

The Technical System and Architecture of Blockchain Privacy Protection based on Encryption

Yujia Xian^{1*}

¹School of Finacial Technology, Hebei Finance University, 071000 Baoding, Hebei, China

Abstract. The concept of blockchain has transformed the trust concept by decentralizing, non-modifiable, and transparent, but there is a certain conflict between the principle of public verifiability and data privacy. As DeFi and cross-institutional data collaboration should grow, it has become a fundamental concern to have the confidentiality of this data without losing verifiability on-chain. The following paper will be a review of blockchain privacy technologies developed in 2020-2025, which will involve the history of zero-knowledge proofs and homomorphic encryption development at the cryptographic primitive level, as well as share new developments such as secure multi-party computation. It points out advances in recursive proof systems, distributed proof generation architectures and scalable multi-party computing systems to overcome bottlenecks in performance. There is a trade-off between privacy, system performance, regulatory compliance, and decentralization in a comparative analysis of technology integration in both public and permissioned chains. Lastly, research directions in the future are suggested in order to overcome issues associated with low proof efficiency, regulatory compliance problems, and migration of post-quantum cryptography. The review offers both theoretical and technical sources on how to develop trusted blockchain infrastructure that would strike the right balance between compliance, high-performance, and data sovereignty.

1 Introduction

Blockchain technology combines cryptography and distributed consensus to establish a distributed ledger technology in an un-trusted context. Its value essence is to have decentralized value exchange. This trust tool is very conditional that the highest level of data transparency: to avoid the possibility of spending twice and maintain consistency of state, each verification node should have access to the entire transaction history. Nonetheless, this

* Corresponding author's email: xianyujia666@outlook.com

openness in comparison with strict user and business requirements of data privacy is what the privacy paradox of blockchain consists of blockchain.

Earlier blockchain used a pseudonym system to grant some form of anonymity but it has over time, failed to stand against the advanced technologies of transaction graph analysis and network correlation. Full disclosure of data in a case of sensitive data application, including finance, healthcare, and supply chain, is a direct violation of the bottom line of commercial secrets and can be an infringement of a data protection law, GDPR. Thus, the challenge of ensuring successful privacy protection and keeping blockchain verifiable has turned into the central concern of the range of user needs to move this technology out of the proof-of-concept stage and implement it in large-scale commercial use.

To systematically address this contradiction, privacy protection research has shifted from simple obfuscation to provably secure cryptographic schemes. From a technical perspective, the main approach is to advance in two aspects: at the lower-level encryption primitives level, such as homomorphic encryption and zero-knowledge proofs, which focus on ciphertext computation and verification; at the protocol-level privacy mechanism level, such as secure multi-party computation (SMPC) and differential privacy, ring signatures, and mixing protocols, which solve the problems of data collaboration and identity concealment in a distributed environment. These technologies aim to reconstruct the privacy boundary without compromising the auditability of the system [1].

The proposed article will provide a systematic review of the recent progress in blockchain privacy protection technologies between 2020 and 2025. The study will initially examine the developmental trajectories of foundational technologies like zero-knowledge proof and homomorphic encryption, and pay attention to elaborating new findings on recursive proof and effective computing infrastructure. The paper will then analyze the combined usage of the protocol-level mechanisms across the established architectures, and compare the trade-off relationships across existing technologies based on performance overheads, compliance, and degree of decentralization. Lastly, the paper will conclude with a summarization of the current issues that the technologies have with regard to the proof efficiency, regulatory flexibility and post-quantum cryptography migration, and prospective research perspectives.

2 Classification of blockchain privacy protection technologies and analysis of mainstream technologies

2.1 Classification framework of blockchain privacy protection technologies

In order to have a better insight into these technologies, in the light of technicality and implementation process, the paper divides the privacy protection technologies into four, including encrypting technology, protocol designing technology, obfuscation technology and identity authentication. Figure 1 provides the system of classification. This classification aids in the understanding of the technical peculiarities and the situations of use [1].

