Abstract

This survey uniquely approaches zero-knowledge proofs (ZKPs) through the lens of folding schemes, offering a fresh framework to analyze efficiency, scalability, and post-quantum resilience. By focusing on folding, we unify diverse protocols, clarify trade-offs, and identify practical engineering constraints, providing both researchers and practitioners with actionable insights. Folding schemes have emerged as the simplest and fastest approach to incrementally verifiable computation (IVC), enabling recursive zero-knowledge arguments with constant recursion overhead. We present a unifying model of folding-based ZKPs across R1CS, Plonkish/CCS, and AIR; synthesize the state of the art from Nova, SuperNova, HyperNova, and cycle-of-curves instantiations to recent post-quantum lattice-based foldings; provide a rigorous comparison of prover time, verifier work, proof size, setup assumptions, and recursion overhead; and map real deployments—including Lurk/Nova, Sonobe-based light clients, and VIMz-style media proofs—to practical constraints. Finally, we highlight open problems such as hybrid elliptic-curve–lattice designs and engineering targets for memory-bounded provers, showing how this folding-centric view advances both theoretical understanding and real-world deployment of ZKPs.