The Fiat Shamir Transformation and its Security Problems

Shymaa M. Arafat
shar.academic@gmail.com, shymaa.arafat@gmail.com

An article submitted in partial fulfilment of the Trailerblazer Tier of the ZK MOOC course 2023

Whenever you study Zero Knowledge you are always told after explaining any interactive protocol that you can turn it into *non interactive* through Fiat Shamir transformation [1], and that the resulting protocol inherits the main security properties (in the random oracle model) of the interactive version; Fig.1. However, that's not the whole story [2]; the security risks of Fiat-Shamir transformation and possible cautions & mitigations have been a subject of research since 2013 [3] and uptill now in 2023 [4,5]. In this article I will try to share what I've learned about Fiat-Shamir transmission from the ZK MOOC course and a few other resources[6].

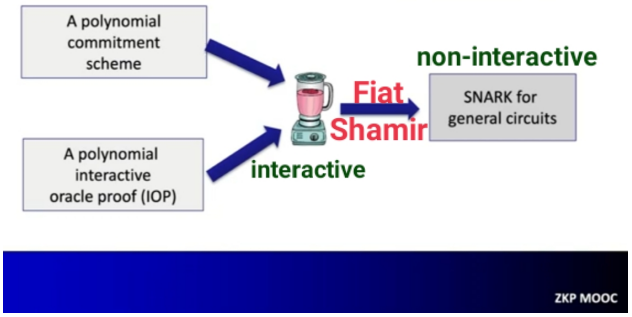

Fig.1 : the SNARK main components, and the Fiat-Shamir transformation

The Fiat Shamir Transformation

We know that the abbreviation "**SNARK**" stands for *Succinct **Non Interactive** Argument of Knowledge*, inspite of the fact that the whole idea of most protocols is a Verifier challenging a Prover with some random challenges. The Fiat-Shamir transformation tell us that if all verifier randomness are public (what is called **public coin** protocol), then we can render such protocol interactive through the use of a strong cryptographic hash function that can be modeled as a random oracle (in practice usually SHA256):

H: M ----> R

The random coin that is supposed to be generated by the verifier is replaced by the hash value of the prover's first message; see Fig.2 from lecture 2 of the course [1] where the prover's first message here is the commitments of the two functions we are checking their equality. The idea behind such replacement, and also behind the widely used term ***"random oracle model"***, comes from the randomness properties of cryptographic hash functions; it simply means that we can treat those functions as *a black box that takes an input and produces a purely random output* that is computational infeasible for an adversary to retrieve back the original input or another input that leads to the same output. Consequently, protocols that depend solely on hash functions (including those that apply Fiat-Shamir transform) are considered post-quantum since

their security depends on this "black box" that we can replace with a stronger one (say double the security bit strength) when our world is opposed to quantum computing threats.

The Fiat-Shamir transform is widely heavily used everywhere; it's said in [2] that almost all SNARKs (with the exception of Groth16 and its variations), and also different FRI deployments in Blockchains use the Fiat-Shamir transformation to become interactive. For its importance that approaches necessity, we have to know more about its security risks.