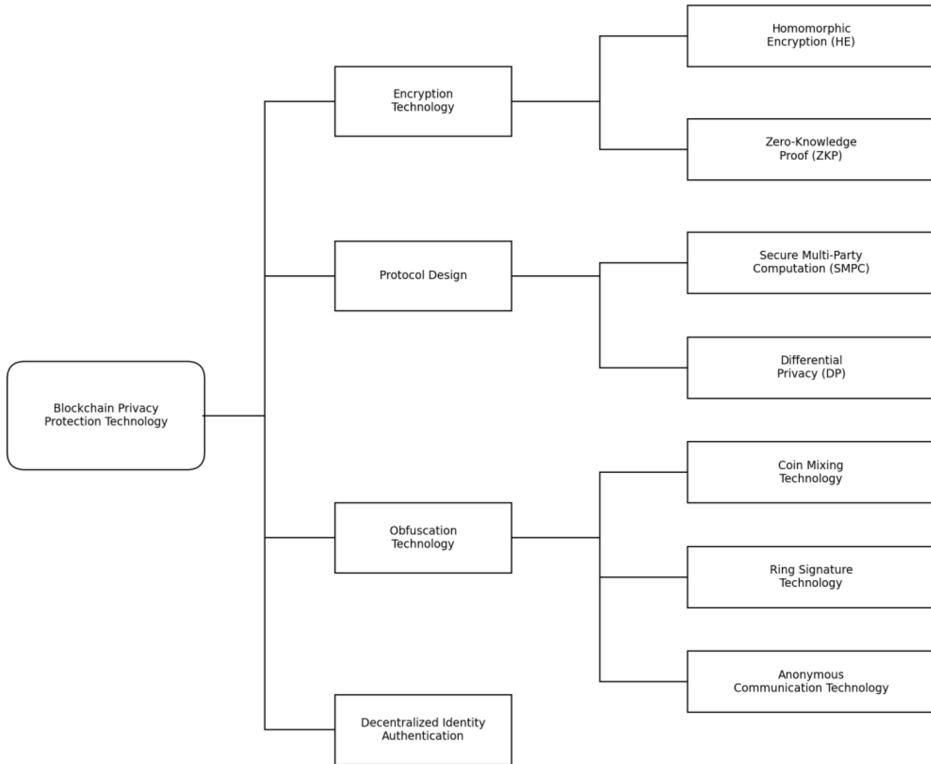


Fig. 1. System of Classifying Blockchain Privacy Protection Technology [1].

2.2 Analysis of mainstream privacy protection technologies

You are free to use colour illustrations for the online version of the proceedings, but any print version will be printed in black and white unless special arrangements have been made with the conference organiser. Please check whether this is the case. If the print version will be black and white only, you should check your figure captions carefully and remove any reference to colour in the illustration and text. In addition, some colour figures will degrade or suffer loss of information when converted to black and white, and this should be considered when preparing them.

2.2.1 Homomorphic Encryption (HE)

Homomorphic encryption is a special encryption technique. It allows the direct calculation of encrypted values, and the contradiction of the safeguarding of privacy and data use is eliminated. This is the additive homomorphism property, $Enc(m_1) \oplus Enc(m_2) = Enc(m_1 + m_2)$. It denotes that the outcome of the ciphertext operation is identical to the encrypted outcome of the plaintext operation. Nevertheless, the existing homomorphic encryption, especially homomorphic encryption with full homomorphic, is computationally quite complicated and consumes a substantial amount of computing power and storage volume. This implies that encryption, computation and decryption are not very fast and efficient. Hence, the homomorphic encryption has a high power that enhances the security of data. Nevertheless, in order to be extensively utilized in blockchain systems, which imply high speed, it continues to have profound performance problems.

Homomorphic encryption allows immediate computation on the ciphertext side, whereas multi-party computation, which is secure, allows joint computation among two or more parties without having those parties share their inputs. Bourse et al. (2020) presented a better variant of TFHE to solve the issue of computational efficiency of the fully homomorphic encryption. They used optimal programmable bootstrapping to obtain programmable ciphertext comparison and table lookup at logic speeds of milliseconds without having to decrypt ciphertext, an important step in implementing it to the challenges of on-chain confidential auctions [2]. The Falcon framework suggested by Wagh et al. (2021) is adapted to the optimal environment of all parties being truthful in the context of the SMPC field. The performance of the inference communication spoken about the deep learning models by order of magnitude has made it possible to carry out privacy-preserving AI computations on the blockchain by improving the calculation protocol of nonlinear functions [3]. Moreover, the Cheetah framework made by Huang et al. (2022) is concerned with the optimization of inference of neural networks in the two-party computing setting. It minimizes the cost of communication of the mobile wallets that are included in privacy computing by simplifying the protocol of a linear layer [4].

