf{num}_c}\right)$. We can have a similar attack with a *post-ask strategy* where $\mathcal{A}$ guesses the response at the same time he forwards the challenge to P. This succeeds with probability $\mathsf{Tail}\left(n, \tau, \frac{1}{\mathsf{num}_r}\right)$.[11] □

[9] In [18], a protocol with two bits of challenges and one bit of response achieving $\alpha = \mathsf{Tail}(n, \tau, \frac{1}{3})$ is proposed. But it actually works with $\mathsf{num}_r = 3$ as it allows response 0, response 1, and no response.

[10] Same remark about [18] as in Th. 7.

[11] Since provers loose the notion of time in the challenge phase, pre-ask and post-ask attacks cannot be detected.

The definition below is adapted from [17]. One difference is that γ' is no longer necessarily $1-\mathsf{negl}$. It also considers extractors just passing the protocol, instead of having to produce the secret; this is clearly more general. Our protocols herein will make the secret extractable though.

Definition 10 (**(γ,γ',m)-soundness).** *We say that a distance-bounding protocol is* (γ,γ',m)-sound *if for any distinguished experiment* $\mathsf{exp}(\mathcal{V})$ *in which* $\mathcal{V}$ *accepts with probability at least* γ*, there exists a PPT algorithm* $\mathcal{E}$ *called* extractor*, with the following property. By* $\mathcal{E}$ *running experiment* $\mathsf{exp}(\mathcal{V})$ *several times, in some executions denoted* $\mathsf{exp}_i(\mathcal{V})$*,* $i=1,\ldots,M$*, for* M *of expected value bounded by* m*, we have that*

$$\Pr\left[\mathsf{Out}_V = 1 : \mathcal{E}(\mathsf{View}_1,\ldots,\mathsf{View}_M) \leftrightarrow \mathcal{V} \middle| \mathsf{Succ}_1,\ldots,\mathsf{Succ}_M\right] \geq \gamma',$$

where View_i *denotes the view of all close-by participants (except* $\mathcal{V}$*) and the transcript seen by* $\mathcal{V}$ *in the run* $\mathsf{exp}_i(\mathcal{V})$*, and* Succ_i *is the event that* $\mathcal{V}$ *accepts in the run* $\mathsf{exp}_i(\mathcal{V})$*.*

In other words, the extractor impersonates the prover to $\mathcal{V}$.[12] In more details, this means that having $\mathcal{V}$ accept in run $\mathsf{exp}_i(\mathcal{V})$ implies the following: a piece of x was given to the close-by participants and it is stored in View_i, and that m such independent pieces, on average, could allow $\mathcal{E}$ to impersonate $P(x)$ to $\mathcal{V}$. This notion is pretty strong as it could offer a guaranty against distance hijacking: a prover making such attack would implicitly leak his credentials.

3 New Highly Efficient, Symmetric Distance-Bounding Protocols

In the idea to outperform SKI and FO, we now advance a family of provably secure symmetric distance-bounding protocols, called DBopt. It includes DB1, DB2, and DB3. Indeed, we will see herein that DB1 is in fact optimal in terms of distance-fraud resistance and security with non-binary challenges. The DB2 and DB3 variants are motivated by the use of binary challenges, which is customary in distance-bounding designs. Whilst DB2 is suboptimal, it still performs well, almost always, i.e., better than SKI and FO. DB3 is optimal but not TF-resistant. The eager reader can directly inspect the performance/security graphs on Fig. 3, page 13, where we plot the (logs of) fraud-resistance thresholds, i.e., $-\log_2\alpha$, $-\log_2\beta$, and $-\log_2\gamma$.

3.1 DBopt

We propose DBopt, a new family of symmetric distance-bounding protocols, as depicted on Fig. 1. It combines ideas taken from SKI [5,6,7,8] and the Swiss-Knife protocol [16] (as used by FO [13]). We use a security parameter s (the length of the secret x, i.e., $x \in \mathcal{K} = \mathbf{Z}_2^s$) and the following parameters based on s: the number of rounds n, the length ℓ_{tag} of tag, a threshold τ, the nonce length ℓ_{nonce}, and a constant q which is a prime power, e.g., $q=2$, $q=3$, or $q=4$. DBopt follows the *common structure* with parameters n, τ, and $\mathsf{num}_c = \mathsf{num}_r = q$.

As in SKI, we assume $L_\mu(x) = (\mu(x),\ldots,\mu(x))$ for some function $x \mapsto \mu(x)$, but μ is not necessarily linear. Concretely, μ is a vector in $\mathbf{Z}_2^s$ and map a *fixed* injection from $\mathbf{Z}_2$ to $\mathsf{GF}(q)$. Hence, $\mu(x) = \mathsf{map}(\mu \cdot x)$ maps a bitstring x to a $\mathsf{GF}(q)$-representation of the bit obtained by the scalar product $\mu \cdot x$. We let $\mathcal{L}$ denote the set of all such possible L_μ mappings (map being fixed). The function f_x maps to different codomains, depending on its inputs: given two nonces N_P and N_V, $L_\mu \in \mathcal{L}$, and $b,c \in \mathsf{GF}(q)^n$, $f_x(N_P,N_V,L_\mu,b) \in \mathsf{GF}(q)^n$ and $f_x(N_P,N_V,L_\mu,b,c) \in \mathsf{GF}(q)^{\ell_{\mathsf{tag}}}$.

[12] Note that cases where there is a close-by prover or a close-by verifier are trivial since they hold the secret x in their view.

