Hyperproofs: Aggregating and Maintaining Proofs in Vector Commitments

Shravan Srinivasan — University of Maryland
Alexander Chepurnoy — Ergo Platform
Charalampos Papamanthou — Yale University
Alin Tomescu — Aptos Labs
Yupeng Zhang — Texas A&M University

USENIX Security '22

Vector Commitment (VC)

- Short commitment to an ordered sequence of values [CF13]
- Open the commitment at specific positions with small proofs
- Should be *position binding*
- Example: Merkle tree

Useful in verifiable storage, stateless blockchains, SNARKs, and more

Problem statement

- **Aggregatable**: combine many proofs into a short proof
- $O(1)$-sized proofs
- aggregatable
- hard-to-maintain: $O(n)$
- RSA: [CF13, BBF19, LM19, CFG+20, TXN20]
- Bilinear: [KZG10, CF13, CDHK15, LM19, GRWZ20, TAB+20]

- **Maintainable**: update all proofs in sublinear time

Problem statement

- **Aggregatable**: combine many proofs into a short proof
- $O(1)$-sized proofs
- aggregatable
- hard-to-maintain: $O(n)$
- RSA: [CF13, BBF19, LM19, CFG+20, TXN20]
- Bilinear: [KZG10, CF13, CDHK15, LM19, GRWZ20, TAB+20]

- **Maintainable**: update all proofs in sublinear time
- $O(\log n)$-sized proofs
- not *efficiently* aggregatable
- easy-to-maintain: $O(\log n)$
- Tree-based: [Mer87, PSTY13, TCZ+20, Tom20, CNR+22]

Problem statement

- **Aggregatable**: combine many proofs into a short proof
- $O(1)$-sized proofs
- aggregatable
- hard-to-maintain: $O(n)$
- RSA: [CF13, BBF19, LM19, CFG$^+$20, TXN20]
- Bilinear: [KZG10, CF13, CDHK15, LM19, GRWZ20, TAB$^+$20]

- **Maintainable**: update all proofs in sublinear time
- $O(\log n)$-sized proofs
- not *efficiently* aggregatable
- easy-to-maintain: $O(\log n)$
- Tree-based: [Mer87, PSTY13, TCZ$^+$20, Tom20, CNR$^+$22]

Use SNARKs to aggregate the Merkle tree?

- 10-92M constraints
- Large memory overhead

Problem statement

- **Aggregatable**: combine many proofs into a short proof
- $O(1)$-sized proofs
- aggregatable
- hard-to-maintain: $O(n)$

- **Maintainable**: update all proofs in sublinear time
- $O(\log n)$-sized proofs
- not *efficiently* aggregatable
- easy-to-maintain: $O(\log n)$

Can we build a VC that is both
efficiently aggregatable and maintainable?

Why do we care about aggregatability and maintainability?

1. Efficiently aggregatable and maintainable VC is an *open* problem
    1.1 Inefficient SNARK-based Merkle aggregation is the alternative
2. Application to *stateless blockchain validation*

Our contributions

- First *efficiently* aggregatable and maintainable VC
- Aggregation $10\times$ to $41\times$ faster than Merkle based alternatives
  - 2 mins vs 20 mins to 1.4 hrs
- Maintainable: $2.58\ ms$ per update
- Small $O(log\ n)$ sized proofs ($1.44\ KiB$)
- Aggregated proof is $52\ KiB$
- Needs $O(n)$-sized trusted public parameters

Vector as **M**ulti **L**inear **E**xtension (MLE)

$$\text{vector } \mathbf{a} = [a_0, \ldots, a_7]$$

$$f(x_3, x_2, x_1) = a_0 \cdot (1 - x_3)(1 - x_2)(1 - x_1) \; +$$

Vector as **M**ulti **L**inear **E**xtension (MLE)

$$\text{vector } \mathbf{a} = [a_0, \ldots, a_7]$$

$$\begin{aligned} f(x_3, x_2, x_1) = a_0 \cdot (1 - x_3)(1 - x_2)(1 - x_1) \; + \\ a_1 \cdot (1 - x_3)(1 - x_2) \quad x_1 \quad + \end{aligned}$$

Vector as **M**ulti **L**inear **E**xtension (MLE)

$$\text{vector } \mathbf{a} = [a_0, \ldots, a_7]$$

$$\begin{aligned} f(x_3, x_2, x_1) = a_0 \cdot & (1 - x_3)(1 - x_2)(1 - x_1) \; + \\ a_1 \cdot & (1 - x_3)(1 - x_2) \; x_1 \; + \\ a_2 \cdot & (1 - x_3) \; x_2 \; (1 - x_1) \; + \end{aligned}$$

Vector as **M**ulti **L**inear **E**xtension (MLE)

$$\text{vector } \mathbf{a} = [a_0, \ldots, a_7]$$

$$\begin{aligned} f(x_3, x_2, x_1) = {} & a_0 \cdot (1 - x_3)(1 - x_2)(1 - x_1) \; + \\ & a_1 \cdot (1 - x_3)(1 - x_2) \; x_1 \; + \\ & a_2 \cdot (1 - x_3) \; x_2 \; (1 - x_1) \; + \\ & a_3 \cdot (1 - x_3) \; x_2 \; x_1 \; + \end{aligned}$$

Vector as **M**ulti **L**inear **E**xtension (MLE)

$$\text{vector } \mathbf{a} = [a_0, \ldots, a_7]$$

$$\begin{aligned} f(x_3, x_2, x_1) = a_0 \cdot (1 - x_3)(1 - x_2)(1 - x_1) &+ a_4 \cdot x_3 (1 - x_2)(1 - x_1) + \\ a_1 \cdot (1 - x_3)(1 - x_2)\, x_1 &+ \\ a_2 \cdot (1 - x_3)\, x_2\, (1 - x_1) &+ \\ a_3 \cdot (1 - x_3)\, x_2\, x_1 &+ \end{aligned}$$