2.2.2 Zero-Knowledge Proof (ZKP)

The technology of zero-knowledge proof hides the content and the identity and lets the prover prove the truthfulness of the statement to the verifier, but does not disclose any particular information. The model behind this is that each prover of a particular statement provides the proof as opposed to the actual data themselves, and the verifier checks the validity of such a proof to be convinced that the statement is indeed true (Figure 2).

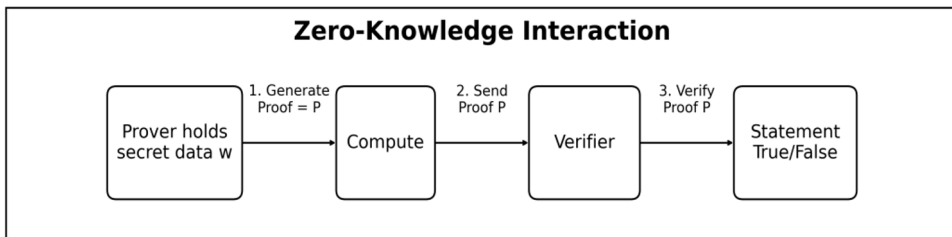


Fig. 2. Zero-knowledge proof basic model [5].

This technology strikes a balance between information verification and privacy protection, and has been applied to privacy-encrypted cryptocurrencies such as Zcash and ZK-Rollups scalability solutions. Zero-knowledge proof is very effective in protecting privacy. However, generating the proof requires a lot of computing resources. The generated proof data is also usually quite large. This results in high computational costs.

Zero-knowledge proof technology has undergone a significant transformation from theoretical exploration to practical application. Zerocash (Zcash), proposed by Ben-Sasson et al., was the first to apply zk-SNARKs on a large scale to blockchain privacy protection, achieving complete concealment of transaction amounts and participating parties [5]. However, its reliance on trusted setups and high computational overhead has become a major bottleneck for its application. As shown in Figure 3, the technological evolution of zero-knowledge proofs follows a clear research trajectory, undergoing a process of evolution from reliance on trusted setups to complete transparency.

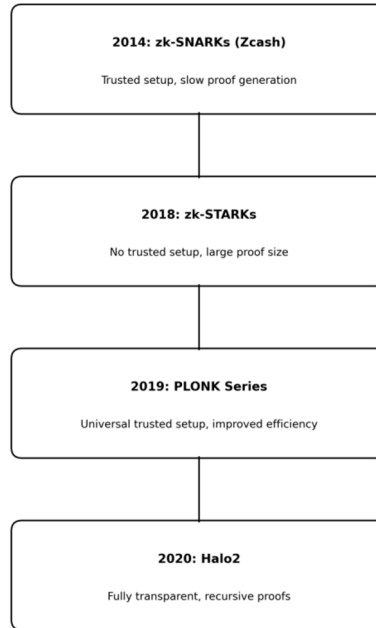


Fig. 3. The evolution path of zero-knowledge proof technology [5, 6].

The Zerocash scheme proposed by Ben-Sasson et al. in 2014 showed in their experimental data that proof generation takes approximately 40 seconds, with a proof size of 288 bytes [5]. Full elimination of a trusted setup, however, came with zk-STARKs, introduced in late 2018 in the work of Ben-Sasson et al., which had a high cost in the increased proof size of around 100 KB. Relative Homogeneous analysis indicates a trade-off between the size of proof and the time of verification (Table 1) [6]. The history of technological evolution is a mathematical optimization process, the complexities of both proof systems and proof systems themselves are reduced (zk-SNARKs, zk-STARKs and Nova). In particular, the Nova system invented by Zhou et al. decreases the time of generating proof by a factor of one due to recursive proof constructions [7]. This is a new recursive type of proof, which improves scalability, and offers a new direction to blockchain privacy protection.

