SnarkExpress (2025.09)

Kurt Pan @ ZKPunk

October 15, 2025

Contents

1 [cryptoeprint:2025/1546] Incrementally Verifiable Computation for NP from Standard Assumptions

In this work, we observe that the Gentry-Wichs barrier can be overcome for IVC for NP. We show the following two results: - Assuming subexponential $i\mathcal{O}$ and LWE (or bilinear maps), we construct IVC for all NP with proof size $\mathsf{poly}(|x_i|, \log T)$. - Assuming subexponential $i\mathcal{O}$ and injective PRGs, we construct IVC for trapdoor IVC languages where the proof-size is $\mathsf{poly}(\log T)$. Informally, an IVC language has a trapdoor if there exists a (not necessarily easy to find) polynomial-sized circuit that determines if a configuration x_i is reachable from x_0 in i steps.

2 [cryptoeprint:2025/1547] Silent Threshold Cryptography from Pairings: Expressive Policies in the Plain Model

In this work, we introduce a new pairing-based approach for constructing threshold signatures and encryption schemes with silent setup. On the one hand, our techniques directly allow us to support expressive policies like monotone Boolean formulas in addition to thresholds. On the other hand, we only rely on basic algebraic tools (i.e., a simple cross-term cancellation strategy), which yields constructions with shorter signatures and ciphertexts compared to previous pairing-based constructions. As an added bonus, we can also prove (static) security under q-type assumptions in the plain model.

Concretely, the signature size in our distributed threshold signature scheme is 3 group elements and the ciphertext size in our distributed threshold encryption scheme is 4 group elements (together with a short tag).

3 [cryptoeprint:2025/1548] Pairing-Based Aggregate Signatures without Random Oracles

In this work, we focus on simple aggregate signatures in the plain model. We construct a pairing-based aggregate signature scheme that supports aggregating an a priori bounded number of signatures N. The size of the aggregate signature is just two group elements. Security relies on the (bilateral) computational Diffie-Hellman (CDH) problem in a pairing group. To our knowledge, this is the first group-based aggregate signature in the plain model where (1) there is no restriction on what type of signatures can be aggregated; (2) the aggregated signature contains a constant number of group elements; and (3) security is based on static falsifiable assumptions in the plain model. The limitation of our scheme is that our scheme relies on a set of public parameters (whose size scales with N) and individual signatures (before aggregation) also have size that scale with N. Essentially, individual signatures contain some additional hints to enable aggregation. Our starting point is a new notion of slotted aggregate signatures. Here, each signature is associated with a "slot" and we only support aggregating signatures associated with distinct slots. We then show how to generically lift a slotted aggregate signature scheme into a standard aggregate signature scheme at the cost of increasing the size of the original signatures.

4 [cryptoeprint:2025/1554] UniCross: A Universal Cross-Chain Payment Protocol with On-demand Privacy and High Scalability

This paper proposes a universal cross-chain payment framework. This framework enables payments across a wide range of blockchains since it is independent of any specific blockchain features. Moreover, this framework provides on-demand privacy and high scalability. To instantiate the framework, we introduce UniCross, a novel universal cross-chain payment protocol. Concretely, we utilize the ring learning with errors (RLWE)-based encryption scheme and propose a new non-interactive zero-knowledge (NIZK) protocol, named HybridProof, to construct UniCross. We formally define the security of the universal cross-chain payment framework and prove the universal composability (UC) security of UniCross. The proof-of-concept implementation and evaluation demonstrate that (1) UniCross consumes up to 78% and 94% less communication and computation cost than the state-of-the-art work; (2) UniCross achieves a throughput (∼360 tps) 36× that of the state-of-the-art work (∼10 tps).

5 [cryptoeprint:2025/1558] Lower Bounding Update Frequency in Short Accumulators and Vector Commitments

We study the inherent limitations of additive accumulators and updatable vector commitments (VCs) with constant-size digest (i.e., independent of the number of committed elements). Specifically, we prove two lower bounds on the expected number of membership proofs that must be updated when a *single* element is added (or updated) in such data structures. Our results imply that when the digest bit length approaches the concrete security level, then the expected number of proofs invalidated due to an append operation for a digest committing to n elements is nearly maximal: $n - \mathsf{negl}(\lambda)$ in the case of exponential-size universes, and $n - o(n)$ for super-polynomial universes. Our results have significant implications for stateless blockchain designs relying on constant-size VCs, suggesting that the overhead of frequent proof updates may offset the benefits of reducing global state storage.

