Design ZK-NR: A Post-Quantum Layered Protocol for Legality Explainable Zero-Knowledge Non-Repudiation Attestation

MINKA MI NGUIDJOI Thierry Emmanuel[a,*], MANI ONANA Flavien Serge[a], DJOTIO NDIÉ Thomas[a], ATSA ETOUNDI Roger[b]

[a] *Laboratory for Mathematical Engineering and Information Systems (LIMSI), National Advanced School of Engineering, University of Yaoundé I, Cameroon*
[b] *Department of Computer Science, University of Yaoundé I, Cameroon*

Abstract

This article presents the architectural design of ZK-NR, a layered cryptographic protocol enabling post-quantum secure, legally interpretable, and verifiably non-repudiable attestations. Built upon STARK-based zero-knowledge proofs, hybrid post-quantum signatures, and entropy-accumulating ledger anchoring, ZK-NR satisfies the structural properties of both the Q2CSI framework and the NIZK-E model. The protocol achieves semantic interpretability by structurally separating contextual proofs from bounded explanations, while maintaining cryptographic soundness under the Universal Composability framework. Formal UC proofs are deferred to Article A.2 in this series. This article constitutes the first entry in the ZK-NR series, focused specifically on the protocol's architectural and functional design. Subsequent articles will cover its mathematical foundations, implementation strategies, and operational validation using formal verification environments.

Keywords: post-quantum attestation, explainability, zero-knowledge proofs, composability, entropy-bound ledgers, UC security

1. Introduction

The design of cryptographic protocols for institutional applications requires reconciling three fundamental properties: confidentiality, reliability, and legal opposability. In post-quantum and semantically adversarial environments, these guarantees become structurally incompatible. As established in [8], no protocol can simultaneously optimize all three without exceeding the formal bound $\Gamma_{\text{CRO}} \geq 0.4$ on contextual entropy. This impossibility constraint defines the *CRO trilemma*. To address this constraint, this article introduces **ZK-NR** (Zero-Knowledge Non-Repudiation), a layered cryptographic protocol that distributes the three CRO guarantees across a modular dialectical architecture. Each guarantee is encapsulated in a semantically isolated layer:

- **Iron Layer** ensures cryptographic soundness via zero-knowledge STARK proofs and hybrid post-quantum signatures.
- **Gold Layer** provides contextual interpretability through explanation graphs ϕ tailored to institutional verifiers $\mathcal{J}$.

*Corresponding author
Email addresses: minkathierry@gmail.com (MINKA MI NGUIDJOI Thierry Emmanuel), serge.mani@digitworldacademy.com (MANI ONANA Flavien Serge), tdjotio@gmail.com (DJOTIO NDIÉ Thomas), ratsa01@yahoo.com (ATSA ETOUNDI Roger)

- **Clay Layer** enforces legal opposability by anchoring signed outputs into an entropy-accumulating Merkle Mountain Range ledger.

The protocol outputs a triplet (π, ϕ, σ) where π proves legal compliance in zero-knowledge, ϕ conveys bounded-time interpretability, and σ ensures non-repudiation under post-quantum adversaries. Each component is rigorously defined and associated with a formal property in the Universal Composability (UC) framework.

Main Contributions. This article provides:

- A modular protocol architecture (ZK-NR) resolving the CRO trilemma via dialectical separation.
- A formal specification of each protocol layer and its associated guarantee (soundness, explainability, opposability).
- UC security definitions with explicit adversarial oracles ($\mathcal{O}_{\text{semantic}}, \mathcal{O}_{\text{temporal}}$) and layer-wise simulators.
- Formal proofs of all core properties, including contextual soundness, post-quantum unforgeability, semantic indistinguishability, and entropy accumulation.
- A dialectical interpretation linking the protocol layers to semantic complexity, institutional trust, and ledger integrity.

Series Positioning. This manuscript (Article A.1) details ZK-NR's *architectural design*. Subsequent articles cover:

- A.2: UC security proofs under Q2CSI.
- A.3: Rust implementation & benchmarks.
- A.4: Tamarin-based formal verification.