Table 1. Comparison of experimental results [5-7]

Technical Solution	Proof generation time	Proof Size	Circuit/Constraint Scale	Trusted Setup Required or Not
zk-SNARKs	40-60s	288bytes	800MB-1.2GB	need
zk-STARKs	fast	100-200KB	2-4GB	unnecessary
Nova	5-8s	10KB/288B	300-500MB	need

2.2.3 Secure Multi-Party Computation (SMPC)

Secure Multi-Party Computation (SMPC) is a combination of secret sharing and distributed computing, allowing distrustful parties to engage in the joint computation of results but without a trusted third party or any disclosure of private inputs. SMPC is an off-chain secret engine in the blockchain system that balances Zero-Knowledge Proofs (ZKP) (added to prove integrity of data) and Homomorphic Encryption (HE) (added to work with encrypted data at

rest). SMPC is a network based scheme that achieves dynamic collaborative privacy and is much more flexible to complex logic at the cost of more communication. With SMPC outcomes anchored on-chain using ZKP, the system establishes an excellent trade-off between privacy and verifiability effectively addressing the blockchain privacy paradox by providing confidential execution and state transition verifiability.

Experimentally, it was shown that although protecting the privacy of smart contracts is at least 50-100 times more expensive than plaintext execution, the Hawk framework of Kosba et al. (2016) partially achieves this protection [8]. It was observed that this overhead is majorly based on how many communication rounds can be made between various parties and this may be a deployment challenge in a consortium blockchain setting. E.g., the issues that can be encountered in real-world applications are: in unreliable network conditions, as many as 70 percent of performance can be lost; every round of communication grows quadratically in terms of the participant count; computation interruptions cannot be restarted. Later studies have aimed at the optimization of communication complexity and the enhancement of fault tolerance of the network to solve these problems. The model put forward by Wagh et al. (2021) is fine-tuned to a three-party honest majority setting [3]. Falcon, by enhancing the computation protocols of non-linear functionalities like ReLU, does away with the costly overheads of obfuscated circuit transformations, enhancing the efficiency of inference of deep learning models tens of times. The comparison of the architecture optimization is presented in Table 2.

Table 2. Architecture Optimization Comparison [3, 9].

Optimization dimensions	Hawk	Falcon
Communication complexity	$O(n^2)$	$O(n)$
Network fault tolerance	low	mid
Main technologies	obfuscation circuit	Secret sharing
Applicable scenarios	General-purpose smart contracts	Privacy-preserving AI training/inference

2.2.4 Differential Privacy (DP)

Differential privacy protects individual information in statistical data by adding random noise, preventing attackers from inferring sensitive information from publicly available data. It can be used in blockchain for publishing public data such as disease statistics, but the added noise reduces statistical accuracy, requiring a trade-off between privacy and data usability. Dwork et al. (2006) laid the mathematical foundation, and subsequent studies determined a reasonable range for the privacy budget ϵ : a loss of 25% in data utility occurs when $\epsilon = 0.1$, the loss decreases to 5% when $\epsilon = 1.0$, and privacy protection is significantly reduced when $\epsilon = 10$. In blockchain data publishing scenarios, a value of ϵ between 0.5 and 2.0 can balance privacy and utility [9]. Zhang and Li proposed a dynamic privacy budget allocation mechanism, with the core algorithm being $M(D) = f(D) + \text{Noise}(\Delta f / \epsilon(t))$, which transforms a static budget into a dynamically adjusted one, thereby improving data availability [10]. Table 3 shows the optimization results of differential privacy parameters.

Table 3. Optimization results of differential privacy parameters.

Privacy budget ϵ	Data utility loss	Applicable scenarios
$\epsilon < 0.1$	~25%	Applicable scenarios
$0.1 \leq \epsilon < 1.0$	~15%	General sensitive data
$\epsilon \geq 1.0$	<5%	Internal data analysis

3 Technology comparison and application analysis

3.1 A comprehensive comparison of privacy protection technologies

To clearly illustrate the characteristics of different privacy protection technologies, this paper will compare them based on factors such as reliance on third parties, the ability to conceal transaction content and addresses, privacy protection effectiveness, security assumptions, performance overhead, and typical applications (Table 4), and Figure 4 shows the reference architecture for blockchain privacy protection integrating ZKP, SMPC, and HE.

Table 4. Comprehensive Comparison of Blockchain Privacy Protection Technologies.