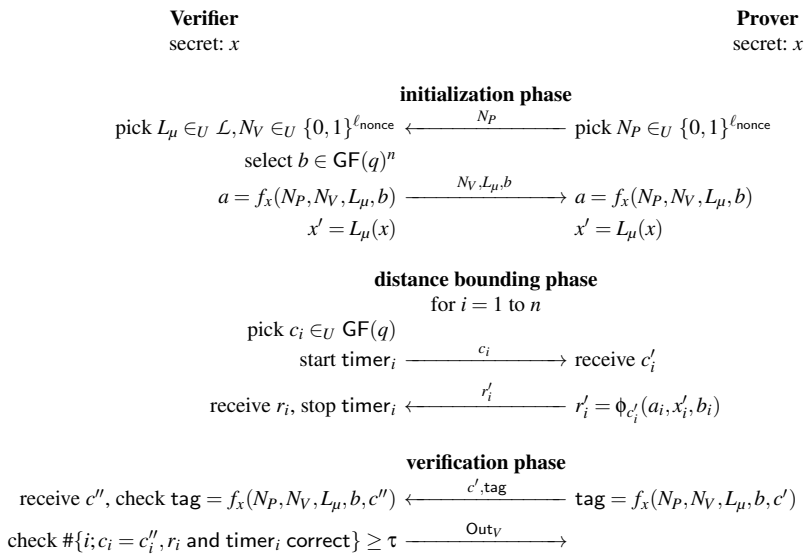

Fig. 1. The DBopt Distance-Bounding Protocols

During the initialization, the participants exchange some nonces N_P, N_V, some $L_\mu \in \mathcal{L}$, and a vector b. The vector b could be fixed in the protocol, but is subject to some constraints as detailed below. V and P compute $a = f_x(N_P, N_V, L_\mu, b)$ and $x' = L_\mu(x)$. In the distance bounding phase, the response function is a linear function $r_i = \phi_{c_i}(a_i, x'_i, b_i)$ defined by the challenge c_i. The verification checks that the participants have seen the same challenges (based on the tag computed by $\mathsf{tag} = f_x(N_P, N_V, L_\mu, b, c)$), counts the number of rounds with a correct and timely response, and accepts if there are at least τ of them.

There are some specificities in each protocol for the selection of q, map, b, and ϕ_{c_i} which are summarized in the following table:

protocol	q	map	b	ϕ_{c_i}
DB1	$q > 2$	$\mathsf{map}(u) \neq 0$	no b used	$\phi_{c_i}(a_i, x'_i, b_i) = a_i + c_i x'_i$
DB2	$q = 2$	$\mathsf{map}(u) = u$	Hamming weight $\frac{n}{2}$	$\phi_{c_i}(a_i, x'_i, b_i) = a_i + c_i x'_i + c_i b_i$
DB3	$q \geq 2$	no map used	Hamming weight n	$\phi_{c_i}(a_i, x'_i, b_i) = a_i + c_i b_i$

Specifically, DB3 is the simplest protocol and is optimal, but it offers no soundness. DB2 works with binary challenges and responses, but it is not optimal. DB1 is optimal but needs $q \geq 3$ since it requires that map is injective from $\mathbf{Z}_2$ to $\mathsf{GF}(q)^*$. They are depicted on Fig. 4–6 in Appendix.

Overall, DBopt is very similar to SKI. Like in SKI, the leak vector x' is fundamental for soundness: the vector x' encodes $\mu \cdot x$, which leaks if the prover reveals his response function. We added a verification step, as in FO (it actually comes from the Swiss-Knife protocol [16]). This verification allows to use better response functions: thanks to the above extra verification, the response function needs no longer resist men-in-the-middle playing with different challenges on the sides of P and V, as it was the case in [1,3]. One particularity is that DB1 mandates $x'_i \neq 0$ so cannot accommodate $q = 2$. If we want $q = 2$, we need for DF-resistance to make sure that r_i really depends on c_i, by introducing the vector b in which exactly half of the coordinates are 0. DB2 can be optimized into DB3 by using $r_i = a_i + c_i$ (so x' unused and $b_i = 1$ for all i) by sacrificing soundness.

DBopt is clearly τ-complete following Def. 2.

Theorem 11 (DF-resistance). *The* DBopt *protocols* α*-resists to distance fraud for*

- *(DB1 and DB3)* $\alpha = \mathsf{Tail}(n, \tau, \frac{1}{q})$ *which is negligible for* $\frac{\tau}{n} > \frac{1}{q} + \mathsf{cte}$*;*
- *(DB2)* $\alpha = \mathsf{Tail}(\frac{n}{2}, \tau - \frac{n}{2}, \frac{1}{2})$ *which is negligible for* $\frac{\tau}{n} > \frac{3}{4} + \mathsf{cte}$.

Due to Th. 7, DB1 and DB3 are optimal for DF-resistance. DB2 is clearly not optimal (as DB3 is better with the same $q = 2$). However, the bound is tight for DB2 as the DF guessing the response matches the α bound: the malicious prover always wins the rounds for which $x' = b_i$ (that is: exactly half of the rounds due to the Hamming weight of b) by sending the response in advance and passes with probability $\alpha = \mathsf{Tail}(\frac{n}{2}, \tau - \frac{n}{2}, \frac{1}{2})$.

Proof. We consider a distinguished experiment $\mathsf{exp}(\mathcal{V})$ with no close-by participant. Due to the distance, the answer r_i to $\mathcal{V}$ comes from far away. Thanks to Lemma 4, it is independent (in the sense of Lemma 4) from c_i. For DB1, since $x'_i \neq 0$ by construction, r_i equals $a_i + c_i x'_i$ with probability $\frac{1}{q}$. The same goes for DB3. For DB2, thanks to the selection of b, this holds for exactly half of the rounds: those such that $x'_i + b_i \neq 0$. So, the probability to succeed in the experiment is bounded as stated. □

As shown in [4], we cannot rely on the PRF assumption alone for DB1 or DB2, since the secret is used as a key of f_x and also outside f_x in x'. The circular-PRF assumption guarantees the PRF-ness of f, even when we encrypt a function $L_\mu(x)$ of the key. In Appendix B, we recall and extend the notion, to accommodate DB1 and DB2.

Theorem 12 (Security). *The* DBopt *protocols are* β*-secure for*

- *(DB1 and DB2)* $\beta = \mathsf{Tail}(n, \tau, \frac{1}{q}) + \frac{r^2}{2} 2^{-\ell_{\mathsf{nonce}}} + (r+1)\varepsilon + r 2^{-\ell_{\mathsf{tag}}}$ *when* f *is a* (ε, T)*-circular-PRF (as defined by Def. 15);*
- *(DB3)* $\beta = \mathsf{Tail}(n, \tau, \frac{1}{q}) + \frac{r^2}{2} 2^{-\ell_{\mathsf{nonce}}} + \varepsilon + 2^{-\ell_{\mathsf{tag}}}$ *when* f *is a* (ε, T)*-PRF.*

There, r *is the number of honest instances of the prover and* T *is a complexity bound on the experiment.* β *is negligible for* $\frac{\tau}{n} > \frac{1}{q} + \mathsf{cte}$*,* r *and* T *polynomially bounded, and* ε *negligible.*

Based on that $\frac{r^2}{2} 2^{-\ell_{\mathsf{nonce}}} + (r+1)\varepsilon + r 2^{-\ell_{\mathsf{tag}}}$ (or the similar term for DB3) can be made negligible against β, DB1, DB2, and DB3 are *optimal* for security due to Th. 9.

Proof. We consider a distinguished experiment $\mathsf{exp}(\mathcal{V})$ with no close-by $P(x)$, no $P^*(x)$, and where $\mathcal{V}$ accepts with probability p. We consider a game Γ_0 in which we simulate the execution of $\mathsf{exp}(\mathcal{V})$ and succeed if and only if Out_V by $\mathcal{V}$ is an acceptance message. Γ_0 succeeds with probability p.

