Designing Efficient Zero-Knowledge Proofs in the Ideal Linear Commitment Model

Jonathan Bootle

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

of

University College London.

Computer Science Department

University College London

September 17, 2018

I, Jonathan Bootle, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.

Abstract

Zero-knowledge proofs are cryptographic protocols enabling a prover to demonstrate to a verifier that a public statement is true, without giving away any of the prover's secret information, or revealing why the statement is true. Since the introduction of zero-knowledge proofs, researchers have found numerous applications to other cryptosystems, such as electronic voting, group and ring signatures, verifiable computation, and cryptocurrencies, and zero-knowledge proofs have become an integral part of blockchain-based cryptocurrencies.

Thus, designing efficient zero-knowledge proofs is an important goal for cryptographers. In recent years, the design space has become extremely large, and to simplify protocol design, designers have begun to separate the process into modular steps. Information theoretic protocols are designed in idealised communication models, and then compiled into real zero-knowledge proofs using suitable cryptographic assumptions.

In this thesis, we investigate a particular model called the Ideal Linear Commitment model, which aims to characterise interactive zero-knowledge protocols where the prover and verifier use homomorphic commitment schemes. We demonstrate the model's power by exhibiting efficient protocols for various useful tasks including NP-Complete problems and some more specialised problems. We demonstrate the model's versatility by explaining how to convert the idealised protocols into real protocols based on two completely different cryptographic assumptions; the discrete logarithm assumption, and the existence of collision-resistant hash-functions.

All in all, we show that the Ideal Linear Commitment model is a useful and effective abstraction for producing zero-knowledge protocols. Furthermore, by identifying the limitations of the model and finding protocols which work outside these constraints, we display special techniques which result in more efficient protocols than ever before.

The results are novel and highly efficient protocols, all of which improve the theoretical state-of-the-art in zero-knowledge research.

Impact Statement

This thesis demonstrates how real cryptographic protocols for a variety of tasks can be designed using a special communication model. It shows that one can separate the algebraic machinery used to design the protocols from the cryptographic assumptions and commitment schemes used to prove that the protocols are secure, and still exhibit highly efficient protocols which improve asymptotically on the prior state-of-the-art. This is likely to benefit the discipline considerably as it lowers the barriers to understanding and producing secure protocols of this type. Cryptographic assumptions can be quite specialised, and it can be difficult to understand the mathematics behind them. Having a framework within which one can prove idealised protocols secure and knowing that the result can be made into a real protocol drastically simplifies the task of protocol designers. Several subtle changes to the Ideal Linear Commitment model were also introduced to make it more realistic and effective.

Furthermore, by showing that certain proofs all fit into a framework, it becomes easier to understand their limitations. Linear algebra is an important part of security proofs in the Ideal Linear Commitment model. Therefore, in future, zero-knowledge protocols can be analysed through the lens of linear algebra. Linear algebra is very well studied, and powerful techniques from this discipline may lead to strong results about zero-knowledge protocols, such as lower bounds on the communication complexity of certain types of interactive zero-knowledge protocol.

The protocols presented in this work led to the creation of the zero-knowledge argument Bulletproofs [1]. Bulletproofs has been implemented by cryptocurrencies including Monero and PIVX. PIVX plans to bring Bulletproofs implementations into common use later in 2018. Following a first successful code audit, Bulletproofs is also set to enter widespread use on Monero's blockchain later in 2018, subject to further successful audits. As a result, the author's work will soon have a sizeable impact on the efficiency of payment systems in the real world, which at the time of writing, amount to a market capitalisation value of roughly two trillion dollars and a daily trade volume of roughly thirty four million dollars [1]. The addition of Bulletproofs will lead to much smaller amounts of proof data being stored on

the blockchains for these cryptocurrencies, which means better performance and functionality. This may help to increase the adoption of cryptocurrencies.

Zero-knowledge proofs are becoming better and better known, not only among research scientists, but increasingly among companies and technology enthusiasts. There are also ongoing standardisation efforts. The implementation of cryptographic protocols is a notoriously difficult task even for experts, and errors can have disastrous consequences. Making protocols easier to design and understand will be of great benefit to interested, non-expert parties who might try to use protocols in the future, promote their usage among wider user communities, or implement them in software or hardware.

[1] Data taken from `coinmarketcap.com`, on the 17th of September 2018.

Acknowledgements

During the four years of study which led to this thesis, I have been incredibly lucky to have the chance to interact with not only my colleagues from UCL, but researchers from NTT Secure Platform Laboratory, IBM Research Zurich, and Microsoft Research Redmond. Our discussions always served to open my mind and help me consider things in new ways. Thanks to my supervisor Jens Groth who was a great mentor and provided me with a better model for how to be an academic researcher than I could ever have hoped for. Thanks to Andrea Cerulli who was always willing to help and discuss things, probably at the expense of his own work. Thanks to Vasilios Mavroudis and David Kohan Marzgao for providing wonderful distractions in the form of exciting mathematical puzzles. Thanks to my mother, whose motivation when I was a child made all of this possible.