Technology Category	Core cryptographic principles	Privacy objectives	Performance overhead (computation/storage)	dependence	Typical applications/Representative solutions
Homomorphic encryption	Ciphertext computation	Hidden transaction details	Extremely high (computational cost)	No	On-chain encrypted data analysis
Zero-Knowledge Proof	Proof/Verification	Hidden transaction details and addresses	High (Proof generation overhead)	No (SNARKs may require a trusted setup)	Zcash, ZK-Rollups Zcash, ZK-Rollups
SMPC	Secret sharing / Obfuscated circuits	Hidden calculation input	High (high communication overhead)	No (depends on the honesty of multiple parties)	Joint Risk Control
Differential Privacy (DP)	Random noise	Protecting individual information in statistical data	medi	No	On-chain aggregated data released

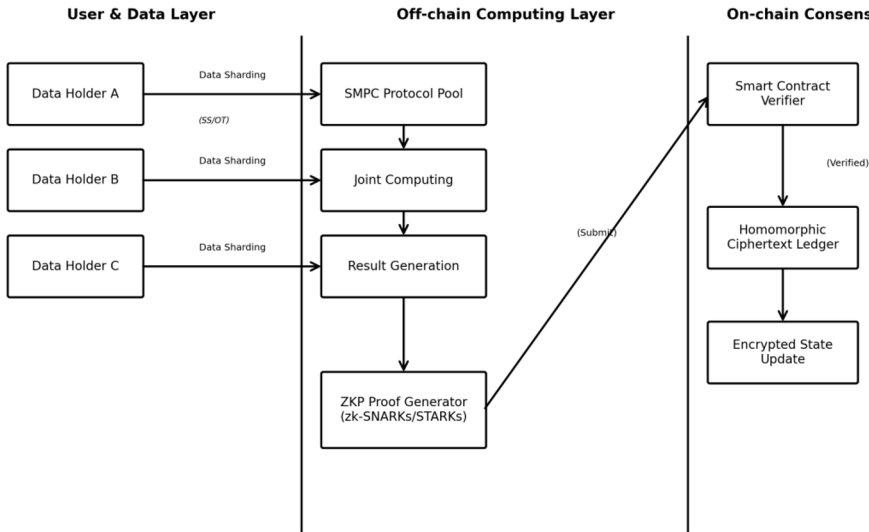


Fig. 4. A Reference Architecture for Blockchain Privacy Protection Integrating ZKP, SMPC, and HE. (Picture credit: Original)

3.2 Analysis of applicable scenarios for different technologies

3.2.1 Non-permissioned blockchain (public blockchain) scenarios

Everyone is free to become a member of the network in a public blockchain environment, making the very essence of privacy defense the deidentifying of the relationship between the on-chain addresses and real-life identities, as well as hiding the transaction quantity and contract logic. A typical example of privacy protection on public blockchains is Tornado Cash. It combines zk-SNARKs and Merkle trees in solving transaction traceability. In case users deposit money, a hash of the credential is created randomly and it is stored in the Merkle tree. In dividing out funds, the user provides a zero-knowledge argument to the degree that he/she holds the leaf node credential, but never discloses which particular node, therefore attaining non-interactive coin mixing. Following this, more general solutions have also been developed in programmable privacy like in the case of the Zexe system, extending protection against privacy even to the level of computation itself [11]. Zexe suggests a decentralized approach to personal computing (according to Zexe) in which users execute arbitrarily complex computations off-chain and present proofs of the validity of the results of the computations (using zk-SNARKs) to the blockchain. This model of off-chain implementation but on-chain verification not just conceals the input data, but the logic of the functions being called, accomplishing the same functional properties as the so-called private smart contracts, and forming the basis of developing privacy-preserving applications of more complex design (such as a decentralized exchange or user-sovereign assets).

Aztec Network is developing an Ethereum-embracing privatized Layer 2 that is founded on the PLONK proof system. The MIT model adopts the UTXO model to store encrypted state and recursive proofs to reduce all transactions into a single validity proof presented to the mainnet and handles the balance leakage problem of account-based models and is programmable privacy-wise[9].