First of all, we reduce to the same game Γ_1 whose success additionally requires that for every (N_P, N_V, L_μ) triplet, there is no more than one instance $P(x)$ and one instance $V(x)$ using this triplet. Since $P(x)$ is honest and selecting the ℓ_{nonce}-bit nonce N_P at random and the same for $V(x)$ selecting N_V, by looking at the up to $\frac{r^2}{2}$ pairs of $P(x)$'s or of $V(x)$'s and the probability that one selection of a nonce repeats, this new game succeeds with probability at least $p - \frac{r^2}{2} 2^{-\ell_{\mathsf{nonce}}}$.

Then, for DB1 and DB2, we outsource the computation of every $a_i + c x'_i$ to the oracle

$$O_{x,f_x}(y, L_\mu, A, B) = (A \cdot L_\mu(x)) + (B \cdot f_x(y))$$

as in Def. 15, with $y = (N_P, N_V, L_\mu, b)$, $A \cdot L_\mu(x) = c(L_\mu(x))_i$, and $B \cdot f_x(y) = (f_x(y))_i$. I.e., $A_i = c e_i$ and $B_i = e_i$, where e_i is the vector having a 1 on its ith component and 0 elsewhere. This can be used with $c = c'_i$ by $P(x)$ (for computing r'_i) or with $c = c_i$ by $V(x)$ (for verifying r_i). Similarly, the computation (by $P(x)$ or $V(x)$) of $\mathsf{tag} = f_x(y)$ can be made by several calls of form $O_{x,f_x}(y, L_\mu, 0, B)$. (We note that the y in this case has incompatible form with the y in the r_i computation.) So, every computation

requiring x is outsourced. Note that queries to the same y must use the same L_μ since this is part of y. So, the first condition in Def. 15 to apply the circular-PRF assumption is satisfied. We consider the event E that there exists in the game some sequence (y, L_μ, A_j, B_j) of queries to O_{x,f_x} sharing the same (y, L_μ) and some λ_j's such that $\sum_j \lambda_j B_j = 0$ and $\sum_j \lambda_j A_j \neq 0$. We need to restrict to the event $\neg E$ to apply Def. 15. We consider the event E' that one instance in $\mathbf{V}$ receives a valid tag which was not computed by the prover $\mathbf{P}$ (i.e., it was forged).

Let c_i'' be the value received by $V(x)$ in the verification phase. We assume that V checks that tag is correct, timer_i is correct, and $c_i = c_i''$, then queries $O_{x,f_x}(y, L_\mu, c_i e_i, e_i)$ only if these are correct. If E happens for some (y, L_μ), due to the property of Γ_1, each i has at most two queries. Since $B_j = e_{i_j}$, $\sum_j \lambda_j B_j = 0$ yields pairs of values j and j' such that $i_j = i_{j'} = i$, $A_j = c_i e_i$, $A_{j'} = c_i' e_i$, $B_j = B_{j'} = e_i$, and $\lambda_j + \lambda_{j'} = 0$. The event E implies that there exists one such pair such that $\lambda_j A_j + \lambda_{j'} A_{j'} \neq 0$. So, $c_i \neq c_i'$. But since V only queries if c_i'' and tag are correct, we have $c_i = c_i'' \neq c_i'$ and tag correct. So, $\mathcal{V}$ must have accepted some tag which was not computed by $P(x)$. So, E implies E'. We now show that $\Pr[E']$ is negligible.

We define Γ_2, the variant of Γ_1, which in turn requires that E' does not occur as an extra condition for success. We let E_j' be the event that tag_j, the jth value tag received by any $V(x)$ in $\mathbf{V}$ is forged. Let $\Gamma_{1,j}$ be the hybrid of Γ_1 stopping right after tag_j is received and succeeding if E_j' occurs but not $E_1', \ldots, E_{j-1}'$.

Clearly, since $E_1' \cup \cdots \cup E_{j-1}'$ does not occur and we stop right after reception of tag_j, E cannot occur. (Remember that for E to occur for the first time upon a query to O_{x,f_x}, there must be a prior tag which was forged.) So, the conditions to apply the circular-PRF security reduction in Def. 15 is satisfied in $\Gamma_{1,j}$. We apply the circular-PRF assumption and replace O_{x,f_x} by $O_{\tilde{x},F}$, loosing some probability ε. We obtain a game $\Gamma_{2,j}$. Clearly, $\Gamma_{2,j}$ succeeds with probability bounded by $2^{-\ell_{\mathsf{tag}}}$ because F is random. So, $\Pr_{\Gamma_{1,j}}[\mathsf{success}] \leq \varepsilon + 2^{-\ell_{\mathsf{tag}}}$ in $\Gamma_{1,j}$.

So, $\Pr[E']$ is bounded by the sum of all $\Pr_{\Gamma_{1,j}}[\mathsf{success}]$, i.e. $\Pr_{\Gamma_1}[E'] \leq r\varepsilon + r2^{-\ell_{\mathsf{tag}}}$ since the number of hybrids is bounded by r. Hence, $\Pr_{\Gamma_2}[\mathsf{success}] \geq p - \frac{r^2}{2} 2^{-\ell_{\mathsf{nonce}}} - r\varepsilon - r2^{-\ell_{\mathsf{tag}}}$.

Now, in the whole game Γ_2 where E' does not occur, we replace O_{x,f_x} by $O_{\tilde{x},F}$ and obtain the simplified game Γ_3. We have $\Pr_{\Gamma_3}[\mathsf{success}] \geq p - \frac{r^2}{2} 2^{-\ell_{\mathsf{nonce}}} - (r+1)\varepsilon - r2^{-\ell_{\mathsf{tag}}}$.

It is now easy to analyze the protocol Γ_3. Thanks to Lemma 4, the response is computed based on information from $P(x)$ (w in Lemma 4) which is independent (in the sense of Lemma 4) from the challenge. Either $P(x)$ was queried with a challenge before, but this could only match the correct one with probability $\frac{1}{q}$ and the adversary would fail with tag otherwise. Or, $P(x)$ leaked nothing about the response to this challenge, and the answer by the adversary can only be correct with probability $\frac{1}{q}$. In any case, his answer is correct with probability $\frac{1}{q}$. So, Γ_3 succeeds with probability up to $\mathsf{Tail}(n, \tau, \frac{1}{q})$.

To sum up, we have $p \leq \mathsf{Tail}(n, \tau, \frac{1}{q}) + \frac{r^2}{2} 2^{-\ell_{\mathsf{nonce}}} + (r+1)\varepsilon + r2^{-\ell_{\mathsf{tag}}}$ for DB1 and DB2.

