## Zero-Knowledge Arguments for Subverted RSA Groups

Dimitris Kolonelos<br>IMDEA Software Institute<br>\& Universidad Politecnica de Madrid

Mary Maller
Ethereum Foundation
\& PQShield

## Mikhail Volkhov

The University of Edinburgh


PKC 2023, Atlanta
10 May 2023

# Hidden Order Groups <br> (aka Groups of Unknown Order) 

Group $\mathbb{G}$ where computing $\operatorname{ord}(\mathbb{G})$ is hard

## Hidden Order Groups <br> (aka Groups of Unknown Order)

Group $\mathbb{G}$ where computing $\operatorname{ord}(\mathbb{G})$ is hard

Group exponentiation: $g^{x} \underbrace{\bmod \operatorname{ord}(g)}=g^{x}$
unknown

## Hidden Order Groups <br> (aka Groups of Unknown Order)

Group $\mathbb{G}$ where computing $\operatorname{ord}(\mathbb{G})$ is hard

Group exponentiation: $g^{x} \underbrace{\bmod \operatorname{ord}(g)}_{\text {unknown }}=g^{x}$

Group exponentiations are (computationally) over $\mathbb{Z}$

## Hidden Order Groups - Applications

Group exponentiations are (computationally) over $\mathbb{Z}$
$\star$ Accumulators [BdM94, BP97, CL02, LLX07, L12, BBF19] \& Set Membership (zk-)proofs
[CLO2, BCFGK19, CDHKKO22]
$\star$ (zk-)Range proofs [B00, L03, G05, CPP17, CKLR21]
$\star$ Vector Commitments [CF13, LM19, BBF19, CFGKN20, TXN20, CFKS22]

* Polynomial commitments \& SNARKs [BFS20, BHRRS21, AGLMS23]
$\star$ Verifiable delay functions [BBBF18, W19, P19] \& time-lock puzzles [RSW96]
$\star$ Additively Homomorphic Encryption [P99, CL15]


## Hidden Order Groups - Applications

Group exponentiations are (computationally) over $\mathbb{Z}$

* Accumulators [BdM94, BP97, CL02, LLX07, L12, BBF19] \& Set Membership (zk-)proofs
[CLO2, BCFGK19, CDHKKO22]
$\star$ (zk-)Range proofs [B00, L03, G05, CPP17, CKLR21]
$\star$ Vector Commitments [CF13, LM19, BBF19, CFGKN20, TXN20, CFKS22]
* Polynomial commitments \& SNARKs [BFS20, BHRRS21, AGLMS23]
* Verifiable delay functions [BBBF18, W19, P19] \& time-lock puzzles [RSW96]
$\star$ Additively Homomorphic Encryption [P99, CL15]


## Hidden Order Groups - Applications

Group exponentiations are (computationally) over $\mathbb{Z}$
$\star$ Accumulators
\& Set Membership (zk-)proofs
[CLO2, BCFGK19, CDHKKO22]

* (zk-)Range proofs [B00, L03, G05, CPP17, CKLR21]
$\star$ Vector Commitments [CF13, LM19, BBF19, CFGKN20, TXN20, CFKS22] next talk!
* Polynomial commitments \& SNARKs [BFS20, BHRRS21, AGLMS23]
$\star$ Verifiable delay functions
\& time-lock puzzles
* Additively Homomorphic Encryption


## Hidden Order Groups - Instantiations

$* \mathbb{Z}_{N}^{*}: N=p \cdot q$ and $p, q$ 'safe' primes
$\& g^{x} \bmod N=g^{x \bmod \phi(N)}$
$\operatorname{ord}(\mathbb{G})=\phi(N)=(p-1)(q-1)$

* Computing $\operatorname{ord}(\mathbb{G}) \leftrightarrow$ Factoring ©
* Structured setup
- Class Groups of imaginary quadratic order
- Complicated Algebra... (see
[BuchamannHamdy01, Straka19])
* Computing $\operatorname{ord}(\mathbb{G}) \leftrightarrow$ Less cryptanalysis effort 6
* Uniformly random setup ©


## Hidden Order Groups - Instantiations

## RSA Groups [RSA78]

## Class Groups

\& $\mathbb{Z}_{N}^{*}: N=p \cdot q$ and $p, q$ 'safe' primes
$\otimes g^{x} \bmod N=g^{x} \bmod \phi(N)$
$\operatorname{ord}(\mathbb{G})=\phi(N)=(p-1)(q-1)$

- Computing ord $(\mathbb{G}) \leftrightarrow$ Factoring $\odot$
* Structured setup