Vector as **M**ulti **L**inear **E**xtension (MLE)

$$\text{vector } \mathbf{a} = [a_0, \ldots, a_7]$$

$$\begin{aligned}
f(x_3, x_2, x_1) = {} & a_0 \cdot (1 - x_3)(1 - x_2)(1 - x_1) + a_4 \cdot x_3 (1 - x_2)(1 - x_1) + \\
& a_1 \cdot (1 - x_3)(1 - x_2)\, x_1 + a_5 \cdot x_3 (1 - x_2)\, x_1 + \\
& a_2 \cdot (1 - x_3)\, x_2 (1 - x_1) + \\
& a_3 \cdot (1 - x_3)\, x_2\, x_1 +
\end{aligned}$$

Vector as **M**ulti **L**inear **E**xtension (MLE)

$$\text{vector } \mathbf{a} = [a_0, \dots, a_7]$$

$$\begin{aligned} f(x_3, x_2, x_1) = {} & a_0 \cdot (1 - x_3)(1 - x_2)(1 - x_1) + a_4 \cdot x_3 (1 - x_2)(1 - x_1) + \\ & a_1 \cdot (1 - x_3)(1 - x_2)\, x_1 + a_5 \cdot x_3 (1 - x_2)\, x_1 + \\ & a_2 \cdot (1 - x_3)\, x_2\, (1 - x_1) + a_6 \cdot x_3\, x_2\, (1 - x_1) + \\ & a_3 \cdot (1 - x_3)\, x_2\, x_1 + \end{aligned}$$

Vector as **M**ulti **L**inear **E**xtension (MLE)

$$\text{vector } \mathbf{a} = [a_0, \ldots, a_7]$$

$$\begin{aligned} f(x_3, x_2, x_1) = {} & a_0 \cdot (1 - x_3)(1 - x_2)(1 - x_1) + a_4 \cdot x_3 (1 - x_2)(1 - x_1) + \\ & a_1 \cdot (1 - x_3)(1 - x_2)\, x_1 + a_5 \cdot x_3 (1 - x_2)\, x_1 + \\ & a_2 \cdot (1 - x_3)\, x_2 \, (1 - x_1) + a_6 \cdot x_3 \, x_2 \, (1 - x_1) + \\ & a_3 \cdot (1 - x_3)\, x_2 \, x_1 + a_7 \cdot x_3 \, x_2 \, x_1 \end{aligned}$$

Vector as Multi Linear Extension (MLE)

$$\text{vector } \mathbf{a} = [a_0, \ldots, a_7]$$

$$\begin{aligned} f(x_3, x_2, x_1) = {} & a_0 \cdot (1 - x_3)(1 - x_2)(1 - x_1) + a_4 \cdot x_3 (1 - x_2)(1 - x_1) + \\ & a_1 \cdot (1 - x_3)(1 - x_2)\, x_1 + a_5 \cdot x_3 (1 - x_2)\, x_1 + \\ & a_2 \cdot (1 - x_3)\, x_2 \,(1 - x_1) + a_6 \cdot x_3\, x_2\, (1 - x_1) + \\ & a_3 \cdot (1 - x_3)\, x_2\, x_1 + a_7 \cdot x_3\, x_2\, x_1 \end{aligned}$$

For any index **i** with binary expansion (i_3, i_2, i_1), we have:

$$f(i_3, i_2, i_1) = a_i$$

Commitment to MLE polynomial [PST13][ZGK+17][ZGK+18]

$$\text{pp} := g^{s_1}, g^{s_2}, g^{s_3}, g^{s_2 s_1}, g^{s_3 s_2}, g^{s_3 s_1}, g^{s_3 s_2 s_1}$$

Commit:

$$\text{pst}(f) = g^{f(s_3, s_2, s_1)}$$

Commitment to MLE polynomial [PST13][ZGK$^+$17][ZGK$^+$18]

$$\mathsf{pp} := g^{s_1}, g^{s_2}, g^{s_3}, g^{s_2 s_1}, g^{s_3 s_2}, g^{s_3 s_1}, g^{s_3 s_2 s_1}$$

Commit:

$$\mathsf{pst}(f) = g^{f(s_3, s_2, s_1)}$$

Homomorphism:

$$\mathsf{pst}(f + f') = \mathsf{pst}(f) \cdot \mathsf{pst}(f')$$

Commitment to MLE polynomial [PST13][ZGK+17][ZGK+18]

$$\mathrm{pp} := g^{s_1}, g^{s_2}, g^{s_3}, g^{s_2 s_1}, g^{s_3 s_2}, g^{s_3 s_1}, g^{s_3 s_2 s_1}$$

Commit:

$$\mathrm{pst}(f) = g^{f(s_3, s_2, s_1)}$$

Homomorphism:

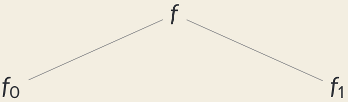
$$\mathrm{pst}(f + f') = \mathrm{pst}(f) \cdot \mathrm{pst}(f')$$