For DB3, we loose $\frac{r^2}{2} 2^{-\ell_{\mathsf{nonce}}}$ from Γ_0 to Γ_1. In Γ_1, we apply the full PRF reduction and loose ε to obtain Γ_2 with a random function. We loose $2^{-\ell_{\mathsf{tag}}}$ more to assume that tag received by $\mathcal{V}$ was not forged in some Γ_3. Then, it is easy to see that either the prover was queried before c_i was known, but this will only succeed if c_i was correctly guessed, or it was queries after, but this will only succeed if the answer r_i was correctly guessed. So, Γ_3 succeeds with a probability bounded by $\mathsf{Tail}(n, \tau, \frac{1}{q})$. (Note that DB3 is insecure without the authenticating tag: the man-in-the-middle can just run the DB phase with the prover, deduce a, then answer all challenges from the verifier.) □

Theorem 13 (Soundness of DB1). *The DB1 scheme is* $(\gamma,\gamma',s+2)$*-sound for any* $\gamma \geq \frac{q}{q-1}p_B$ *and* γ' *such that* $\gamma' = (1-\gamma^{-1}p_B)^s$*, where* $p_B = \max_{a+b\leq n} p_B(a,b)$ *and*

$$p_B(a,b) = \sum_{u+v\geq\tau-a} \binom{n-a-b}{u}\binom{b}{v}\left(1-\frac{1}{q}\right)^{b+u-v}\left(\frac{1}{q}\right)^{n-a-b-u+v}$$

More precisely, any collusion fraud with a success probability $\gamma \geq \frac{p_B}{1-\frac{1}{q}-\varepsilon}$ *leaks one random* $(\mu,\mu\cdot x)$ *pair with probability at least* $\frac{1}{q}+\varepsilon$*. Assuming* $p_B = p_B(0,0)$,[13] *this compares* γ *to* $\frac{q}{q-1}\mathsf{Tail}(n,\tau,\frac{q-1}{q})$.

For instance, for $\gamma = sp_B$ and $\frac{\tau}{n} > \frac{q}{q-1} + \mathsf{cte}$, γ is negligible and γ' is greater than a constant.

If we applied the same proof as for SKI from [17, Th.14], we would not get such a good result. We would rather obtain $\mathsf{Tail}(\frac{n}{2},\tau-\frac{n}{2},\frac{q-1}{q})$. So, our proof of Th. 13 is substantially improved.

Proof. We consider a distinguished experiment $\mathsf{exp}(\mathcal{V})$ where $\mathcal{V}$ accepts with probability $p \geq \gamma$.

The verifier $\mathcal{V}$ has computed some a and x'. We apply Lemma 4. We let $\mathsf{Resp}_i(c)$ be the value of the response r_i arriving to $\mathcal{V}$ when c_i is replaced to c in the simulation. We show below that we can always compute $\mathsf{Resp}_i(c) - \mathsf{Resp}_i(c')$ for any (c,c') pair from a straightline simulation (i.e., without rewinding). Let View_i be the view of close-by participants $\mathcal{A}$ until the time before c_i arrives, and w_i be the extra information (independent from c_i, in the sense of Lemma 4) arriving from far-away. Due to Lemma 4, we have $\mathsf{Resp}_i(c) = \mathsf{Algo}(\mathsf{View}_i, c, w_i)$. So, we can easily compute $\mathsf{Resp}_i(c) - \mathsf{Resp}_i(c')$ without rewinding. The answer by a far-away participant is independent from c_i, so $\mathsf{Resp}_i(c) - \mathsf{Resp}_i(c') = 0$: we can compute $\mathsf{Resp}_i(c) - \mathsf{Resp}_i(c')$ as well.

We say that c is correct in the ith round if $\mathsf{Resp}_i(c) = a_i + cx'_i$. We let C_i be the set of correct c's for the ith round. We let S be the set of all i's such that $c_i \in C_i$. Finally, we let R (resp. R') be the set of all i's for which $\#C_i = q$ (resp. $\#C_i \leq 1$). I.e., all c's are correct in the ith round for $i \in R$ and at most one is correct for $i \in R'$.

By definition, the probability that $\#S \geq \tau$ is $p \geq \gamma$. We see that $\frac{\mathsf{Resp}_i(c)-\mathsf{Resp}_i(c')}{c-c'} = x'_i$ if $i \in R$, for any $c \neq c'$. If the left-hand side leads to the same value ξ_i for each $c \neq c'$, we say that the round i votes for $x'_i = \xi_i$. If the (c,c') pairs do not lead to the same value in $\mathsf{GF}(q)$, we say that the round i does not vote. So, we can always compute the vote ξ_i from the views of close-by participants. The majority of the available $\mathsf{map}^{-1}(\xi_i)$ shall decode $\mu\cdot x$.

For DB1, we can prove that if the round i votes for some ξ_i such that $\xi_i \neq x'_i$, then we must have $i \in R'$. Indeed, if round i votes for some ξ_i and $\#C_i \geq 2$, it means that there exist two different challenges c and c' such that the responses $\mathsf{Resp}_i(c)$ and $\mathsf{Resp}_i(c')$ are correct. So, $\mathsf{Resp}_i(c) = a_i + cx'_i$ and $\mathsf{Resp}_i(c') = a_i + c'x'_i$. The vote ξ_i is $\frac{\mathsf{Resp}_i(c)-\mathsf{Resp}_i(c')}{c-c'}$ which is thus equal to x'_i. So, an incorrect vote cannot have two correct challenges: it must be for $i \in R'$. The majority of the votes does not give x'_i only when $\#R \leq \#R'$. So, we shall bound $\Pr[\#R \leq \#R']$.

Let $I, I' \subseteq \{1,\ldots,n\}$ such that $\#I \leq \#I'$ and $I \cap I'$ is empty. Let $p_B(a,b)$ be the probability that at least τ rounds succeed, when we know that a rounds succeed with probability 1, b rounds succeed with probability $\frac{1}{q}$, and the other succeed with probability $1-\frac{1}{q}$. We have $\Pr[\#S \geq \tau, R = I, R' = I'] = \Pr[\#S \geq \tau | R = I, R' = I']\Pr[R = I, R' = I']$ and $\Pr[\#S \geq \tau | R = I, R' = I'] \leq p_B(\#I, \#I') \leq p_B$ since we have $\#I$ correct rounds for sure and it remains to pick u correct challenges (out of at most $q-1$) among the $i \notin I \cup I'$ rounds, and v correct challenges (out of at most 1) among the $i \in I'$ rounds, for all u and v such that $u+v \geq \tau - \#I$. By summing over all choices for I and I', we obtain that $\Pr[\#S \geq \tau, \#R \leq \#R'] \leq p_B$ So, $\Pr[\#R > \#R' | \#S \geq \tau] \geq 1-\gamma^{-1}p_B$. So, when the experiment succeeds,

[13] this is actually confirmed by experiment for the data we use.

the extracting algorithm gets a random pair $(\mu, \mu \cdot x)$ with probability at least $1-\gamma^{-1}p_B$. This is better than just guessing $\mu \cdot x$ when $\gamma > \frac{q}{q-1}p_B$.