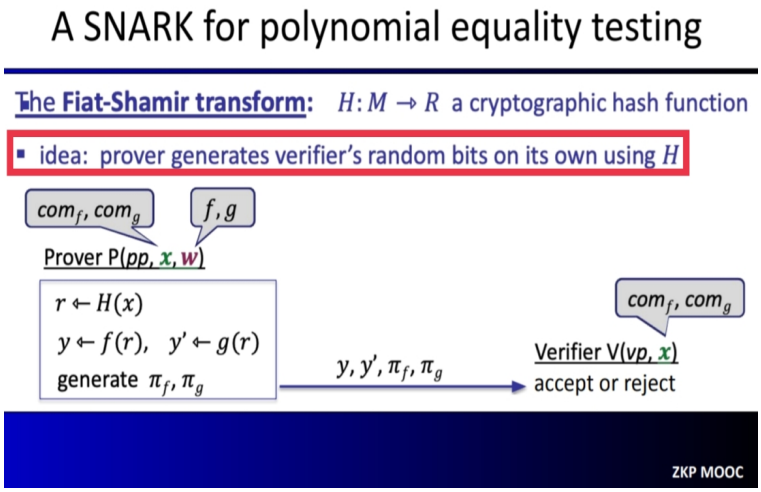

Fig.2: turning a polynomial equality testing interactive using Fiat-Shamir

Doubts & Theoretical Contradictions- an ongoing research

The original theorem tell us that this works if the hash function H is modeled as a random oracle, and if the degree of the committed polynomial/message "d" is much smaller than the field size "p" such that d/p is so small to be considered negligible [1]. However, the doubts in [3] came from an older 2003 paper which proved that "*there exists a computationally sound argument on which the Fiat-Shamir heuristic is never sound, when instantiated with any actual efficient hash function*", and was discussed thoroughly in [6]. The authors in [4] seems to follow the same analogy of [3] to show the black-box impossibility of a quantum Fiat-Shamir transform, I'm really not sure if this gets us back to a kind of contradiction about being post quantum when depending on Fiat-Shamir.

A paper in 2018 [7] showed that the Fiat-Shamir transform can be soundly applied (in the plain model) to a richer class of protocols when applied Round-by-Round; infact [6] was followed in the next day [8] by the same lecturer explaining these positive results (being one of the authors) and it was even more clear for me in the ZK MOOC maybe more up-to-date version[2].

More theoritical details & results were explained in [9], from which catched my interest that the same idea of randomness in the random oracle model leads to 1/Q security loss for Q queries; Fig.3; which I think is closely connected to the attack introduced in[10] that causes security loss $1/2(Q^{\mathcal{M}}/\mathcal{M}^{\mathcal{M}+t})$ where Q is the number of queries to the random oracle, and t is the

folding degree (notice now the degree of folding becomes also a concern) of a (2$\mathcal{M}$+1)-move[1] protocol.

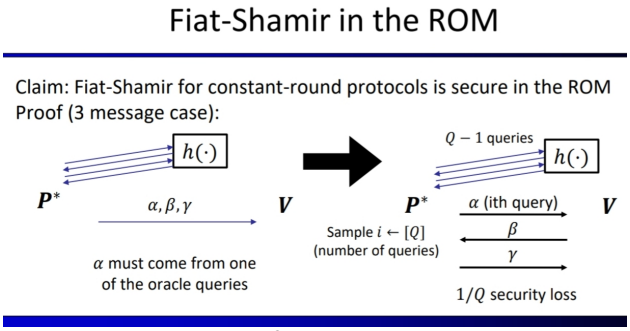

Fig. 3 security loss for Q queries to Fiat-Shamir

An extra doubt crossed by mind, though maybe not related, about FRI and the d/p condition in the above theorem doesn't exactly encourage the possible use of small field sizes as an advantage in this case. The lecturer [9] also mentioned the security risks when applying recursion on a Fiat-Shamir transformed protocol (recursive SNARKs) and that it becomes like running the hash function on its own description.

Still there are plenty of newer research about Fiat-Shamir security in lattices [11], A lot more others are there [12,13] that I only read their abstracts, and even if I could have access to the full papers they are much heavier cryptographic dose for me and beyond my scope of interest; so let me just point out their links for the interested reader and then concentrate on telling you what I have learned...