6 [cryptoeprint:2025/1566] Lattice-based Threshold Blind Signatures

We present the first construction of a threshold blind signature secure in the post-quantum setting, based on lattices. We prove its security under an interactive variant of the SIS assumption introduced in [Agrawal et al., CCS'22]. Our construction has a reasonable overhead of a factor of roughly 1.4 X to 2.5 X in signature size over comparable non-threshold blind signatures over lattices under heuristic but natural assumptions.

7 [cryptoeprint:2025/1569] How Hard Can It Be to Formalize a Proof? Lessons from Formalizing CryptoBox Three Times in EasyCrypt

We present a new security proof for the generic construction of a PKAE scheme from a NIKE and AE scheme, written in a code-based, game-playing style à la Bellare and Rogaway, and compare it to the same proof written in the style of state-separating proofs, a methodology for developing modular game-playing security proofs. Additionally, we explore a third "blended" style designed to avoid anticipated difficulties with the formalization. Our findings suggest that the choice of definition style impacts proof complexity—including, we argue, in detailed pen-and-paper proofs—with trade-offs depending on the proof writer's goals.

8 [cryptoeprint:2025/1576] Compressed verification for post-quantum signatures with long-term public keys

A method to replace large public keys in GPV-style signatures with smaller, private verification keys. This significantly reduces verifier storage and runtime while preserving security. Applied to the conservative, short-signature schemes Wave and Squirrels.

9 [cryptoeprint:2025/1580] IronDict: Transparent Dictionaries from Polynomial Commitments

We present IronDict, a transparent dictionary construction based on polynomial commitment schemes. Transparent dictionaries enable an untrusted server to maintain a mutable dictionary and provably serve clients lookup queries. Our construction makes black-box use of a generic multilinear polynomial commitment scheme and inherits its security notions, i.e. binding and zero-knowledge. We implement our construction with the recent KZH scheme and find that a dictionary with 1 billion entries can be verified on a consumer-grade laptop in 35 ms, a 300× improvement over the state of the art, while also achieving 150,000× smaller proofs (8 KB). In addition, our construction ensures perfect privacy with concretely efficient costs for both the client and the server. We also show fast-forwarding techniques based on incremental verifiable computation (IVC) and checkpoints to enable even faster client auditing.

10 [cryptoeprint:2025/1588] Query-Optimal IOPPs for Linear-TIme Encodable Codes

We present the first IOPP for a linear-time encodable code that achieves linear prover time and $O(\lambda)$ query complexity, for a broad range of security parameters λ. No prior work is able to simultaneously achieve this efficiency: it either supports linear-time encodable codes but with worse query complexity [FICS; ePrint 2025], or achieves $O(\lambda)$ query complexity but only for quasilinear-time encodable codes [Minzer, Zheng; FOCS 2025]. Furthermore, we prove a matching lower bound that shows that the query complexity of our IOPP is asymptotically optimal (up to additive factors) for codes with constant rate. We obtain our result by tackling a ubiquitous subproblem in IOPP constructions: checking that a batch of claims hold. Our novel solution to this subproblem is twofold. First, we observe that it is often sufficient to ensure that, with all but negligible probability, most of the claims hold. Next, we devise a new 'lossy batching' technique which convinces a verifier of the foregoing promise with lower query complexity than that required to convince it that all the claims hold. This method differs significantly from the line-versus-point test used to achieve query-optimal IOPPs (for quasilinear-time encodable codes) in prior work [Minzer, Zheng; FOCS 2025], and may be of independent interest. Our IOPP can handle all codes that support efficient codeswitching [Ron-Zewi, Rothblum; JACM 2024], including several linear-time encodable codes. Via standard techniques, our IOPP can be used to construct the first (to the best of our knowledge) IOP for NP with $O(n)$ prover time and $O(\lambda)$ query complexity. We additionally show that our IOPP (and by extension the foregoing IOP) is round-by-round tree-extractable and hence can be used to construct a SNARK in the random oracle model with $O(n)$ prover time and $O(\lambda \log n)$ proof size.