Commitment to vector $\mathbf{a}$ in Hyperproof is $\mathrm{pst}(f)$

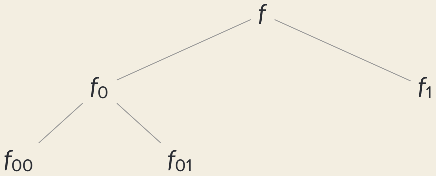
Multi **L**inear **T**ree (MLT)

f

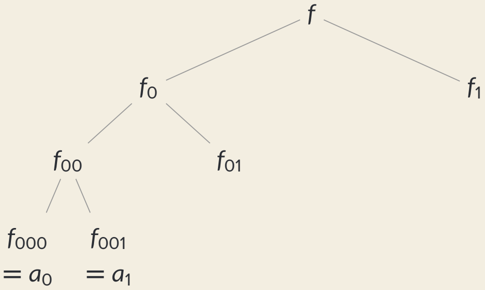
Multi Linear Tree (MLT)



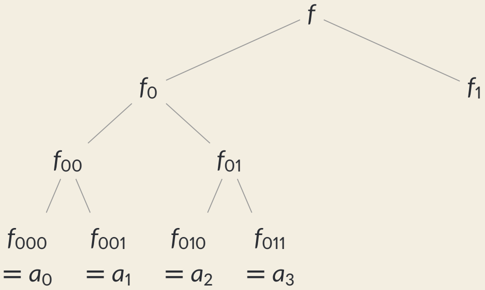
Multi Linear Tree (MLT)



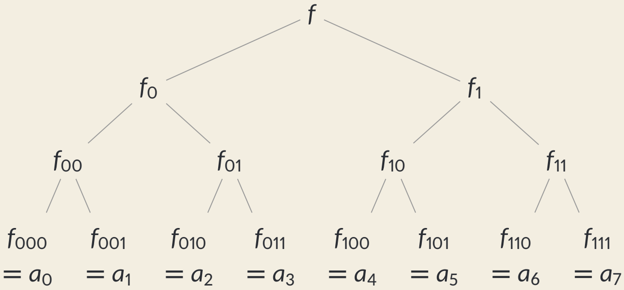
Multi Linear Tree (MLT)



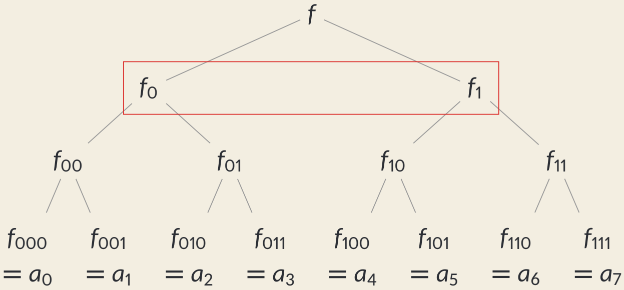
Multi Linear Tree (MLT)



Multi Linear Tree (MLT)

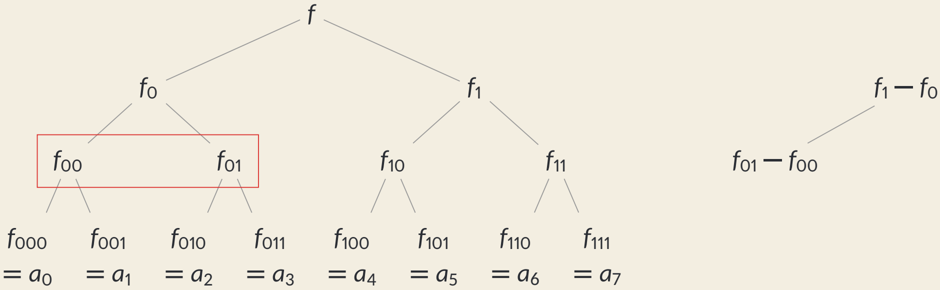


Multi Linear Tree (MLT)

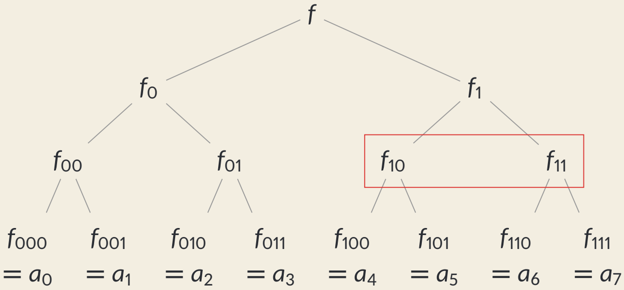


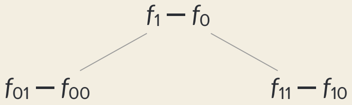
$f_1 - f_0$

Multi Linear Tree (MLT)

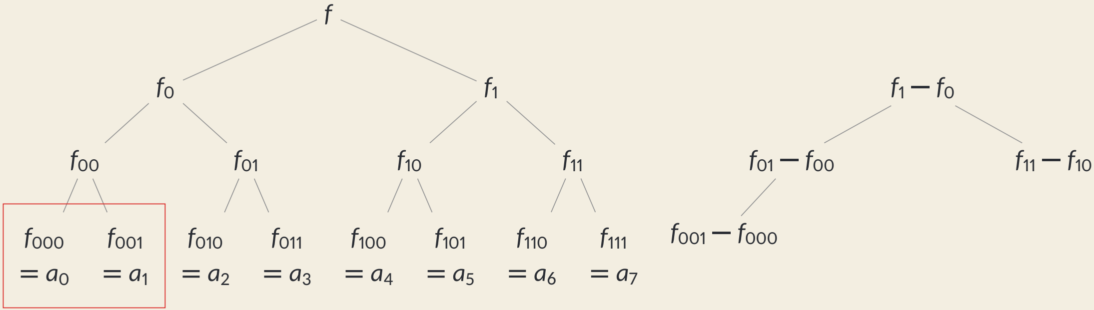


Multi Linear Tree (MLT)