We can do M many such accepting experiments, collect some $(\mu, \mu \cdot x)$ until we have s vector μ spanning $\mathsf{GF}(q)^s$, and reconstruct x with probability at least $\gamma' = \left(1-\gamma^{-1}p_B\right)^s$. The probability that m samples in $\mathsf{GF}(q)^s$ do not generate this space is $p_m \leq q^{s-m}$ (the number of hyperplane, q^s-1 times the probability that the m samples are all in this hyperplane, which is q^{-1} to the power m). So, the expected M until we generate the space is bounded by $s + \sum_{m \geq s} q^{s-m} \leq s+2$. Hence, after at most $s+2$ iterations on average, we can recover x by solving a linear system. This defines the extractor.

We can also push the extraction further when $1-\gamma^{-1}p_B > \frac{1}{q}$ by solving an instance of the Learning Parity with Noise problem (LPN), which would still be feasible by the practical parameters s. Extraction can also work with a complexity overhead bounded by $O(s^j)$ and a probability of at least

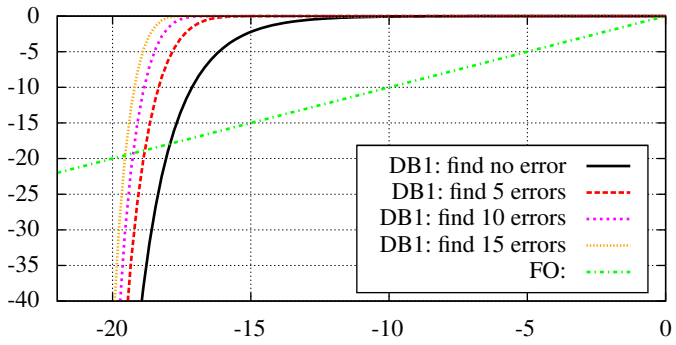

Fig. 2. $\log_2 \gamma'$ in terms of $\log_2 \gamma$ for FO and DB1 with $q=3$, $s=80$, $n=92$, $\tau=82$.

$\gamma' = \mathsf{Tail}(s, s-j, 1-\gamma^{-1}p_B)$, by finding at most j errors by exhaustive search or LPN solving algorithms. On Fig. 2 we plot $-\log_2 \gamma'$ in terms of $-\log_2 \gamma$ for an instance of DB1 and $j \in \{0,5,10,15\}$. This is why we say that γ compares to $\frac{q}{q-1}p_B$ (which is 2^{-20} on Fig. 2) in practice.

The maximum $p_B = p_B(a,b)$ is always reached for $a=b$. Indeed, for all the values plotted on Fig. 3 with $n \geq 6$, we saw it was reached for $a=b=0$. In this case, we have $p_B = p_B(0,0) = \mathsf{Tail}(n, \tau, \frac{q-1}{q})$. □

Below, we prove that the result is tight for DB1 using $q=3$. Whether this it tight for other q is an open question. Whether it is optimal for protocols following the common structure is also open.

DB1's tightness of the soundness proof. To show that the result is tight for DB1 with $q=3$, we mount a (non-leaking) terrorist fraud succeeding with probability $\gamma = \mathsf{Tail}(n, \tau, \frac{q-1}{q})$: let the malicious prover give to the adversary the tables for $c_i \mapsto r_i + e_i(c_i)$ for every round i. For each such i, randomly pick one entry for which $e_i(c_i)$ is a random *nonzero* value and let it be 0 for other, two entries. With such tables as a response function, the adversary passes the DB phase with probability γ. (Other phases are done by relaying messages.) Since the verifier accepts with negligible probability γ, the adversary learns as much as if Out_V was always set to 0.

For $q=3$ and each i, based on random $a_i \in \mathsf{GF}(q)$, $x'_i \in \mathsf{GF}(q)^*$, and $c \mapsto e_i(c)$ as distributed above, we can easily see that the distribution of the transmitted table is independent from x'_i: for $x'_i = 1$, the

table of $c \mapsto a_i + cx'_i$ defined by a random a_i is randomly picked from

$$\begin{pmatrix} 0\mapsto 0 \\ 1\mapsto 1 \\ 2\mapsto 2 \end{pmatrix}, \begin{pmatrix} 0\mapsto 1 \\ 1\mapsto 2 \\ 2\mapsto 0 \end{pmatrix}, \begin{pmatrix} 0\mapsto 2 \\ 1\mapsto 0 \\ 2\mapsto 1 \end{pmatrix}.$$

When adding the random table $e_i(c)$, it becomes a uniformly distributed random table among those with an output set of cardinality 2. For $x'_i = 2$, the table of $a_i + cx'_i$ is randomly picked from

$$\begin{pmatrix} 0\mapsto 0 \\ 1\mapsto 2 \\ 2\mapsto 1 \end{pmatrix}, \begin{pmatrix} 0\mapsto 2 \\ 1\mapsto 1 \\ 2\mapsto 0 \end{pmatrix}, \begin{pmatrix} 0\mapsto 1 \\ 1\mapsto 0 \\ 2\mapsto 2 \end{pmatrix},$$

but adding $e_i(c)$ leads to the same distribution as for $x'_i = 1$. So, the above attack does not leak and is a valid terrorist fraud. Th. 13 essentially says that there is no valid terrorist fraud with a larger γ. So, the result is tight for DB1 with $q = 3$.

The same proof technique leads to the following result for DB2.

Theorem 14 (Soundness of DB2). *For $\frac{\tau}{n} > \frac{3}{4}$, the DB2 scheme is $(\gamma, \gamma', s+2)$-sound for any $\gamma \geq 2\mathsf{Tail}(\frac{n}{2}, \tau - \frac{n}{2}, \frac{1}{2})$ and $\gamma' = (1 - \gamma^{-1}\mathsf{Tail}(\frac{n}{2}, \tau - \frac{n}{2}, \frac{1}{2}))^s$.*

Again, it is open whether this is optimal for a protocol with binary challenges. The bound is pretty tight for DB2: a malicious adversary could leak the $c_i \mapsto r_i$ tables for a random selection of half of the rounds, and leak the table with one bit flipped for the others. This will not leak x'_i and will pass with probability $\gamma = \mathsf{Tail}(\frac{n}{2}, \tau - \frac{n}{2}, \frac{1}{2})$.