## Subverted RSA Groups

* Natural setting when the setup is adversarial
$\star$ E.g. $N=p q$ but not safe primes, $N=p_{1} p_{2} \ldots p_{k}, N=p_{1}^{k_{1}} p_{2}^{k_{2}}$
* Class Groups of imaginary quadratic order
- Complicated Algebra... (see
[BuchamannHamdy01, Straka19])
* Computing $\operatorname{ord}(\mathbb{G}) \leftrightarrow$ Less cryptanalysis effort ©
* Uniformly random setup ©


## Hidden Order Groups - Instantiations

## RSA Groups

## Class Groups

[BW88]
\& $\mathbb{Z}_{N}^{*}: N=p \cdot q$ and $p, q$ 'safe' primes
$\otimes g^{x} \bmod N=g^{x \bmod \phi(N)}$
$\operatorname{ord}(\mathbb{G})=\phi(N)=(p-1)(q-1)$

- Computing ord $(\mathbb{G}) \leftrightarrow$ Factoring $\odot$
* Structured setup
* Class Groups of imaginary quadratic order
- Complicated Algebra... (see
[BuchamannHamdy01, Straka19])
* Computing $\operatorname{ord}(\mathbb{G}) \leftrightarrow$ Less cryptanalysis effort ©
* Uniformly random setup


## Subverted RSA Groups

* Natural setting when the setup is adversarial
$\star$ E.g. $N=p q$ but not safe primes, $N=p_{1} p_{2} \ldots p_{k^{\prime}} N=p_{1}^{k_{1}} p_{2}^{k_{2}}$


## Paillier Encryption \& Key Subversion

```
Fully additive variant [DJ10, L17]
RSA modulus
\(\star \operatorname{KeyGen}\left(1^{\lambda}\right): N=p \cdot q, h \leftarrow g^{N} \in \mathbb{Z}_{N^{\prime}}^{2} \mathrm{pk}=(N, h), \mathrm{sk}=(p, q)\)
*Enc(pk, \(m\) ) : ct \(\leftarrow(N+1)^{m} h^{r} \bmod N^{2}\)
\(* \operatorname{Dec}(\mathrm{sk}, \mathrm{ct}):\left(\left[\mathrm{ct} \cdot \mathrm{ct}^{\left[N^{-1} \bmod \phi(N)\right] \cdot N} \bmod N^{2}\right]-1\right) / N\)
```


## Paillier Encryption \& Key Subversion

Fully additive variant [DJ10, L17]

## RSA modulus

*KeyGen $\left(1^{\lambda}\right): N=p \cdot q, h \leftarrow g^{N} \in \mathbb{Z}_{N^{\prime}}^{2}, \mathrm{pk}=(N, h)$, sk $=(p, q)$
*Enc(pk, $m$ ) : ct $\leftarrow(N+1)^{m} h^{r} \bmod N^{2}$
$* \operatorname{Dec}(\mathrm{sk}, \mathrm{ct}):\left(\left[\mathrm{ct} \cdot \mathrm{ct}^{\left[N^{-1} \bmod \phi(N)\right] \cdot N} \bmod N^{2}\right]-1\right) / N$

Why would anyone subvert their own key?

Encrypt-(under-your-own-key)-and-prove

- MPC Ceremony for RSA modulus generation [HMRT19]
- Threshold ECDSA [CGGMP20, BMP22]
- E-Voting [DJ01]


## Zero-Knowledge Proofs [GMR89]



* (Proof/)Argument of Knowledge ((P/)AoKs): if Verify $(x, \pi)=1$ then $\mathscr{P}$ knows a s.t. $R(x, \quad)=1$
* Zero-Knowledge (ZK): VV learns nothing about
* Non-interactive (NI): $\mathscr{P}$ generates $\pi$ without any interaction with $\mathscr{V}$
* (Public/)Designated-Verifier (DV) [DFNO6, PsV06]: (every $\mathscr{V} /$ ) only one $\mathscr{V}$ holding vsk can verify $\pi$


## Zero-Knowledge Proofs [GMR89]



* (Proof/)Argument of Knowledge ((P/)AoKs): if Verify $(x, \pi)=1$ then $\mathscr{P}$ knows a s.t. $R(x, \quad)=1$
* Zero-Knowledge (ZK): VV learns nothing about
* Non-interactive (NI): $\mathscr{P}$ generates $\pi$ without any interaction with $\mathscr{V}$
* (Public/)Designated-Verifier (DV) [DFN06, PsV06]: (every $\mathscr{V} /$ ) only one $\mathscr{V}$ holding vsk can verify $\pi$