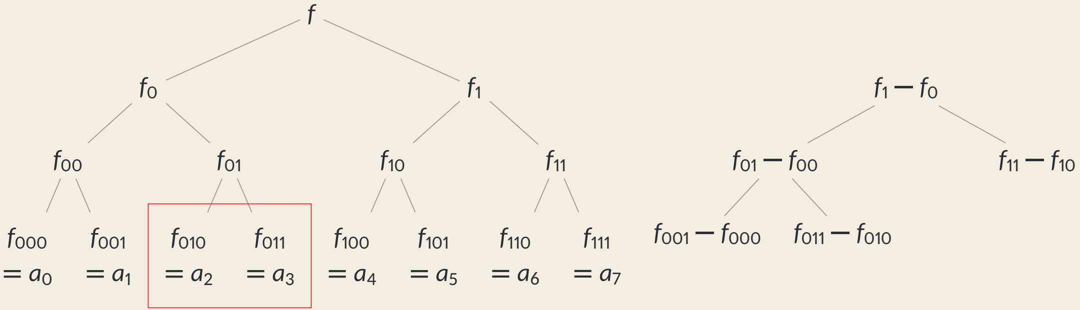




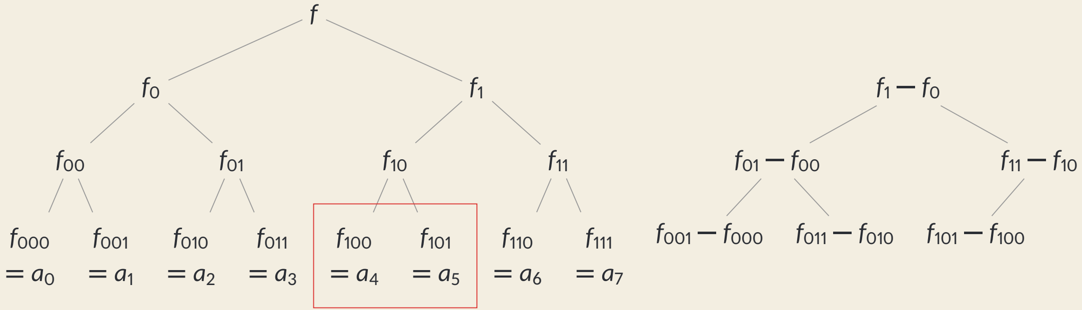
Multi Linear Tree (MLT)



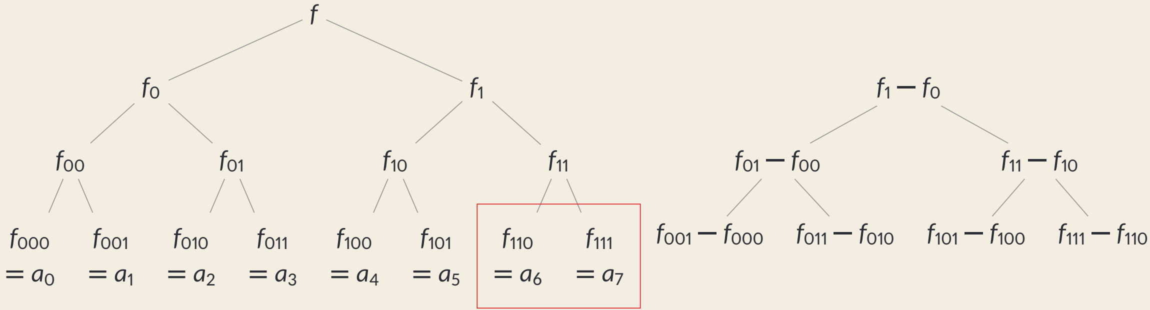
Multi Linear Tree (MLT)



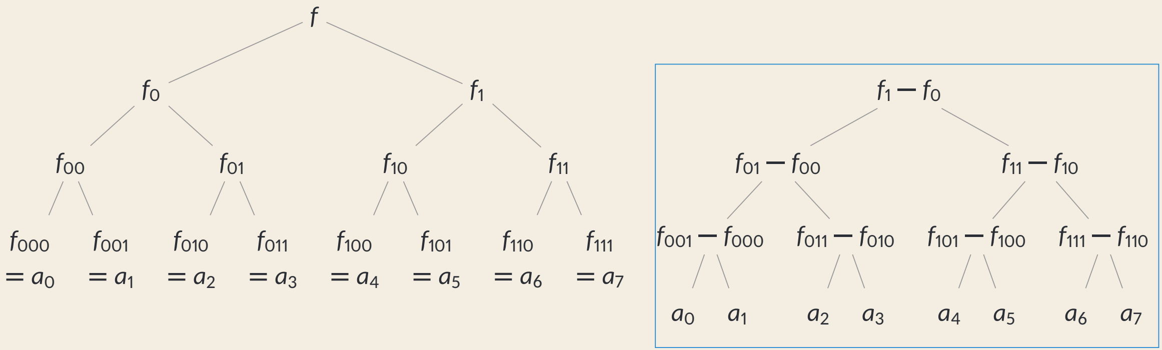
Multi Linear Tree (MLT)



Multi Linear Tree (MLT)