3.2 Performance Comparisons

Fig. 3 plots the resistance of DB1, DB2, and DB3 compared with the protocols SKI [5,6,7,8] and FO [13].[14] In these figures, we assume a noise level of $p_{\mathsf{noise}} = 5\%$ and we adjust τ in terms of the number of rounds n such that $\mathsf{Tail}(n, \tau, 1 - p_{\mathsf{noise}}) \approx 99\%$, for τ-completeness; i.e., we admit a false rejection rate if below 1%. We plot then $-\log_2 \alpha$, $-\log_2 \beta$, and $-\log_2 \gamma$ in terms of n, assuming that the residual terms (such as ε and $2^{-\ell_{\mathsf{tag}}}$ from the PRF and $2^{-\ell_{\mathsf{nonce}}}$ from the nonce) can be neglected. We used the following dominant security parameters:

protocol	α	β	γ
SKI	$\mathsf{Tail}(n, \tau, 3/4)$	$\mathsf{Tail}(n, \tau, 2/3)$	$\mathsf{Tail}(\frac{n}{2}, \tau - \frac{n}{2}, 2/3)$
FO	$\mathsf{Tail}(n, \tau, 3/4)$	$\mathsf{Tail}(n, \tau, 3/4)$	n/a
DB1	$\mathsf{Tail}(n, \tau, 1/q)$	$\mathsf{Tail}(n, \tau, 1/q)$	$\frac{q}{q-1}\mathsf{Tail}(n, \tau, 1 - 1/q)$
DB2	$\mathsf{Tail}(\frac{n}{2}, \tau - \frac{n}{2}, 1/2)$	$\mathsf{Tail}(n, \tau, 1/2)$	$\mathsf{Tail}(\frac{n}{2}, \tau - \frac{n}{2}, 1/2)$
DB3	$\mathsf{Tail}(n, \tau, 1/2)$	$\mathsf{Tail}(n, \tau, 1/2)$	n/a

As we can see, our protocols are better than SKI and FO on *all* curves.

DB3 is not plotted on the third graph since it is not sound. FO has an incomparable TF-resistance notion and is not plotted either. Indeed, the FO curve cannot identify any threshold γ like other protocols (see Fig. 2). TF-resistance therein follows another philosophy: in order to pass a DB run, the FO protocol always leaks with a probability $\gamma' = \gamma$, no matter the number of rounds. Although this is an interesting idea, the price to pay is a much lower resistance to man-in-the-middle, as observed in [17].

Since we consider online attacks, security levels of 2^{-10} or 2^{-20} should suffice, i.e., better (online) security may be ambitious. We now report the minimal number of rounds to attain such security:

[14] We take the FO protocol as described in [17] since the original one from [13] introduces two counters and has an incorrect parameter p_e. The one from [17] has been shown to provide an optimal expression for p_e.

	security level 2^{-10}		
	DF	security	soundness
SKI	84	48	181
FO	84	84	n/a
DB1 $q=3$	14	14	54
DB1 $q=4$	12	12	91
DB2	69	24	79
DB3	24	24	n/a

	security level 2^{-20}		
	DF	security	soundness
SKI	151	91	315
FO	151	151	n/a
DB1 $q=3$	24	24	92
DB1 $q=4$	20	20	152
DB2	123	43	131
DB3	43	43	n/a

Interpretation of results. As we can see in the table above, DB1 with $q=4$ is the best choice for distance fraud and security. Unfortunately, its (non-tightly) proven soundness requires more rounds. DB1 with $q=3$ seems to be the best compromise. But if we want to use binary challenges, we shall choose between DB2 (suboptimal for DF-resistance) and DB3 (not sound).

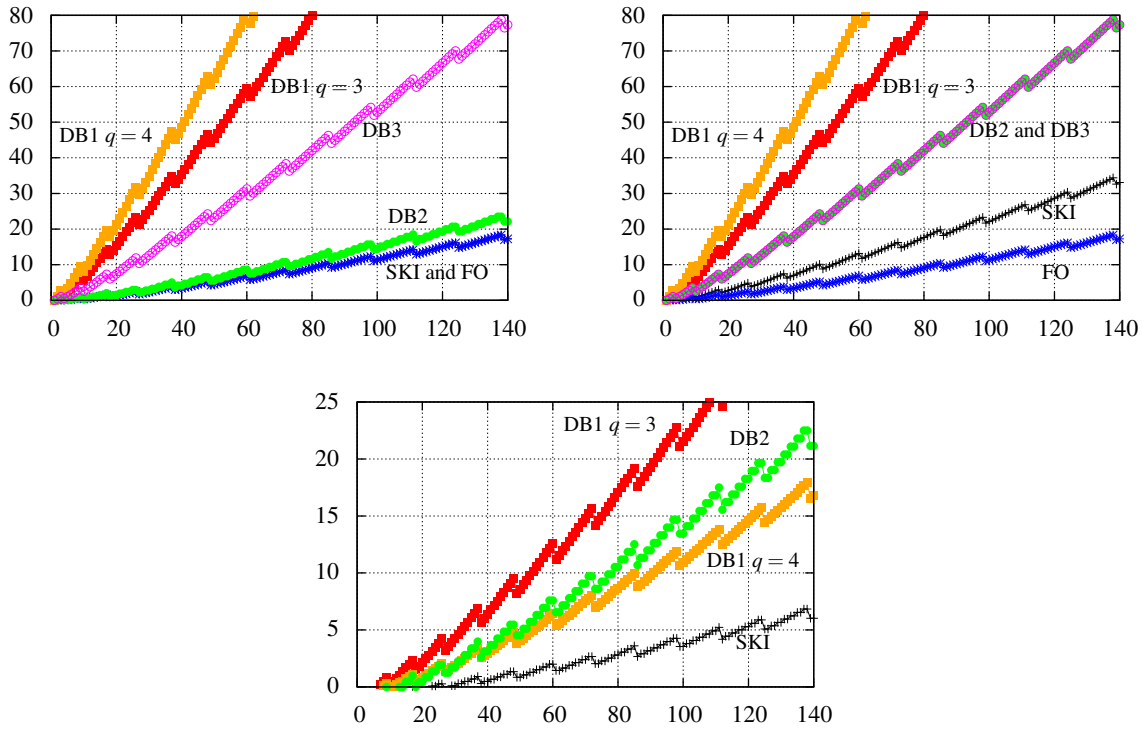

Fig. 3. Distance fraud resistance (top left) and security (top right), in equivalent bitlength, with respect to the number of rounds n. This assumes a τ-completeness level of 99% and $p_{\mathsf{noise}} = 5\%$. The bottom curve gives the soundness level. (Note that DB3 is not sound and that FO follows another TF-resistance philosophy.)