Adaptive Security

For the resulting hash to be purely random and protect from an adversary capable of choosing the public input to the verifier, vp in Fig.2, then vp must also be fed as an input (with the prover's first message) to the random oracle hash function; **r <--H(x,vp)**. This comes from [12] where the author's proved that "***adaptive soundness*** (and non adaptive zero knowledge[2]) is achieved ***so long as the challenge is obtained by hashing both the prover's first round and the instance being proven***"

The Grinding Attack:

I guess the core concept of the grinding attack is how catastrophic the situation can get if the prover ***can try different messages offline*** to find a match, instead of only able to try online,

---

1 From [15] "we use the convention that n-round PCIP has 2n+1 moves", so $\mathcal{M}$ is the number of rounds

[2] I think real zero knowledge meaning not revealing info about data being proved is not the main target of most Blockchain deployments that aims at scaling and use only the term ZK maybe because it was publicly promoted as a title, or maybe because it's the title of the science that developed those solutions.

especially with nowadays available mining devices that is mastered to try all possible hashes. Meaning that in the interactive version of the protocol the malicious prover will have to confront the verifier with every false attempt; a situation that can be handled by just limiting the number of failed proofs or keeping the history of previous attempts in some function, or reputation, or even sometimes the needed number of attempts is infeasible to happen (ex. 1 million Ethereum blocks takes 3yrs, Bitcoin blocks maybe 20yrs).

You may tell your self that this render the problem to the hash function resistance against brute force attacks, which applies to all applications that deploy cryptographic hash functions applications not just Fiat-Shamir transformation, and protocol designers increase security by simply increasing the number of bits of the hash function to increase the effort needed to find an input that leads to the hash which could be the original value or a hash collision (another input that leads to the same hash). The problem here is that compared to a Birthday attack[3] with the same effort (hash computational power), ***a Fiat-Shamir grinding attack has a probability of 2^{10} to succeed when the Birthday attack has a probability of 2^{-20} to succeed; that's 1/1000 compared to 1/million***, a dramatic security loss. In general, for an (n-bit hash strength) a Fiat-Shamir adversary possesing 2^b hash power a grinding attack succeeds with probability 2^{b-n} (even with adaptive soundness that doesn't hardens it), while a birthday attack success probability is $2^{2(n-b)}$; currently it's considered safe enough to use 128bit security strength (256 bit hash functions) for Fiat-Shamir transformation. I couldn't find the original paper, but more explainatory details can be found in [14].

The risk of grinding attacks on many rounds protocols

The security loss when applying Fiat-Shamir to many round protocols could be catastrophic since the attacker can silently grind on each round separately[4] ; ie, if the computational hash effort to brute force round 1,2,...,i is respectively 2^{m1},2^{m2},...2^{mi} then rendering the protocol non-interactive using Fiat-Shamir transform gives an adversary to try offline (grind silenty) on the each round separately, the 1^{st} then the 2^{nd} then the 3^{rd} and so on leading to a total effort $2^{m1}+2^{m2}+...+2^{mi}$ instead of $2^{m1+m2+....+mi}$.

Round-by-Round Soundness

The thoerem in [7] states that "The Fiat-Shamir transform is provably (adaptively) sound when applied to round-by-round sound interactive proofs using a hash family satisfying a restricted form of correlation intractability"; it simply means that an adversary cannot grind on each round separately [14]. Examples of protocols that have been proven to be round-by-round sound are the Sumcheck protocol and Bullet proofs, while a post-quantum protocol (supposed to be

---

[3] A birthday attack is an interesting famous result that in a random population sample of 25 person's there's a probability more than 1/2 (>50%) to find two that have the same birthday; the result led to the fact that a brute force attack on a cryptographically secure hash function of n bits could succeed after $2^{n/2}$ with probability≥50% and thus hash functions with 256bits are said to provide 128bit security (likely 2^{128} trials are enough to a break them)

[4] Think of it as breaking each stick separately easier than breaking a bunch of them at once

deployed for maximum security) like FRI that is heavily used as non-interactive using Fiat-Shamir has not proven to be round-by-round sound uptill now [2]; you can read more about round-by-round soundness in [15].

The Bottom-line is ...