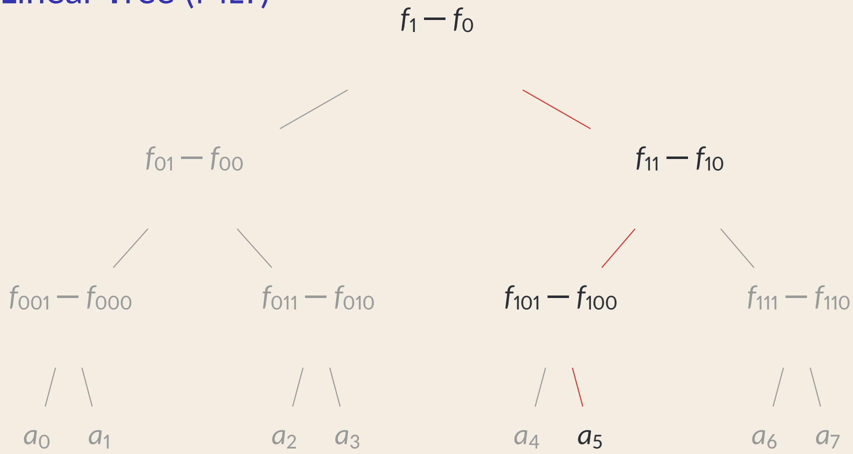


Multi Linear Tree (MLT)

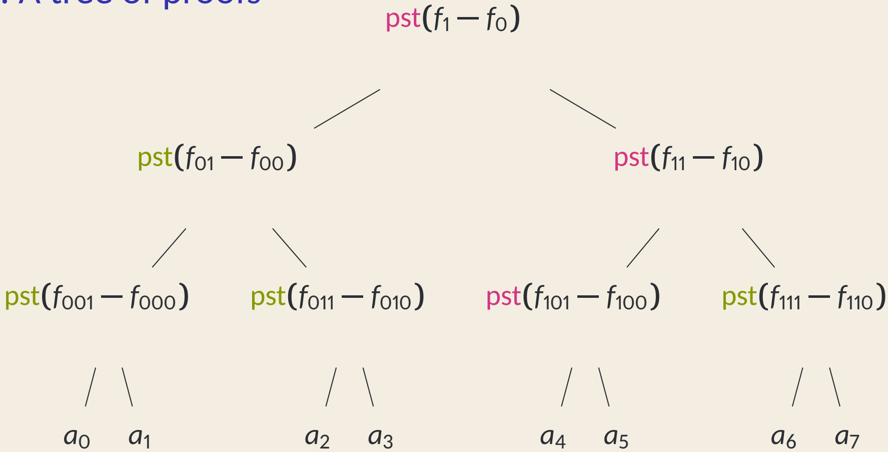


MLT

Multi Linear Tree (MLT)