3.2.2 *Permissioned blockchain (consortium blockchain/private blockchain) scenarios*

Consortium blockchains are based on isolating data and sharing values between institutions. Hyperledger Fabric manages identity and data privacy with the help of Identity Mixer (Idemix) and private data collections [12]. Idemix would utilize the zero-knowledge proof to conduct anonymous identity authentication, where members can verify that they have certain characteristics but not identify themselves. The sensitive information stored as private data sets is synchronized between the authorized nodes over the Gossip protocol but it is only the hash that can be added to a blockchain to accomplish physical data insularity.

The FATE model used by WeBank shows how blockchain and SMPC/HE have been strongly combined with homomorphic encryption and multi-party computation protocols to enable collaborative modeling using local data without the required data being transferred out of its physical location in cross-institutional credit risk control. Such usability of data in the absence of a visibility model has been proven in domains like medical data collaboration and data sharing in the government.

As a coordination layer, blockchain holds authorization logs, model version hash, and contribution proofs, addressing the issue of data siloing and complying with requirements.

4 Challenges and future research directions

4.1 The main challenges currently faced

4.1.1 *Performance and scalability overhead*

Privacy protection often comes at the cost of performance. The generation process of zero-knowledge proofs involves a large number of multi-scalar multiplications and number-theoretic transforms, requiring extremely high computational resources. For example, generating a complex zkEVM circuit proof may consume substantial memory and computational resources, which severely limits its application in high-frequency trading scenarios [6]. The computational overhead of fully homomorphic encryption is even thousands of times higher than that of plaintext computation, and is currently limited to simple logical operations. Partially Homomorphic Encryption (PHE) and Leveled Fully Homomorphic Encryption (LHE) have achieved a certain degree of practicality for specific simple operations, laying the groundwork for future performance optimization.

4.1.2 *Quantum computing threats and the costs of migrating to post-quantum cryptography*

Many existing privacy protection schemes (such as those based on elliptic curve zk-SNARKs and ring signatures) rely on the difficulty of the discrete logarithm problem or the integer factorization problem. The emergence of Shor's algorithm makes these schemes vulnerable to future quantum computers. If a timely migration to quantum-resistant algorithms is not implemented, existing on-chain private historical data will face the risk of being decrypted in the future.

4.1.3 *balance between regulatory compliance and privacy*

Privacy technologies in the financial application should balance user data privacy and legal requirements like Anti-Money Laundering (AML) and Know Your Customer (KYC).

Regulatory compliance is usually in conflict with pure anonymity. This is dealt with by ZKPs via Selective Disclosure. As an example, the application of the so-called Viewing Keys means that the user gets privacy with the rest of the world but the authorized regulators or auditors are allowed to access the detailed information about the transactions. By this, institutions can demonstrate that they meet the mandates of such frameworks as MiCA or AML6 without disclosing delicate trade secrets.

4.2 Future research directions and prospects

Further studies will still be done to reduce the overhead of ZKP proving. Recursive proofs and folding schemes include Nova and SuperNova will continue to be of primary interest, because this provides the possibility of constant-size verification and the capability to support proofs of unlimited computational depth. At the same time, hardware acceleration, as well as distributed proving network, will play a crucial role in enhancing system throughput. Moreover, we shall see the advent of new popular areas of research such as Cross-chain Privacy and Privacy as a Service (PaaS): Cross-chain privacy is designed to break blockchain silos, such that interoperability of high-performance privacy technology may happen smoothly across heterogeneous networks, whereas PaaS models provide more liberal points of entry to enterprises who may implement new blockchain-based privacy infrastructure and technologies.

As quantum-resistant blockchain privacy technology, the move to quantum-resistant algorithms with the publication of post-quantum cryptography standards by NIST in 2024 has become urgent. Lattice-based cryptography schemes of zero-knowledge proof and ring signature will be a hot topic in research. As an illustration, the lattice-based linkable ring signature scheme introduced by Liu et al. tries to minimize the size of signatures and the cost of verification and quantum resistance to be more accommodating of the storage capacity of blockchain [6]. Future studies will be done in order to maximize the efficiency of these post-quantum primitives and come up with higher-level transition protocols so that existing blockchain systems can be safely upgraded.