[**cryptoeprint:2024/474**]

11 [cryptoeprint:2025/1593] Leveraging Smaller Finite Fields for More Efficient ZK-Friendly Hash Functions

In this paper, we examine the applicability of the generalized subfield construction and the possibility of improvements on ZK-friendly hash functions. As a case study, we focus on a recent ZK-friendly hash function Vision Mark-32 presented by Ashur et al. in [IACR Preprint 2024/633]. In particular, instead of using a 24×24 MDS matrix over $\mathbb{F}_{2^{32}}$ for a 24×1 column input over $\{0,1\}^{32}$, we suggest separating the 24×1 column input over $\{0,1\}^{32}$ into four 24×1 subcolumns over $\{0,1\}^8$ and then using a 24×24 MDS matrix over $\mathbb{F}_{2^8}$ for each subcolumn. This method still keeps the maximum diffusion property without any compromise and provides simplicity and efficiency. For example, it is possible to significantly decrease the required LUT values to 265 from about 9200 and FF values to 102 from about 4600 for the hardware implementation. We also highlight that we do not need any additional tricks such as NTT for field multiplications. We also push the theoretical boundaries of the generalized subfield construction to see how much small finite fields we can use, examine the arithmetization complexity, and discuss its applicability to other ZK-friendly hash functions.

12 [cryptoeprint:2025/1596] On GPU acceleration of PQC algorithms

This paper investigates their acceleration using GPUs. We implemented Dilithium, FrodoKEM, and SPHINCS+ on GPUs using CUDA and benchmarked them together with an existing GPU implementation of Kyber on a Tesla A100 and on a RTX 2070 Super. Dilithium performed convincingly on both GPUs, achieving speed-ups in key generation, signing, and verify by factors of around 820, 2,724 and 2,609 on the A100 and 198, 714 and 802 on the RTX 2070 using the optimal batch sizes. SPHINCS+ achieved speed-ups by factors of around 715, 4,114 and 5,915 on the A100 and 193, 193 and 134 on the RTX 2070. FrodoKEM's key generation, encapsulation, and decapsulation on the A100 were accelerated by factors of 9,989, 4,726, and 3,566. It performed speed-up factors of 107, 108, and 206 on the RTX 2070, respectively. We compared to Kyber's acceleration factors of 476, 513 and 1782 on the A100 and 18.5, 17.4 and 184.5 on the RTX 2070. In addition, we investigated the effect of using a variable set of CUDA streams for FrodoKEM. Here, using 8 streams, a speedup of another 2% could be achieved.

13 [cryptoeprint:2025/1646] Scalable zkSNARKs for Matrix Computations: A Generic Framework for Verifiable Deep Learning

Sublinear proof sizes have recently become feasible in verifiable machine learning (VML), yet no approach achieves the trio of strictly linear prover time, logarithmic proof size and verification time, and architecture privacy. Hurdles persist because we lack a succinct commitment to the full neural network and a framework for heterogeneous models, leaving verification dependent on architecture knowledge. Existing limits motivate our new approach: a unified proof-composition framework that casts VML as the design of zero-knowledge succinct non-interactive arguments of knowledge (zkSNARKs) for matrix computations. Representing neural networks with linear and non-linear layers as a directed acyclic graph of atomic matrix operations enables topology-aware composition without revealing the graph. Modeled this way, we split proving into a reduction layer and a compression layer that attests to the reduction with a proof of proof. At the reduction layer, inspired by reduction of knowledge (Crypto '23), root-node proofs are reduced to leaf-node proofs under an interface standardized for heterogeneous linear and non-linear operations. Next, a recursive zkSNARK compresses the transcript into a single proof while preserving architecture privacy. Complexity-wise, for a matrix expression with M atomic operations on $n \times n$ matrices, the prover runs in $O(Mn^2)$ time while proof size and verification time are $O(\log(Mn))$, outperforming known VML systems. Honed for this framework, we formalize relations directly in matrices or vectors—a more intuitive form for VML than traditional polynomials. Our LiteBullet proof, an inner-product proof based on folding and its connection to sumcheck (Crypto '21), yields a polynomial-free alternative. With these ingredients, we reconcile heterogeneity, zero-knowledge, succinctness, and architecture privacy in a single VML system.

14 [cryptoeprint:2025/1653] Distributed SNARK via folding schemes

In this paper, we propose a novel distributed SNARK system constructed by compiling distributed PIOP with additively homomorphic polynomial commitment, rather than distributed polynomial commitment. The core technical component is distributed SumFold, which folds multiple sum-check instances into one. After the folding process, only one prover is required to perform polynomial commitment openings. It facilitates compilation with SamaritanPCS, which is a recently proposed additively homomorphic multilinear polynomial commitment scheme. The resulting SNARK system is specifically optimized for data-parallel circuits. Compared to prior HyperPlonk-based distributed proof systems (e.g., Hyperpianist and Cirrus), our construction achieves improvements in both proof size and prover time.