4 Conclusion

We provided the provably secure symmetric protocols DB1, DB2, and DB3 which require fewer rounds than the only two existing, provably secure protocols, SKI and FO. Prior to this, we have

revised the formal model for distance-bounding protocols in a way which is closer to (the state of the art of) interactive proofs. We also studied optimality of all provably secure DB protocols, existing and advanced herein. Some open challenges remain: 1. identify an optimal and *sound* protocol for $\mathsf{num}_c = \mathsf{num}_r = 2$; 2. study the optimality of soundness; 3. implement these protocols.

References

1. G. Avoine, C. Lauradoux, B. Martin. How Secret-Sharing can Defeat Terrorist Fraud. In *ACM Conference on Wireless Network Security WISEC'11*, Hamburg, Germany, pp. 145–156, ACM, 2011.
2. G. Avoine, A. Tchamkerten. An Efficient Distance Bounding RFID Authentication Protocol: Balancing False-Acceptance Rate and Memory Requirement. In *Information Security ISC'09*, Pisa, Italy, Lecture Notes in Computer Science 5735, pp. 250–261, Springer-Verlag, 2009.
3. A. Bay, I. Boureanu, A. Mitrokotsa, I. Spulber, S. Vaudenay. The Bussard-Bagga and Other Distance-Bounding Protocols under Attacks. In *INSCRYPT'12*, Beijing, China, Lecture Notes in Computer Science 7763, pp. 371–391, Springer-Verlag, 2012.
4. I. Boureanu, A. Mitrokotsa, S. Vaudenay. On the Pseudorandom Function Assumption in (Secure) Distance-Bounding Protocols - PRF-ness alone Does Not Stop the Frauds! In *LATINCRYPT'12*, Santiago, Chile, Lecture Notes in Computer Science 7533, pp. 100–120, Springer-Verlag, 2012.
5. I. Boureanu, A. Mitrokotsa, S. Vaudenay. Secure & Lightweight Distance-Bounding. In *Lightweight Cryptography for Security and Privacy LightSec'13*, Gebze, Turkey, Lecture Notes in Computer Science 8162, pp. 97–113, Springer-Verlag, 2013.
6. I. Boureanu, A. Mitrokotsa, S. Vaudenay. Practical & Provably Secure Distance-Bounding. Eprint technical report, 2013. http://eprint.iacr.org/2013/465.pdf
7. I. Boureanu, A. Mitrokotsa, S. Vaudenay. Towards Secure Distance Bounding. In *Fast Software Encryption'13*, Singapore, Lecture Notes in Computer Science 8424, pp. 55–67, Springer-Verlag, 2013.
8. I. Boureanu, A. Mitrokotsa, S. Vaudenay. Practical & Provably Secure Distance-Bounding. To appear in the proceedings of ISC'13.
9. S. Brands, D. Chaum. Distance-Bounding Protocols (Extended Abstract). In *Advances in Cryptology EUROCRYPT'93*, Lofthus, Norway, Lecture Notes in Computer Science 765, pp. 344–359, Springer-Verlag, 1994.
10. H. Chernoff. A Measure of Asymptotic Efficiency for Tests of a Hypothesis Based on the sum of Observations. *Annals of Mathematical Statistics*, vol. 23 (4), pp. 493-507, 1952.
11. C.J. F. Cremers, K.B. Rasmussen, B. Schmidt, S. Čapkun. Distance Hijacking Attacks on Distance Bounding Protocols. In *IEEE Symposium on Security and Privacy S&P'12*, San Francisco, California, USA, pp. 113–127, IEEE Computer Society, 2012.
12. Y. Desmedt. Major Security Problems with the "Unforgeable" (Feige-)Fiat-Shamir Proofs of Identity and How to Overcome Them. In *Congress on Computer and Communication Security and Protection Securicom'88*, Paris, France, pp. 147–159, SEDEP Paris France, 1988.
13. M. Fischlin, C. Onete. Terrorism in Distance Bounding: Modelling Terrorist-Fraud Resistance. In *Applied Cryptography and Network Security ACNS'13*, Banff AB, Canada, Lecture Notes in Computer Science 7954, pp. 414–431, Springer-Verlag, 2013.
14. G.P. Hancke. Distance Bounding for RFID: Effectiveness of Terrorist Fraud. In *Conference on RFID-Technologies and Applications RFID-TA'12*, Nice, France, pp. 91–96, IEEE, 2012.
15. W. Hoeffding. Probability Inequalities for Sums of Bounded Random Variables. *Journal of the American Statistical Association*, vol. 58, pp. 13–30, 1963.
16. C.H. Kim, G. Avoine, F. Koeune, F.-X. Standaert, O. Pereira. The Swiss-Knife RFID Distance Bounding Protocol. In *Information Security and Cryptology ICISC'08*, Seoul, Korea, Lecture Notes in Computer Science 5461, pp. 98–115, Springer-Verlag, 2009.
17. S. Vaudenay. On Modeling Terrorist Frauds. In *Provable Security ProvSec'13*, Melaka, Malaysia, Lecture Notes in Computer Science 8209, pp. 1–20, Springer-Verlag, 2013.
18. T.-Y. Youn, D. Hong. Authenticated Distance Bounding Protocol with Improved FAR: Beyond the Minimal Bound of FAR. *IEICE Transactions on Communications*, vol. E97-B (5), pp. 930–935, 2014.

A Security Models to be Revisited.

The model in [6,8] factors all the previously enumerated common frauds into three possible threats:

– *Distance fraud*. This is the classical notion, but concurrent runs with many participants is additionally considered. I.e., it includes other possible provers (with other secrets) and verifiers. Consequently, this generalized distance fraud also includes distance hijacking.
– *Man-in-the-middle*. This formalization considers an adversary working in two phases. During a *learning phase*, this adversary can interact with many honest provers and verifiers. Then, the *attack phase* contains a far away honest prover of given ID and possibly many other honest provers and other verifiers. The goal of the adversary is to make the verifier accept the proof with ID. Clearly, this generalizes mafia fraud (capturing relay attacks) and includes impersonation fraud.
– *Collusion fraud*. This formalization considers a far-away prover holding x who helps an adversary to make the verifier accept. This might be in the presence of many other honest participants. However, there should be no man-in-the-middle attack stemming from this malicious prover. I.e., one should not extract from this prover any advantage to (later) run a man-in-the-middle attack.