## Zero-Knowledge Proofs [GMR89]



* (Proof/)Argument of Knowledge ((P/)AoKs): if Verify $(x, \pi)=1$ then $\mathscr{P}$ knows a s.t. $R(x, \quad)=1$
* Zero-Knowledge (ZK): VV learns nothing about
* Non-interactive (NI): $\mathscr{P}$ generates $\pi$ without any interaction with $\mathscr{V}$
* (Public/)Designated-Verifier (DV) [DFNO6, PsV06]: (every $\mathscr{V} /$ ) only one $\mathscr{V}$ holding vsk can verify $\pi$


## Zero-Knowledge Proofs [GMR89]



* (Proof/)Argument of Knowledge ((P/)AoKs): if Verify $(x, \pi)=1$ then $\mathscr{P}$ knows a s.t. $R(x, \quad)=1$
* Zero-Knowledge (ZK): VV learns nothing about
* Non-interactive (NI): $\mathscr{P}$ generates $\pi$ without any interaction with $\mathscr{V}$
* (Public/)Designated-Verifier (DV) [DFN06, PsV06]: (every $\mathscr{V} /$ ) only one $\mathscr{V}$ holding vsk can verify $\pi$


## Our contributions

ネ A DV-zk-AoK of a pre-image of any (additive) homomorphism under subverted RSA groups.
e.g. $x: y=g^{x}(\bmod N), x:$ ct $=\operatorname{Paillier} . \operatorname{Enc}(x)$ for any $N$
$\star A$ DV-zk range argument for any (additive) homomorphism under subverted RSA groups. e.g. $x: \mathrm{ct}=\operatorname{Paillier} . \operatorname{Enc}(x) \wedge x \in[A, B]$ for any $N$

* Technically: A new extraction technique for proving knowledge-soundness.


## Possible Approaches

## $N$ : arbitrary chosen by $\mathscr{P}$

## $y=g^{x}(\bmod N)$

* $\Sigma$-protocols
$\rightarrow \lambda$ repetitions ( $x \lambda$ efficiency overhead) [BCK10, TW12]
* General purpose NIZK (e.g. SNARK)
$\rightarrow$ Very expensive to encode RSA operations ( $\sim 80$ million gates for ar. circuits) [OWWB20]
\& Prove correctness of $N[C M 99, \ldots]$ \& proof for non-subverted RSA groups:
$\rightarrow$ Proofs of correct moduli very expensive
* More elaborate approaches
$\rightarrow$ See the paper for discussion...


## Possible Approaches

## $N$ : arbitrany chosen by op

$y=g^{x}(\bmod N)$
$\star$-protocols
$\rightarrow \lambda$ repetitions ( $x \lambda$ efficiency overhead) [BCK10, TW12]

* General purpose NIZK (e.g. SNARK)
$\rightarrow$ Very expensive to encode RSA operations ( $\sim 80$ million gates for ar. circuits) [OWWB20]
\& Prove correctness of $N[C M 99, \ldots]$ \& proof for non-subverted RSA groups:
$\rightarrow$ Proofs of correct moduli very expensive
* More elaborate approaches
$\rightarrow$ See the paper for discussion...


## Sigma-Protocols over HoGs (and pitfails)

$N$ : arbitrary chosen by $\mathscr{P}$

$y=g^{x}(\bmod N)$



## Sigma-Protocols over HoGs (and pitfails)

$N$ : arbitrary chosen by $\mathscr{P}$

$y=g^{x}(\bmod N)$



Non-interactive via Fiat-Shamir

## Sigma-Protocols over HoGs (and pitfails)

$N$ : arbitrary chosen by $\mathscr{P}$

$y=g^{x}(\bmod N)$

(Knowledge) Soundness. Rewind to obtain 2 accepting transcripts on $a$ :


$$
g^{s}={ }_{?} a \cdot y^{c}(\bmod N)
$$

## Sigma-Protocols over HoGs (and pitfails)

$N$ : arbitrary chosen by $\mathscr{P}$

$y=g^{x}(\bmod N)$

(Knowledge) Soundness. Rewind to obtain 2 accepting transcripts on $a$ :


## Sigma-Protocols over HoGs (and pitfails)

$N$ : arbitrary chosen by $\mathscr{P}$

$y=g^{x}(\bmod N)$

(Knowledge) Soundness. Rewind to obtain 2 accepting transcripts on $a$ :


## Sigma-Protocols over HoGs (and pitfails)

$N$ : arbitrary chosen by $\mathscr{P}$

$y=g^{x}(\bmod N)$

(Knowledge) Soundness. Rewind to obtain 2 accepting transcripts on $a$ :