We know it's pretty useful and sometimes crucial to turn an interactive argument of Knowledge non-interactive using the Fiat-Shamir transform, however you have to watch out for security losses ***even in the random oracle mode***l. So, when designing protocols, and when choosing which protocol to deploy, ***you have to prove the round-by-round soundness of the protocol before applying the Fiat-Shamir transform; and also do not forget to feed the verifier public input to the used hash to achieve adaptive soundness***.

Is there an alternative?

This Berkeley MOOC course is my first real attempt about Zero Knowledge; ie, I not claim an expert here, but according to [2] all practical implentations except Groth16 use Fiat-Shamir; however, according to [9] theorists have never stopped searching for alternatives. There are efforts on falsifiable assumptions, efforts based on using Encryption in place of hashing (with adding a constant like 1 to provide circular security since Encrypt (f(x)) should be completely different from Encrypt (f(X)+1)), and a lot much more; the term ***Correlation Intractability (IC)*** is excessively used to describe the security strength of the used hash. From this huge pile, I bring this newest one [16],just released May 2023, the paper suggests a new generic transform to replace Fiat-Shamir; "A Generic Transform from Multi-round Interactive Proof to NIZK". In their transform, as opposed to the random oracle model [9,17] in Fiat-Shamir, the zero-knowledge property is in the standard model, and the adaptive soundness is in the non-programmable random oracle model (NPROM) [18].

References

[1] lecture 2 of ZK Berkeley MOOC; https://youtu.be/Sv99taTJJmM
[2] lecture 8 of ZK Berkeley MOOC; https://youtu.be/A3edAQDPnDY
[3] https://www.researchgate.net/publication/262398051_Why_Fiat-Shamir_for_Proofs_Lacks_a_Proof
[4] https://www.researchgate.net/publication/359758217_Fiat-Shamir_for_Proofs_Lacks_a_Proof_Even_in_the_Presence_of_Shared_Entanglement
[5] https://www.researchgate.net/publication/366468861_Fiat-Shamir_Transformation_of_Multi-round_Interactive_Proofs
[6] Ron Rothblum, "The Fiat Shamir Transform", BIU 2019; https://youtu.be/9cagVtYstyY
[7] R.Canetti, Y. Chen, J. Holmgren, A. Lombardi, G. Rothblum, Ron Rothblum, "Fiat-Shamir From Simpler Assumptions", Oct 2018; https://eprint.iacr.org/2018/1004.pdf
[8] Ron Rothblum, "Fiat Shamir from Theory to Practice", BIU 2019; https://youtu.be/yWq-_hiTgJw

[9] lecture 11 of ZK Berkeley MOOC; https://youtu.be/CIGnBb8B0rQ
[10] "Fiat-Shamir Transformation of Multi-round Interactive Proofs"; https://link.springer.com/chapter/10.1007/978-3-031-22318-1_5
[11] https://www.researchgate.net/publication/357699480_Full_Leakage_resilience_of_Fiat-Shamir_signatures_over_lattices
[12] On Adaptive Security of Delayed-Input Sigma Protocols and Fiat-Shamir NIZKs; https://eprint.iacr.org/2020/831
[13] https://www.researchgate.net/publication/363267504_What_Makes_Fiat-Shamir_zkSNARKs_Updatable_SRS_Simulation_Extractable
[14] https://crypto.stackexchange.com/questions/97735/grinding-in-the-fiat-shamir-heuristic
[15] https://www.semanticscholar.org/paper/On-Round-By-Round-Soundness-and-State-Restoration-Holmgren/a33836943673c5f62b4273863745080f12321878
[16] https://www.researchgate.net/publication/370452121_A_Generic_Transform_from_Multi-round_Interactive_Proof_to_NIZK; Full paper: https://link.springer.com/chapter/10.1007/978-3-031-31371-4_16
[17] https://crypto.stackexchange.com/questions/879/what-is-the-random-oracle-model-and-why-is-it-controversial
[18] https://crypto.stackexchange.com/questions/37247/what-is-the-non-programmable-random-oracle-model