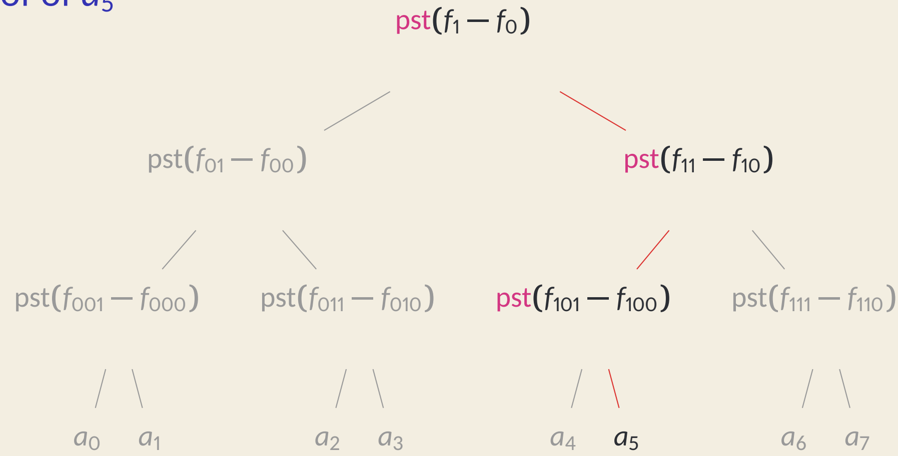


MLT: A tree of proofs



MLT computation $O(n \log n)$ time

Proof of a_5



$O(\log n)$ proof size and verification time

Maintainability

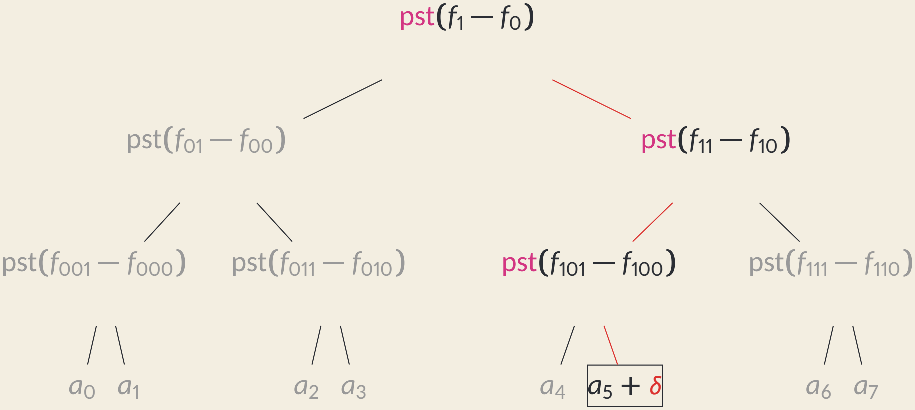
$$\mathrm{pst}(f_1 - f_0)$$

$$\mathrm{pst}(f_{01} - f_{00}) \qquad \mathrm{pst}(f_{11} - f_{10})$$

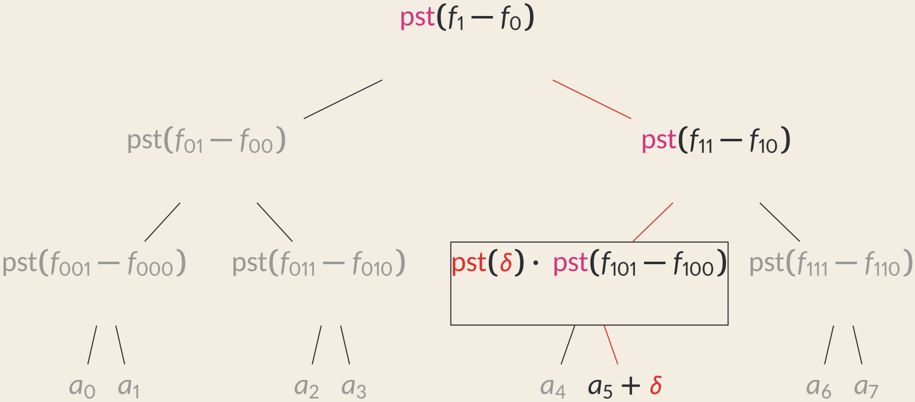
$$\mathrm{pst}(f_{001} - f_{000}) \qquad \mathrm{pst}(f_{011} - f_{010}) \qquad \mathrm{pst}(f_{101} - f_{100}) \qquad \mathrm{pst}(f_{111} - f_{110})$$

$$a_0 \quad a_1 \qquad a_2 \quad a_3 \qquad a_4 \quad a_5 \qquad a_6 \quad a_7$$

Maintainability



Maintainability



Maintainability

$$\mathrm{pst}(f_1 - f_0)$$

$$\mathrm{pst}(f_{01} - f_{00}) \qquad \boxed{\mathrm{pst}(\delta \cdot x_1) \cdot \mathrm{pst}(f_{11} - f_{10})}$$

$$\mathrm{pst}(f_{001} - f_{000}) \qquad \mathrm{pst}(f_{011} - f_{010}) \qquad \mathrm{pst}(\delta) \cdot \mathrm{pst}(f_{101} - f_{100}) \qquad \mathrm{pst}(f_{111} - f_{110})$$

$$a_0 \quad a_1 \qquad a_2 \quad a_3 \qquad a_4 \quad a_5 + \delta \qquad a_6 \quad a_7$$

Maintainability

$$\mathrm{pst}(\delta \cdot (1 - x_2) x_1) \cdot \mathrm{pst}(f_1 - f_0)$$

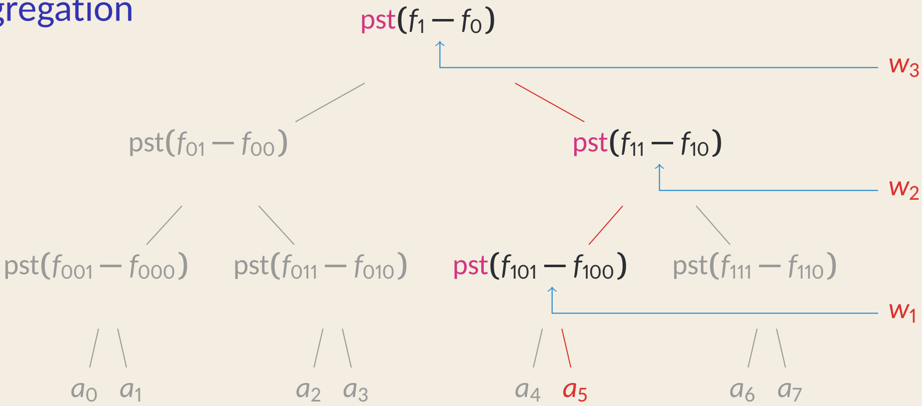
$$\mathrm{pst}(f_{01} - f_{00}) \qquad \mathrm{pst}(\delta \cdot x_1) \cdot \mathrm{pst}(f_{11} - f_{10})$$

$$\mathrm{pst}(f_{001} - f_{000}) \quad \mathrm{pst}(f_{011} - f_{010}) \quad \mathrm{pst}(\delta) \cdot \mathrm{pst}(f_{101} - f_{100}) \quad \mathrm{pst}(f_{111} - f_{110})$$

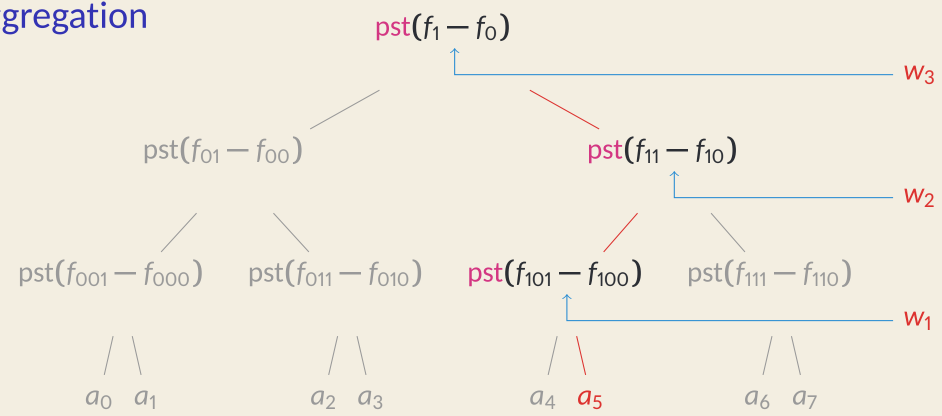
$$a_0 \quad a_1 \qquad a_2 \quad a_3 \qquad a_4 \quad a_5 + \delta \qquad a_6 \quad a_7$$