15 [cryptoeprint:2025/1663] IVC in the Open-and-sign Random Oracle Model

To mitigate the theoretical challenges, we present the Open-and-Sign Random Oracle Model (osROM) as an extension to the signed random oracle of Chiesa and Tromer (ICS '10). This model, while strictly harder to instantiate than the Random Oracle Model, allows the design of protocols that can efficiently verify calls to the oracle and support straight-line extractors. As a result, IVC constructions in the osROM can be shown to have provable security for polynomial depths of computation. Under our new model, we construct a framework to build secure IVC schemes from simple non-interactive reductions of knowledge. Our construction natively supports cycles of elliptic curves in the style of Ben-Sasson et al. (CRYPTO '14), thus answering the practical challenge outlined above. Finally, we analyze the HyperNova (CRYPTO '24) IVC scheme in the osROM and show that it is secure over a two-cycle of elliptic curves, for polynomial depths of computation.

16 [cryptoeprint:2025/1698] SNARK Lower Bounds via Communication Complexity

We initiate the study of lower bounding the verification time of Succinct Non-interactive ARguments of Knowledge (SNARKs) built in the Polynomial Interactive Oracle Proof + Polynomial Commitment Scheme paradigm. The verification time of these SNARKs is generally dominated by the polynomial commitment scheme, and so we want to understand if polynomial commitment schemes admit lower bounds on the verification time. By recognizing that polynomial commitment schemes are also often built by applying cryptography to some information-theoretic core protocol, we seek to separate this core from the cryptography in a way that meaningfully captures the verification time required by the polynomial commitment scheme verifier. We provide strong evidence that several polynomial commitment schemes have (nearly) optimal verifier times. Our evidence comes from connecting polynomial commitment schemes to certain information-theoretic protocols known as communication protocols from the field of communication complexity, a link which we believe to be of independent interest. Through this lens, we model the verifier work in the cryptographic protocols as information (i.e., number of bits) exchanged between parties in the communication protocols, allowing us to leverage lower bounds from communication complexity. These lower bounds give strong evidence that the verifier time in these polynomial commitment schemes must be at least the number of bits exchanged in the communication protocol. We extract the communication protocol cores of three polynomial commitment schemes and lower bound the bits exchanged in these cores. The lower bounds we obtain match (up to poly-logarithmic factors) the best-known (asymptotic) verification times of the polynomial commitment schemes we examine in this work. Specifically, we show that for univariate/multilinear polynomials of size $N = 2^n$: - the communication core of Hyrax PCS (Wahby et al., S&P 2016) requires $\Omega(\sqrt{N})$ bits to be exchanged; - the communication core of Bulletproofs PCS (Bootle et al., EUROCRYPT 2016; Bünz et al., S&P 2018) requires $\Omega(N)$ bits to be exchanged; and - the communication core of Dory PCS (Lee, TCC 2021) requires $\Omega(\log(N))$ bits to be exchanged. Our results strongly suggest a negative answer to a longstanding open question on whether the Bulletproofs verifier can be made sublinear time.

17 [cryptoeprint:2025/1709] The zkVot Protocol: A Distributed Computation Protocol for Censorship Resistant Anonymous Voting

zkVot is a client side trustless distributed computation protocol that utilizes zero knowledge proving technology. It is designed to achieve anonymous and censorship resistant voting while ensuring scalability.

18 [cryptoeprint:2025/1712] The Syndrome-Space Lens: A Complete Resolution of Proximity Gaps for Reed-Solomon Codes

We resolve the Correlated Agreement (CA) problem for Reed-Solomon codes up to the information-theoretic capacity limit by introducing a fundamental change of basis: from the traditional evaluation domain to the syndrome space. Viewed through this "Syndrome-Space Lens," the problem of proximity testing transforms into a transparent question of linear-algebraic geometry: a single affine line of syndromes traversing a family of low-dimensional subspaces. This new perspective makes a sharp phase transition at the capacity boundary visible, allowing for a complete characterization of the problem's behavior across all parameter regimes, yielding short, self-contained proofs.