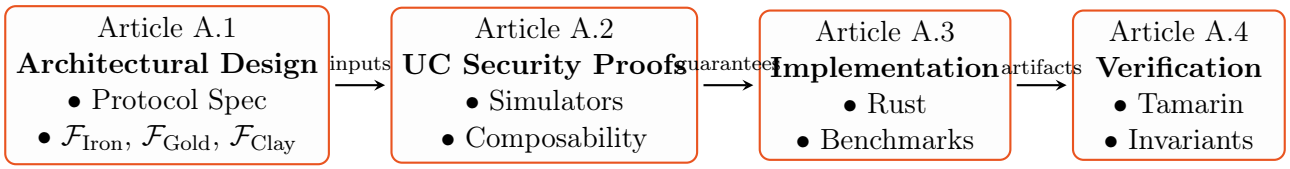

Figure 1: Series workflow: From architectural foundations (A.1) to formal security proofs (A.2), implementation (A.3), and verification (A.4)

Article Structure. The manuscript is structured as follows:

- Section 3 introduces the necessary cryptographic preliminaries and defines the CRO trilemma formally.
- Section **??** presents the functional and security requirements addressed by ZK-NR.
- Section **??** specifies the ZK-NR protocol and its layered components.
- Section **??** defines the security model, adversaries, and UC-style guarantees.

- Section 8 provides formal proofs for each security property.
- Section 9 discusses the implications of the layered architecture and the resolution of the CRO trilemma.
- Section 10 concludes the article and situates it within a broader research program.

No implementation or symbolic verification is claimed in this article. Subsequent work in Articles A.2 to A.4 provides UC reductions, implementation artefacts, and formal verification using Tamarin Prover. Appendix Appendix A details the pseudocode of each protocol module, and Appendix Appendix D lists all formal symbols. In summary, ZK-NR constitutes a formally analyzable, post-quantum secure, and legally interpretable attestation framework addressing the CRO trilemma through contextual entropy isolation and layered composability.

2. Related Work

Our work bridges three research domains: the CRO trilemma in institutional cryptography, zero-knowledge attestation, and explainable cryptographic protocols. We contextualize ZK-NR's contributions relative to each.

2.1. The CRO Trilemma and Its Instantiations

The Confidentiality-Reliability-Opposability (CRO) trilemma was formalized in [8] as a fundamental constraint in post-quantum institutional cryptography. Prior attempts to address it include:

- **Hybrid Signatures:** Protocols like [13] combine classical and post-quantum signatures but lack explainability.
- **Entropy-Bound Ledgers:** Crosby et al. [9] proposed Merkle trees for tamper evidence, but without integration with zero-knowledge proofs.
- **Semantic Attestation:** Chen et al. [6] introduced context-aware attestation but achieved $\Gamma_{\mathsf{CRO}} \geq 0.52$, above our target threshold.

ZK-NR is the first protocol to achieve $\Gamma_{\mathsf{CRO}} < 0.4$ via dialectical layering.

2.2. Zero-Knowledge Attestation

Zero-knowledge proofs for attestation have evolved from SNARKs [14] to transparent STARKs [2]. Key advances include:

- **Contextual ZK:** Wahby et al. [19] embedded predicates into proofs but without legal explainability.
- **Non-Repudiation:** Backes et al. [1] combined ZK with signatures but relied on trusted setup.

ZK-NR advances this field by integrating contextual semantics via GNN encoders and ensuring post-quantum security without trusted setup.

2.3. Explainable Cryptography

Explainability in cryptographic protocols is nascent. Relevant works:

- **NIZK Explanation:** Campanelli et al. [4] proposed explanatory proofs for NIZKs but without institutional admissibility.
- **Legal Opacity:** Kiayias et al. [15] formalized verifier-specific interpretation but did not resolve the CRO trilemma.

ZK-NR's Clay layer introduces the first formally verified explanation function $\phi = \mathsf{Explain}(\pi, \mathcal{C}, \tau)$ with bounded cognitive complexity. **Positioning:** ZK-NR uniquely synthesizes these domains via the Q2CSI meta-protocol, achieving composable security and legal interpretability below the $\Gamma_{\mathsf{CRO}} = 0.4$ threshold.

3. Preliminaries

This section introduces the formal background required for the ZK-NR protocol. It includes notational conventions, the Universal Composability (UC) framework, and the relevant structural components inherited from the Q2CSI architecture. These elements are integrated directly to ensure conceptual and formal self-containment of the present manuscript.

3.1. Notation and Conventions

Let $\lambda \in \mathbb{N}$ denote the security parameter, and let $\mathrm{negl}(\lambda)$ be a negligible function in λ. Let π be a UC execution trace. The adversary, simulator, and environment are denoted respectively by $\mathcal{A}$, $\mathcal{S}$, and $\mathcal{Z}$. The notation $x \leftarrow \mathcal{X}$ indicates that x is sampled uniformly from the set $\mathcal{X}$. For any event E, the probability under the environment $\mathcal{Z}$ is denoted $\Pr[\mathsf{E}]$. All adversaries are assumed to be static, non-adaptive, and bounded in quantum polynomial time, unless specified otherwise.