Update all proofs in $O(\log n)$ time

Aggregation

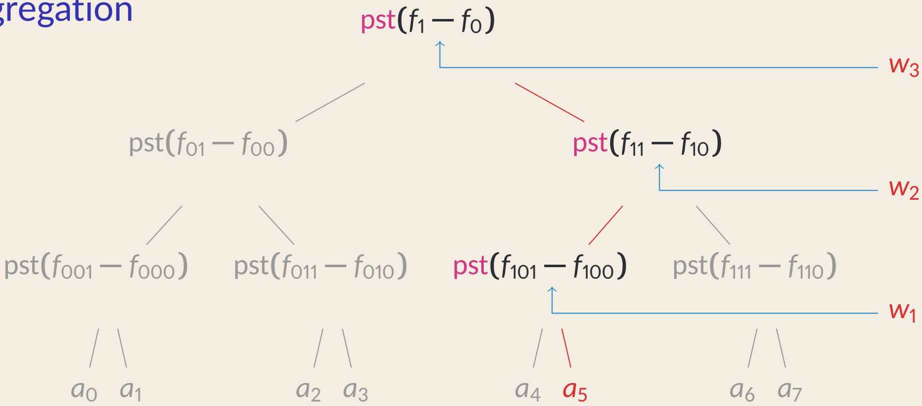


Aggregation



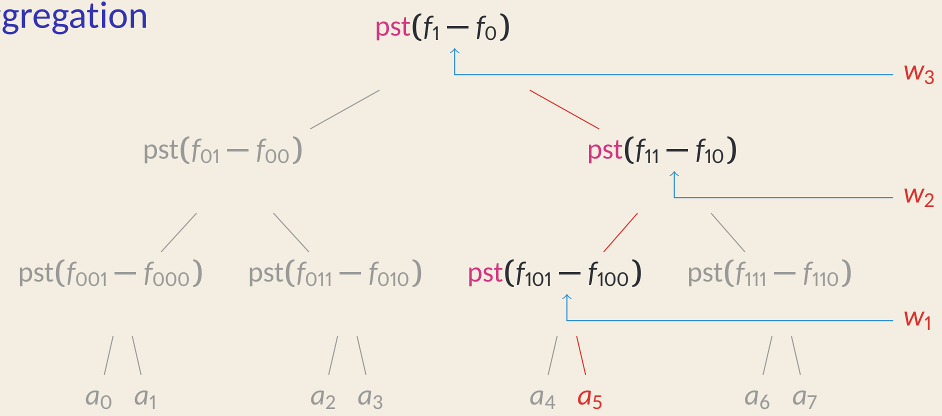
$$e(\mathrm{pst}(f)/g^{a_5}, g) = e(w_3, \qquad) \cdot e(w_2, \qquad) \cdot e(w_1, \qquad)$$

Aggregation



$$e(\mathrm{pst}(f)/g^{a_5}, g) = e(w_3, g^{(s_3-1)}) \cdot e(w_2, g^{(s_2-0)}) \cdot e(w_1, g^{(s_1-1)})$$

Aggregation



$$e(\mathrm{pst}(f)/g^{a_5}, g) = e(w_3, g^{(s_3-1)}) \cdot e(w_2, g^{(s_2-0)}) \cdot e(w_1, g^{(s_1-1)})$$

Inner-pairing product between the proof and the selectors

Proof aggregation

Say, ℓ is the height of MLT and the batch size is k

$$[w_{i,1}, \cdots, w_{i,\ell}] \qquad e(\mathsf{pst}(f)/g^{a_i}, g) = \prod_{i=1}^{\ell} e(w_{i,\ell}, g^{s-i_\ell})$$

- Prove knowledge of all $w_{i,j}$'s using inner-product argument (IPA) based on [BMM+21]
- Random linear combination of the pairing equations

$O(log\ (k \cdot \ell))$ aggregated proof size and $O(k \cdot \ell)$ aggregation time

Proof aggregation

Say, ℓ is the height of MLT and the batch size is k

$$\left\{[w_{i,1}, \cdots, w_{i,\ell}]\right\}_{i\in[k]} \qquad \left\{e(\mathsf{pst}(f)/g^{a_i}, g) = \prod_{i=1}^{\ell} e(w_{i,\ell}, g^{s-i_\ell})\right\}_{i\in[k]}$$

- Prove knowledge of all $w_{i,j}$'s using inner-product argument (IPA) based on [BMM+21]
- Random linear combination of the pairing equations

$O(log\,(k\cdot\ell))$ aggregated proof size and $O(k\cdot\ell)$ aggregation time

Benchmarks

transactions = 1024, height of the tree = 30, single threaded

Operation	**Hyperproofs**	**Merkle w/ Poseidon**
Agg	123 s	20 min
Aggr. Ver	17.4 s	0.11 s
Aggr. proof size	52 KiB	192 B

Benchmarks

transactions = 1024, height of the tree = 30, single threaded

Operation	**Hyperproofs**	**Merkle w/ Poseidon**
Agg	123 s	20 min
Aggr. Ver	17.4 s	0.11 s
Aggr. proof size	52 KiB	192 B
Block proposal (**P**)	2.23 min	81 min
Block validation (**V**)	17.5 s	0.18 s
Proof maintenance (**M**)	5.14 s	4.7 s
Total ($\mathbf{P} + (20 \cdot \mathbf{V}) + \mathbf{M}$)	8 min	81 min

Faster end-to-end performance than Merkle tree

Conclusion

- VC with efficient aggregation + maintainability + unstealability
- Unstealability incentivizes proof computation
- End-to-end performance is $10\times$ to $41\times$ faster than Merkle
- Algebraic alternative for Merkle tree

References I

[BBF19] Dan Boneh, Benedikt Bünz, and Ben Fisch.
Batching Techniques for Accumulators with Applications to IOPs and Stateless Blockchains.
In Alexandra Boldyreva and Daniele Micciancio, editors, *Advances in Cryptology – CRYPTO 2019*, pages 561–586, Cham, 2019. Springer, Springer International Publishing.
https://eprint.iacr.org/2018/1188.