19 [cryptoeprint:2025/1715] UltraMixer: A Compliant Zero-Knowledge Privacy Layer for Tokenized Real-World Assets

We present UltraMixer, a noncustodial privacy layer natively compatible with ERC-3643. Compliance is enforced at the boundary via zero-knowledge proofs of whitelist membership, while in-mixer transfers and atomic trades operate over commitments with nullifiers to prevent double-spend. A generalized UTXO encoding supports heterogeneous assets (fungible and nonfungible) under a unified commitment scheme. For selective disclosure, UltraMixer provides a verdict-only Δ-Window Proof of Holding that attests to continuous ownership across a time interval without revealing balances, identities, or linkages. Gas-aware batching and composable emergency controls (pause, freeze/unfreeze, force-transfer) preserve practicality and governance.

The resulting architecture delivers regulator-compatible confidentiality for permissioned RWA markets.

20 [cryptoeprint:2025/1723] Space-Deniable Proofs

We introduce and construct a new proof system called Non-interactive Arguments of Knowledge or Space (NArKoS), where a space bounded prover can convince a verifier they know a secret, while having access to sufficient space allows one to forge indistinguishable proofs without the secret. An application of NArKoS are space-deniable proofs, which are proofs of knowledge (say for authentication in access control) that are sound when executed by a lightweight device like a smart-card or an RFID chip that cannot have much storage, but are deniable (in the strong sense of online deniability) as the verifier, like a card reader, can efficiently forge such proofs. We construct NArKoS in the random oracle model using an OR-proof combining a sigma protocol (for the proof of knowledge of the secret) with a new proof system called simulatable Proof of Transient Space (simPoTS). We give two different constructions of simPoTS, one based on labelling graphs with high pebbling complexity, a technique used in the construction of memory-hard functions and proofs of space, and a more practical construction based on the verifiable space-hard functions from TCC'24 where a prover must compute a root of a sparse polynomial. In both cases, the main challenge is making the proofs efficiently simulatable.

21 [cryptoeprint:2025/1724] Efficient Aggregate Anonymous Credentials for Decentralized Identity

In this paper, we first introduce what we call aggregate tag-based signatures and describe an efficient instantiation. We then leverage the latter together with structure-preserving signatures and signatures of knowledge to construct an efficient aggregate anonymous credential scheme. We finally, formally evaluate the security of the proposed schemes and run benchmarks to showcase the practicality of the resulting scheme and its relevance for decentralized identity applications.

22 [cryptoeprint:2025/1732] Zero-Knowledge AI Inference with High Precision

In this work, we present ZIP, an efficient and precise commit and prove zero-knowledge SNARK for AIaaS inference (both linear and non-linear layers) that natively supports IEEE-754 double-precision floating-point semantics while addressing reliability and privacy challenges inherent in AIaaS. At its core, ZIP introduces a novel relative-error-driven technique that efficiently proves the correctness of complex non-linear layers in AI inference computations without any loss of precision, and hardens existing lookup-table and range proofs with novel arithmetic constraints to defend against malicious provers. We implement ZIP and evaluate it on standard datasets (e.g., MNIST, UTKFace, and SST-2). Our experimental results show, for non-linear activation functions, ZIP reduces circuit size by up to three orders of magnitude while maintaining the full precision required by modern AI workloads.

23 [cryptoeprint:2025/1741] Full L1 On-Chain ZK-STARK+PQC Verification on Solana: A Measurement Study

This work investigates whether a fully on-chain pipeline that verifies both a ZK-STARK and a post-quantum signature can operate within Solana L1's compute and memory constraints. Our prototype adapts Winterfell 0.12 with a dedicated SHA-256 hashv syscall path to reduce hashing overhead, suppresses inlining in FRI hotspots to respect SBF (Solana BPF) stack limits, and uses a custom bump allocator synchronized with requested heap frames.

24 [cryptoeprint:2025/1759] Plonk is Simulation Extractable in ROM Under Falsifiable Assumptions

Solving a long-standing open problem, Faonio, Fiore, and Russo proved that the widely used Plonk zk-SNARK is simulation extractable. However, their proof assumes both the random oracle model (ROM) and the algebraic group model. We prove that the same holds in the ROM under falsifiable assumptions. We combine the template of Faust et al., who proved that simulation extractability follows from knowledge soundness, (weak) unique response, and trapdoorless zero-knowledge, with the recent result of Lipmaa, Parisella, and Siim (Crypto 2025), who proved that Plonk has knowledge soundness in the ROM under falsifiable assumptions. For this, we prove that Plonk satisfies new variants of the weak unique response and trapdoorless zero-knowledge properties. We prove that several commonly used gadgets, like the linearization trick, are not trapdoorless zero-knowledge when considered as independent commit-and-prove zk-SNARKs.