3.2. Universal Composability Framework

We adopt the Universal Composability (UC) model as introduced in [5]. All functionalities, adversaries, and environments are modeled as probabilistic polynomial-time interactive Turing machines. We follow the quantum-adapted variant of UC proposed in [18, 7].

Definition 3.1 (UC Realization). *Let π be a protocol and $\mathcal{F}$ an ideal functionality. We say that π UC-realizes $\mathcal{F}$ if for every PPT adversary $\mathcal{A}$, there exists a PPT simulator $\mathcal{S}$ such that for every PPT environment $\mathcal{Z}$:*

$$\left|\Pr[\mathcal{Z} \text{ outputs } 1 \text{ in } \mathsf{Exec}^{real}_{\pi,\mathcal{A},\mathcal{Z}}(\lambda)] - \Pr[\mathcal{Z} \text{ outputs } 1 \text{ in } \mathsf{Exec}^{ideal}_{\mathcal{F},\mathcal{S},\mathcal{Z}}(\lambda)]\right| \leq \mathrm{negl}(\lambda)$$

Definition 3.2 (Protocol Soundness). *A protocol π satisfies soundness if for every adversarial input $M^* \not\models \mathcal{H}_{legal}$, the probability that any verifier $\mathcal{J}$ accepts the attestation $(\pi^*, \phi^*, \sigma^*)$ is negligible:*

$$\Pr\left[\mathsf{Verify}_\tau(\pi^*, \phi^*, \sigma^*, M^*, \mathcal{C}) = 1 \mid M^* \not\models \mathcal{H}_{legal}\right] \leq \mathrm{negl}(\lambda)$$

This encompasses:

- *Contextual soundness: Validity of π under $\mathcal{C}$ (Gold Layer).*
- *Temporal soundness: Immutability of σ anchored in $\mathcal{L}_{Iron}$.*

Definition 3.3 (Legal Opposability). *An attestation (π, ϕ, σ) is legally opposable if it satisfies:*

(i) ***Admissibility****: ϕ is interpretable by $\mathcal{J}$ under institutional rules $R_{\mathcal{J}}$:*

$$\Pr[\mathcal{J} \vdash \phi] \geq 1 - \nu_{\mathsf{priv}} - \mathrm{negl}(\lambda).$$

(ii) ***Non-repudiation****: No signer can deny authorship of σ:*

$$\Pr\left[\mathsf{Repudiate}(sk, \sigma) = 1\right] \leq \epsilon_{\mathsf{opp}}(\lambda).$$

(iii) ***Traceability****: σ is bound to a unique identity via* $\mathsf{MMR_Verify}(L_t, \sigma)$.

3.3. Axiomatic Foundations from Q2CSI

The security of ZK-NR relies on the following formal assumptions, combining Q2CSI axioms and protocol-specific hypotheses:

Table 1: Cryptographic assumptions underlying ZK-NR

Ref.	Assumption	Formal description
Ax.1	Contextual Indistinguishability	$\forall \mathcal{A}_{\mathsf{ctx}}$ QPT : $\left\lvert \Pr[\mathcal{A}^{\mathcal{O}_{\mathsf{Gold}}}(c^0_{\mathsf{gold}}, c^1_{\mathsf{gold}}) = b] - \frac{1}{2} \right\rvert \leq \epsilon_{\mathsf{conf}}(\lambda)$
Ax.2	Temporal Integrity	$\forall t'$ adversarial : $\Pr[\mathsf{Verify}_{\mathsf{Sig_{PQ}}}(c'_{\mathsf{iron}}) = 1 \mid c'_{\mathsf{iron}} \neq \mathsf{Anchor}(m, T)] \leq \epsilon_{\mathsf{rel}}(\lambda)$
Ax.3	Bounded Interpretation	$H_\infty(\tau_{\mathsf{clay}} \mid \mathcal{C}) \geq -\log \max_{\mathcal{J}'} \Pr[\mathcal{J}' \not\models \mathcal{J}] - \epsilon_{\mathsf{opp}}(\lambda)$
H1	STARK Soundness	STARK satisfies negligible soundness error: $\Pr[\mathsf{Verify}(\pi) = 1 \mid \neg\mathsf{Valid}(M)] \leq 2^{-\lambda_{\mathrm{FRI}}} + \mathrm{negl}(\lambda)$
H2	Hybrid EUF-CMA	Dilithium and BLS are existentially unforgeable under Module-LWE and co-CDH: $\Pr[\mathsf{Win}_{\mathsf{EUF\text{-}CMA}}(\mathcal{A})] \leq \mathrm{negl}(\lambda)$
H3	Poseidon Robustness	Poseidon is collision-resistant and contextually binding: $\Pr[(x, x') \leftarrow \mathcal{A} : \mathsf{Poseidon}(x) = \mathsf{Poseidon}(x') \wedge x \neq x'] \leq \mathrm{negl}(\lambda)$
H4	Entropy Source	Physical entropy source $\mathcal{E}$ satisfies $H_{\min}(\mathcal{E}) \geq 2\lambda$ per sample