[BMM+21] Benedikt Bünz, Mary Maller, Pratyush Mishra, Nirvan Tyagi, and Psi Vesely.
Proofs for Inner Pairing Products and Applications.
In *Advances in Cryptology – ASIACRYPT 2021*, 2021.

References II

[CDHK15] Jan Camenisch, Maria Dubovitskaya, Kristiyan Haralambiev, and Markulf Kohlweiss.
Composable and Modular Anonymous Credentials: Definitions and Practical Constructions.
In Tetsu Iwata and Jung Hee Cheon, editors, *Advances in Cryptology – ASIACRYPT 2015*, pages 262–288, Berlin, Heidelberg, 2015. Springer Berlin Heidelberg.

[CF13] Dario Catalano and Dario Fiore.
Vector Commitments and Their Applications.
In Kaoru Kurosawa and Goichiro Hanaoka, editors, *International Workshop on Public Key Cryptography*, pages 55–72, Berlin, Heidelberg, 2013. Springer, Springer Berlin Heidelberg.

References III

https://eprint.iacr.org/2011/495.

[CFG+20] Matteo Campanelli, Dario Fiore, Nicola Greco, Dimitris Kolonelos, and Luca Nizzardo.
Vector Commitment Techniques and Applications to Verifiable Decentralized Storage.
Cryptology ePrint Archive, Report 2020/149, 2020.
https://eprint.iacr.org/2020/149.

[CNR+22] Matteo Campanelli, Anca Nitulescu, Carla Ràfols, Alexandros Zacharakis, and Arantxa Zapico.
Linear-map Vector Commitments and their Practical Applications.
Cryptology ePrint Archive, Paper 2022/705, 2022.
https://eprint.iacr.org/2022/705.

References IV

[GRWZ20] Sergey Gorbunov, Leonid Reyzin, Hoeteck Wee, and Zhenfei Zhang. Pointproofs: Aggregating Proofs for Multiple Vector Commitments. In *Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security*, CCS '20, page 2007–2023, New York, NY, USA, 2020. Association for Computing Machinery. https://eprint.iacr.org/2020/419.

[KZG10] Aniket Kate, Gregory M. Zaverucha, and Ian Goldberg. Constant-Size Commitments to Polynomials and Their Applications. In Masayuki Abe, editor, *Advances in Cryptology - ASIACRYPT 2010*, pages 177–194, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg. https://www.iacr.org/archive/asiacrypt2010/6477178/6477178.pdf.

References V

[LM19] Russell W.F. Lai and Giulio Malavolta.
Subvector Commitments with Application to Succinct Arguments.
In Alexandra Boldyreva and Daniele Micciancio, editors, *Advances in Cryptology – CRYPTO 2019*, pages 530–560, Cham, 2019. Springer, Springer International Publishing.
https://eprint.iacr.org/2018/705.

[Mer87] Ralph C. Merkle.
A Digital Signature Based on a Conventional Encryption Function.
In Carl Pomerance, editor, *Advances in Cryptology — CRYPTO '87*, pages 369–378, Berlin, Heidelberg, 1987. Springer Berlin Heidelberg.

References VI

[PST13] Charalampos Papamanthou, Elaine Shi, and Roberto Tamassia. Signatures of Correct Computation. In Amit Sahai, editor, *Theory of Cryptography*, pages 222–242, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg. https://eprint.iacr.org/2011/587.

[PSTY13] Charalampos Papamanthou, Elaine Shi, Roberto Tamassia, and Ke Yi. Streaming Authenticated Data Structures. In Thomas Johansson and Phong Q. Nguyen, editors, *Advances in Cryptology – EUROCRYPT 2013*, pages 353–370, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg.

References VII

[TAB+20] Alin Tomescu, Ittai Abraham, Vitalik Buterin, Justin Drake, Dankrad Feist, and Dmitry Khovratovich.
Aggregatable Subvector Commitments for Stateless Cryptocurrencies.
In Clemente Galdi and Vladimir Kolesnikov, editors, *Security and Cryptography for Networks*, pages 45–64, Cham, 2020. Springer International Publishing.

[TCZ+20] Alin Tomescu, Robert Chen, Yiming Zheng, Ittai Abraham, Benny Pinkas, Guy Golan Gueta, and Srinivas Devadas.
Towards Scalable Threshold Cryptosystems.
In *2020 IEEE Symposium on Security and Privacy (SP)*, pages 877–893, May 2020.

References VIII

[Tom20] Alin Tomescu.
How to Keep a Secret and Share a Public Key (Using Polynomial Commitments).
PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 2020.

[TXN20] Alin Tomescu, Yu Xia, and Zachary Newman.
Authenticated Dictionaries with Cross-Incremental Proof (Dis)aggregation.
Cryptology ePrint Archive, Report 2020/1239, 2020.
`https://ia.cr/2020/1239`.

References IX

[ZGK+17] Yupeng Zhang, Daniel Genkin, Jonathan Katz, Dimitrios Papadopoulos, and Charalampos Papamanthou.
vSQL: Verifying Arbitrary SQL Queries over Dynamic Outsourced Databases.
In *2017 IEEE Symposium on Security and Privacy (SP)*, pages 863–880, May 2017.

[ZGK+18] Yupeng Zhang, Daniel Genkin, Jonathan Katz, Dimitrios Papadopoulos, and Charalampos Papamanthou.
vRAM: Faster Verifiable RAM with Program-Independent Preprocessing.
In *2018 IEEE Symposium on Security and Privacy (SP)*, pages 908–925, May 2018.