In Vaudenay [17], the last threat model is replaced by a notion coming from interactive proofs:

– *Soundness*. For all experiment with a verifier $\mathcal{V}$, there exists an extractor such that the following holds: if this extractor is given as input several views of all participants which were close to $\mathcal{V}$ in several executions and which made him accept therein, then this extractor reconstructs the secret x. This was further shown to generalize collusion-fraud resistance [17].

In Section 2, we refine these models in a more natural way, including at its basis a stronger, inner sense of interactive proofs. Indeed, distance-bounding (DB) should ideally behave like a traditional interactive proof system as it really is a *proof of proximity*. In this sense, it must satisfy: 1. **completeness** (i.e., an honest prover close to the verifier will certainly pass the protocol); 2. **soundness** (i.e., if the verifier accepts the protocol, then we could extract from close-by participants the information to define a successful prover); 3. **security** (i.e., no participant shall be able to extract some information from the honest prover to make the verifier accept). These properties are similar to what is required in *identification protocols*. They differ in that in DB we face the introduction of the notion of proximity.

More precisely, in the above approach, distance fraud (as in Def. 6) does not capture distance hijacking anymore, distance hijacking being now captured by soundness. This makes proofs simpler. To this end, we also formalize in Def. 8 security without a learning phase, and we extend in Def. 10 the definition of soundness in such a way that the extraction of the secret is no longer necessary.

B Circular-Keying PRF

We revise the notion of circular-keying in pseudorandom functions introduced in [6,8].

Definition 15 (Circular PRF). *We consider come parameters s, n_1, n_2, and q. Given $\tilde{x} \in \{0,1\}^s$, a function L from $\{0,1\}^s$ to $\mathsf{GF}(q)^{n_1}$, and a function F from $\{0,1\}^*$ to $\mathsf{GF}(q)^{n_2}$, we define an oracle $O_{\tilde{x},F}$ by $O_{\tilde{x},F}(y,L,A,B) = A \cdot L(\tilde{x}) + B \cdot F(y)$, using the dot product over $\mathsf{GF}(q)$. We assume that L is taken from a set of functions with polynomially bounded representation. Let $(f_x)_{x\in\{0,1\}^s}$ be a family of functions from $\{0,1\}^*$ to $\{0,1\}^{n_2}$. We say that the family f is a (ε,T)-circular-PRF if for any distinguisher limited to a complexity T, the advantage for distinguishing O_{x,f_x}, $x \in_U \{0,1\}^s$, from $O_{\tilde{x},F}$, $\tilde{x} \in_U \{0,1\}^s$, where F is uniformly distributed, is bounded by ε. We require two conditions on the list of queries:*

– *for any pair of queries (y,L,A,B) and (y',L',A',B'), if $y = y'$, then $L = L'$;*

– *for any* $y \in \{0,1\}^*$*, if* (y, L, A_i, B_i)*,* $i = 1, \ldots, \ell$ *is the list of queries using this value* y*, then*

$$\forall \lambda_1, \ldots, \lambda_\ell \in \mathsf{GF}(q) \quad \sum_{i=1}^{n} \lambda_i B_i = 0 \implies \sum_{i=1}^{n} \lambda_i A_i = 0 \tag{1}$$

over the $\mathsf{GF}(q)$*-vector space* $\mathsf{GF}(q)^{n_2}$ *and* $\mathsf{GF}(q)^{n_1}$.

Link between notions of circular keying. This definition extends the one from [6,8] in the following sense: 1. the function L (the leak of x) is arbitrary instead of being linear; 2. this arbitrary function L now requires the first condition i.e., the same F-input implies the same leak function L. In [6,8], it was shown that the natural construction $f_x(y) = H(x,y)$ is circular-PRF in the random oracle model, with the definition from [6,8]. We can easily see that the same proof holds with Def. 15. It would be interesting to make other constructions without random oracles.

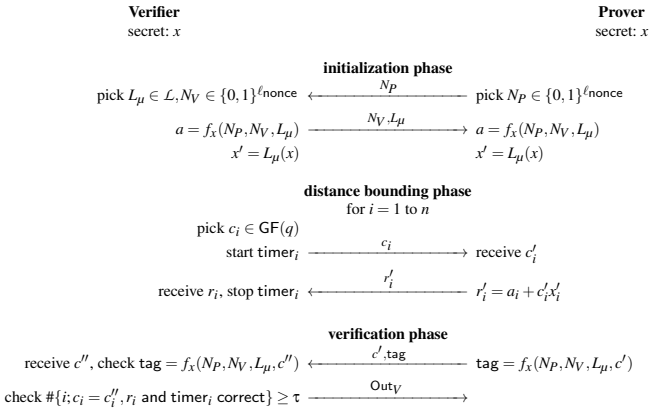

Fig. 4. The DB1 Distance-Bounding Protocol

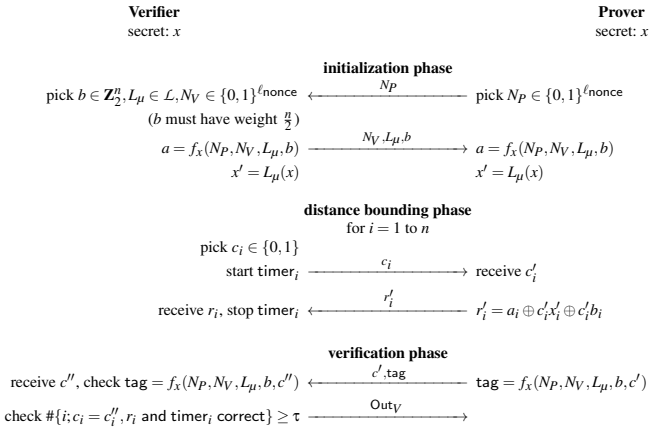

Fig. 5. The DB2 Distance-Bounding Protocol

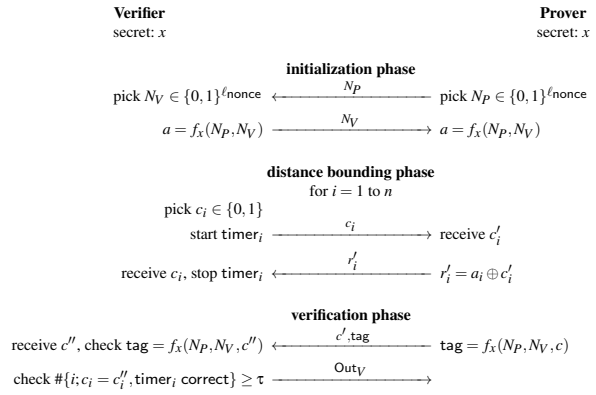

Fig. 6. The DB3 Distance-Bounding Protocol with $q = 2$.