Cannot divide with $c-c^{\prime}$ in the exponent: $\phi\left(N^{2}\right)$ is secret
$\rightarrow$ Unable to extract unless $c \in \mathscr{C}=\{0,1\} \rightarrow$ soundess error $=1 / 2 \rightarrow$ requires $\lambda$ repetitions

## Sigma-Protocols over HoGs (and pitfails)

## $N$ : arbitrary chosen by $\mathscr{P}$

$y=g^{x}(\bmod N)$

(Knowledge) Soundness. Rewind to obtain 2 accepting transcripts on $a$ :


$$
g^{s}={ }_{?} a \cdot y^{c} \quad(\bmod N)
$$



Cannot divide with $c-c^{\prime}$ in the exponent: $\phi\left(N^{2}\right)$ is secret
$\rightarrow$ Unable to extract unless $c \in \mathscr{C}=\{0,1\} \rightarrow$ soundess error $=1 / 2 \rightarrow$ requires $\lambda$ repetitions

## Our Protocol (1): A new extraction approach



## Our Protocol (1): A new extraction approach



```
Bezout's theorem (Informal):
If \(\operatorname{gcd}\left(\delta_{2}, \ldots, \delta_{M}\right)=1\) then there exist \(k_{2}, \ldots, k_{M}\) such that \(k_{2} \delta_{2}+\ldots+k_{M} \delta_{M}=1\) (over \(\mathbb{Z}\) )
```


## Our Protocol (1): A new extraction approach

 Assume $M$ accepting transcripts on $a$ we get:$\left.\left.\begin{array}{l}\text { 1. }\left\{a, c^{(1)}, s^{(1)}\right\}: g^{s^{(1)}}=a \cdot y^{c^{(1)}} \\ \text { 2. }\left\{a, c^{(2)}, s^{(2)}\right\}: s^{s^{(2)}}=a \cdot y^{c^{(2)}}-s^{(1)} \\ \vdots \\ \text { M. }\left\{a, c^{(M)}, s^{(M)}\right\}: g^{s^{(M)}}=a \cdot y^{c^{(M)}}\end{array}\right\} \Rightarrow \begin{array}{c}g^{\beta_{2}}=y^{\delta_{2}} \\ g^{\beta_{3}}=y^{\delta_{3}} \\ \vdots \\ g^{\beta_{M}}=y^{\delta_{M}}\end{array}\right\} \Rightarrow c^{(1)}-g_{i=2}^{\sum_{i} k_{i} \beta_{i}}=y^{\sum_{i=2}^{M} k_{i} \delta_{i}}=y \quad(\bmod N)$

```
Bezout's theorem (Informal):
If \(\operatorname{gcd}\left(\delta_{2}, \ldots, \delta_{M}\right)=1\) then there exist \(k_{2}, \ldots, k_{M}\) such that \(k_{2} \delta_{2}+\ldots+k_{M} \delta_{M}=1\) (over \(\mathbb{Z}\) )
```


## Our Protocol (1): A new extraction approach

$\begin{array}{cc}\text { Assume } M \text { accepting transcripts on } a \text { we get: } & \left.\beta_{i}:=s^{(i)}-s^{(1)}\right) \\ \left.\left.\left.\begin{array}{l}\text { 1. }\left\{a, c^{(1)}, s^{(1)}\right\}: g^{s^{(1)}}=a \cdot c^{c^{(1)}} \\ \text { 2. }\left\{a, c^{(2)}, s^{(2)}\right\}: g^{(1)} \\ \vdots \\ s^{s^{(2)}}=a \cdot y^{\left(c^{(2)}\right.} \\ \text { M. }\left\{a, c^{(M)}, s^{(M)}\right\}: g^{s^{(M)}}=a \cdot y^{c^{(M)}}\end{array}\right\} \Rightarrow \begin{array}{c}g^{\delta_{2}} \\ g^{\beta_{3}}=y^{\delta_{3}} \\ \vdots \\ g^{\beta_{M}}=y^{\delta_{M}}\end{array}\right\} \Rightarrow g \sum_{i=2}^{M} k_{i} \beta_{i}=y \sum_{i=2}^{M} k_{i} \delta_{i}=y(\bmod N)\right]\end{array}$

## Bezout's theorem (Informal):

If $\operatorname{gcd}\left(\delta_{2}, \ldots, \delta_{M}\right)=1$ then there exist $k_{2}, \ldots, k_{M}$ such that $k_{2} \delta_{2}+\ldots+k_{M} \delta_{M}=1$ (over $\mathbb{Z}$ )

How can we guarantee that $\operatorname{gcd}\left(\delta_{2}, \ldots, \delta_{M}\right)=1$ ?