Contents

List of Figures

List of Tables

Chapter 1

Conclusion

In this thesis, we introduced the ILC model; an idealised communication model which can be used to construct information theoretic zero-knowledge proofs. We aimed to demonstrate the usefulness and power of the ILC model by constructing zero-knowledge protocols for a wide-variety of applications. We wanted to give zero-knowledge protocols with better efficiency than ever seen before. We wanted to distill the ideas underlying a long seqence of works on interactive zero-knowledge proofs, so that the techniques used to design and analyse them could be viewed through a common lens, and in doing so, make it easier to design these protocols.

In order to solve these problems, we presented changes to the original ILC model to bring it closer to real protocols, and gave compilations from ILC protocols to real zero-knowledge protocols based on hash functions and error-correcting codes. The compilations separated the cryptographic and non-cryptographic parts of the design process and simplify the protocol design process. In particular, designing our ILC protocols and proving them secure was a matter of applying linear algebra and simple lemmata about polynomial identity testing. Proving that the ILC protocols could be securely compiled into real arguments was more complicated, but was done once and for all, and the compilations can be reused for many ILC protocols in the future. We presented protocols with state-of-the-art communication complexity and round complexity, and showed that the ILC model is powerful enough to reason about both general NP-Complete statements like arithmetic circuit satisfiability. We also gave ILC protocols for simpler statements such as polynomial evaluation or range proofs, in a manner that leads to highly efficient protocols. This included the framing of general relations to capture a class of zero-knowledge protocols characterised by low-degree polynomials, formalising the techniques used in such protocols, and providing a generic protocol for reasoning about such relations, which can be used to give batch-proofs for many statements at the same time.

Using only this methodology, we were able to present some protocols with state-of-the-art communication complexity. Examples include a discrete-logarithm based polynomial evaluation argument, with a better asymptotic communication complexity than observed prevously, and a discrete-logarithm based membership argument, whose asymptotic communication complexity has improved constants over previous work, and which has highly tuneable parameters. These are of practical significance as they can be used as part of membership and non-membership arguments both in the designs of other primitives, like group and ring signatures, and in applications such as preventing double-spending in cryptocurrencies. Since ILC protocols can also be compiled based on hash functions and error-correcting codes, we also obtain some completely new arguments for polynomial evaluation and membership based on the existence of collision resistant hash functions.

The path that led to this work was trying to find the techniques common to all interactive zero-knowledge protocols based on the discrete logarithm assumption, and other homomorphic commitment schemes over fields. Through the ILC model, this work shows that indeed, a great many discrete logarithm arguments follow the same basic design paradigms, and that surprisingly, the same style of protocol and design techniques extend beyond the discrete logarithm setting to another commitment scheme which is not homomorphic!

We also presented some extra techniques which fall outside the ILC model, namely, a recursive argument to show that committed values have a particular scalar product, and a field extension technique which boosts the soundness of ILC protocols over small fields. This is at once a strength and a weakness of using idealised communication models. Protocols inside such models are highly constrained, which makes them easier to design and reason about, but may also limit their performance and utility. The fact that the most efficient protocol in this thesis, the logarithmic-communication argument for arithmetic circuit satisfiability, does not lie within the main model of communication, is a limitation. However, once a suitable model has been identified, one can also try to design useful protocols by attempting to create protocols outside the model.

There are other zero-knowledge protocols [**?**], some based on lattices, and some based on the Strong RSA assumption, which seem to work on the same basis as ILC protocols. That is, the prover commits to certain vectors, and the verifier picks a random challenge, and uses structured linear combinations of the committed vectors in a number of verification equations. Unlike in the ILC model, in which all elements belong to a field and the notion of size is not important, these settings require careful consideration of the size of committed elements to ensure zero-knowledge, and often for soundness too. The model falls short of capturing these protocols. Improving the model to take this into account, in particular for lattice-based protocols which may enjoy post-quantum security guarantees, is an attractive target for future research.

Another avenue that was not investigated is restricting the verifier's ILC queries. In all of the ILC protocols presented in this thesis, the coefficients of the verifier's linear queries are given by a linearly-independent set of polynomials evaluated at uniformly random challenges chosen by the verifier. The queries have a carefully chosen algebraic structure. For every protocol that we give, the query matrix appears to be a form of strongly universal hash function. The compilation from ILC protocols to discrete-logarithm based protocols requires restrictions on the rank and dimension of the matrix, and that a related system of linear equations can be solved. These conditions are treated in an ad-hoc manner outside of the proofs that the protocols are secure in the idealised model. There is still a gap between the model and the compiled protocols, and the communication model can be refined further. One could hope that such strong algebraic restrictions lead to interesting results, such as lower bounds on the communication complexity of ILC protocols, as linear algebra is an old discipline with many results that one could hope to apply to the structure of the query matrices.

Bibliography

[1] Benedikt Bünz, Jonathan Bootle, Dan Boneh, Andrew Poelstra, Pieter Wuille, and Greg Maxwell. Bulletproofs: Short proofs for confidential transactions and more. *IACR Cryptology ePrint Archive*, 2017:1066, 2017.