5 Conclusion

Protection of privacy by blockchain is an imperative on a transition stage on the way towards the theories to the reality. Homomorphic encryption, zero-knowledge proofs, and SMPC have attained breakthroughs in both performance and usability, and can be used to deal with the privacy paradox of blockchain. Recursive proof systems such as Nova greatly lower the cost of verification, distributed proof systems such as Pianist solve the bottlenecks of large-scale computations, and efficient proving systems based on SMPC frameworks such as Falcon and MVOC can scale up complex AI computations that preserve privacy. Nevertheless, all the challenges, such as computational efficiency, regulatory compliance, and quantum security, limit the widespread use of such technologies. The next generation will not concern personal technologies that are still isolated but rather systemized engineering that comprises optimization of underlying cryptographic primitives, specialized hardware acceleration, protocol layer compliance functionality design and integration across layers architecture. Through the creation of an advanced regulatory framework between transparency and privacy, the following generation of blockchain systems will be able to better fulfill the deep-rooted interests of the digital economy, ensuring free movement of value without giving up on the data-sovereignty.

References

1. P.L. Tan, T. Xu, S.J. Yang, Z.H. Tao, Review of research on blockchain privacy protection technologies. *Appl. Res. Comput.* **41**, 2261–2269 (2024)
2. F. Bourse, O. Sanders, J. Traoré, Improved secure integer comparison via homomorphic encryption, in *Proceedings of the Cryptographers' Track at the RSA Conference (CT-RSA 2020)*, San Francisco, CA, USA, February 24–28 (2020), B391–416
3. S. Wagh, S. et al., Falcon: Honest-majority maliciously secure framework for private deep learning, in *Proceedings on Privacy Enhancing Technologies (PETs)*, Virtual Event, July 12–16 (2021), B188–208
4. Z. Huang, W. Lu, C. Hong, J. Ding, Cheetah: Lean and fast secure two-party deep neural network inference, in *31st USENIX Security Symposium (USENIX Security 22)*, Boston, MA, USA, August **10–12** (2022), B809–826
5. E. Ben-Sasson, A. Chiesa, M. Garman, M. Green, I. Miers, E. Tromer, M. Virza, Zerocash: Decentralized anonymous payments from Bitcoin, in *Proceedings of the 2014 IEEE Symposium on Security and Privacy (SP)*, Berkeley, CA, USA, May **18–21** (2014), B459–474
6. W. Liu, Z. Liu, K. Nguyen, G. Yang, Y. Yu, A lattice-based key-insulated and privacy-preserving signature scheme with publicly derived public key, in *Proceedings of the 25th European Symposium on Research in Computer Security (ESORICS 2020)*, Part II, Guildford, UK, September 14–18 (2020), B357–377
7. A. Kothapalli, S. Setty, I. Tzialla, Nova: Recursive zero-knowledge arguments from folding schemes, in *Advances in Cryptology – CRYPTO 2022*, Santa Barbara, CA, USA, August **15–19** (2022), B359–388
8. A. Kosba, A. Miller, E. Shi, Z. Wen, C. Papamanthou, Hawk: The blockchain model of cryptography and privacy-preserving smart contracts, in *Proceedings of the 2016 IEEE Symposium on Security and Privacy (SP)*, San Jose, CA, USA, May 22–26 (2016), B839–858
9. C. Dwork, Differential privacy, in *Proceedings of the 33rd International Colloquium on Automata, Languages and Programming (ICALP)*, Venice, Italy, July **10–14** (2006), pp. 1–12
10. J. Zhang, N. Li, H. Chen, Dynamic differential privacy for data streams. *IEEE Trans. Knowl. Data Eng.* **30**, 2314–2328 (2018).
11. S. Bowe, A. Chiesa, M. Green, I. Miers, P. Mishra, H. Wu, ZEXE: Enabling decentralized private computation, in *Proceedings of the 2020 IEEE Symposium on Security and Privacy (SP)*, San Francisco, CA, USA (2020), B947–964
12. C. Stamatellis, P. Papadopoulos, N. Pitropakis, S. Katsikas, W.J. Buchanan, A privacy-preserving healthcare framework using Hyperledger Fabric. *Sensors* **20**, 6587 (2020)
13. T. Liu, et al., Pianist: Scalable zkRollups via fully distributed zero-knowledge proofs, in *Proceedings of the 45th IEEE Symposium on Security and Privacy (SP)*, San Francisco, CA, USA, May 20–23 (2024), B1777–1793