## Our Protocol (2): Our core technical Lemma

How can we guarantee that $\operatorname{gcd}\left(\delta_{2}, \ldots, \delta_{M}\right)=1$ ?

## Our Protocol (2): Our core technical Lemma

How can we guarantee that $\operatorname{gcd}\left(\delta_{2}, \ldots, \delta_{M}\right)=1$ ?

Caveat: Prover can choose the 'type' of $c$ to answer so that never gcd $\left(\left(c^{(i)}-c^{(1)}\right)_{i=2}^{M}\right)=1$ (e.g. only even $c$ )

## Our Protocol (2): Our core technical Lemma

How can we guarantee that $\operatorname{gcd}\left(\delta_{2}, \ldots, \delta_{M}\right)=1$ ?

Caveat: Prover can choose the 'type' of $c$ to answer so that never gcd $\left(\left(c^{(i)}-c^{(1)}\right)_{i=2}^{M}\right)=1$ (e.g. only even $c$ ) Our approach: Partially hide $c$ from the prover $\rightarrow c=\langle\vec{d}, \vec{b}\rangle$ where hidden, $\vec{b}$ sampled during the protocol

## Our Protocol (2): Our core technical Lemma

How can we guarantee that $\operatorname{gcd}\left(\delta_{2}, \ldots, \delta_{M}\right)=1$ ?

Caveat: Prover can choose the 'type' of $c$ to answer so that never $\operatorname{gcd}\left(\left(c^{(i)}-c^{(1)}\right)_{i=2}^{M}\right)=1$ (e.g. only even $c$ ) Our approach: Partially hide $c$ from the prover $\rightarrow c=\langle\vec{d}, \vec{b}\rangle$ where hidden, $\vec{b}$ sampled during the protocol

Our Information-Theoretical Lemma (Informal):
Let $\vec{d}=\left(d_{1}, \ldots, d_{\lambda}\right) \in\left(\{0,1\}^{\lambda}\right)^{\lambda}$ uniformly random, $\vec{b}=\left(b_{1}, \ldots, b_{n}\right) \in(\{0,1\})^{\lambda}$ and $c=\langle\vec{d}, \vec{b}\rangle$ then for any distribution of $\vec{b}$ one can obtain $M=$ poly $(\lambda)$ transcripts such that:

$$
\operatorname{Pr}\left[\operatorname{gcd}\left(\left(c^{(i)}-c^{(1)}\right)_{i=2}^{M}\right)=1\right]=1-\operatorname{neg|}(\lambda)
$$

## Our Protocol (3): Bootstraping via DV

Our approach:
$\otimes$ Partially hide $c$ from the prover: $c=\langle\vec{d}, \vec{b}\rangle$ where hidden, $\vec{b}$ sampled during the protocol
*Then Core-Lemma + gcd-extraction technique $\rightarrow$ Knowledge-Sound Protocol

## Our Protocol (3): Bootstraping via DV

Our approach:
$\otimes$ Partially hide $c$ from the prover: $c=\langle\vec{d}, \vec{b}\rangle$ where hidden, $\vec{b}$ sampled during the protocol *Then Core-Lemma + gcd-extraction technique $\rightarrow$ Knowledge-Sound Protocol

Question: How to hide $\vec{d}=\left(d_{1}, \ldots, d_{\lambda}\right)$ from $\mathscr{P}$ ?
$\rightarrow \mathscr{V}$ encrypts it $\rightarrow$ DV-model

## Performance-Extensions-Limitations

More on the paper:
$\star$ Range proofs over Subverted RSA groups

* Malicious and reusable DV keys

Implementation and Performance:

* Paillier Range proof (with malicious-verifier security):
$T(\mathscr{P})=192 \mathrm{~ms}, \mathrm{~T}(\mathscr{V})=125 \mathrm{~ms},|\pi|=11.05 \mathrm{~KB}$
Limitations:
* Designated-Verifier model
* Relatively expensive DV KeyGen for malicious zk
* Polynomial-reusability of DV keys (if verification oracles queries are assumed)


## Conclusions and summary

Summary:
$\star$ A new general extraction method for $\Sigma$-protocols.
$\star$ DV-AoK and range proof-protocols for subverted RSA groups.

Open questions:
\&Efficient Public-Verifier protocols for Subverted RSA groups?
*Apply the extraction technique to other contexts (e.g. lattice-based zk-proofs)?

## Thank you!

Full version: httos://eprint.iacr.oro/2023/364
Implementation: httos://github.com/volhovm/rsa-zkos-imol