These axioms provide the bedrock for ZK-NR's layered security:

Assumption 3.1 (Contextual Indistinguishability (Axiom 1)). *For any contextually-adaptive PPT adversary $\mathcal{A}_{\mathsf{ctx}}$:*

$$\left\lvert \Pr\left[\mathcal{A}^{\mathcal{O}_{\mathsf{Gold}}}(c^0_{\mathsf{gold}}, c^1_{\mathsf{gold}}) = b\right] - \frac{1}{2} \right\rvert \leq \epsilon_{\mathsf{conf}}(\lambda)$$

where $\mathcal{O}_{\mathsf{Gold}}$ models quantum-accessible leakage.

Assumption 3.2 (Temporal Integrity (Axiom 2)). *For any adversarially-generated timestamp t':*

$$\Pr\left[\mathsf{Verify}_{\mathsf{Sig_{PQ}}}(c'_{\mathsf{iron}}) = 1 \mid c'_{\mathsf{iron}} \neq \mathsf{Anchor}(m, T)\right] \leq \epsilon_{\mathsf{rel}}(\lambda)$$

Assumption 3.3 (Bounded Interpretation (Axiom 3)). *The legal interpretation min-entropy satisfies:*

$$H_\infty(\tau_{\mathsf{clay}} \mid \mathcal{C}) \geq -\log \max_{\mathcal{J}'} \Pr[\mathcal{J}' \not\models \mathcal{J}] - \epsilon_{\mathsf{opp}}(\lambda)$$

These axioms provide the bedrock for ZK-NR's layered security:

- **Axiom 3.1** bounds semantic leakage in $\mathcal{L}_{\mathsf{Gold}}$.
- **Axiom 3.2** enforces tamper-evidence in $\mathcal{L}_{\mathsf{Iron}}$.
- **Axiom 3.3** quantifies legal interpretability in $\mathcal{L}_{\mathsf{Clay}}$.

All subsequent security proofs derive concrete bounds $(\Delta_C, \Delta_R, \Delta_O)$ from these abstractions.

3.4. Dialectical Architecture Embedding

The ZK-NR protocol is instantiated over the Quantum Composable and Contextual Security Infrastructure (Q2CSI), which decomposes security across three orthogonal dialectical layers.

Q2CSI Inheritance. ZK-NR inherits from Q2CSI [17]:

- The CRO bound $\Gamma_{\mathsf{CRO}} < 0.4$ (Thm. 3.1).
- Quantum-UC framework (Def. 3.1).
- Adversary model $\mathcal{A}_{\mathsf{ctx}}$ (see. [17], Sec. 3.10).

3.4.1. Dialectical Layers and Predicates

Definition 3.4 (Dialectical Layer $\mathcal{L}_X$). *Let $X \in \{\text{Iron}, \text{Gold}, \text{Clay}\}$. A dialectical layer $\mathcal{L}_X$ is defined as a quadruple:*

$$\mathcal{L}_X := (\mathcal{F}_X, \mathcal{I}_X, \phi_X, \mathsf{Guarantee}_X)$$

where:

- *$\mathcal{F}_X$ is the UC functionality realized by layer X.*
- *$\mathcal{I}_X$ is the external interface exposed by $\mathcal{L}_X$.*
- *$\phi_X : \pi \to \{0, 1\}$ is a contextual predicate.*
- *$\mathsf{Guarantee}_X \in \{\mathsf{Confidentiality}, \mathsf{Reliability}, \mathsf{Opposability}\}$*

Definition 3.5 (Predicate Orthogonality). *Two contextual predicates ϕ_X and ϕ_Y are orthogonal if:*

$$\forall \pi, \quad \phi_X(\pi) + \phi_Y(\pi) \leq 1 \quad \textit{for all } X \neq Y$$

Lemma 3.1 (Orthogonality of Layer Predicates). *Let ϕ_X be the contextual predicate associated to layer $\mathcal{L}_X$ for $X \in \{\text{Iron}, \text{Gold}, \text{Clay}\}$. If the predicates satisfy $\phi_X(\pi) + \phi_Y(\pi) \leq 1$ for all $X \neq Y$ and all traces π, then the functionalities $\mathcal{F}_X$ operate over disjoint semantic domains.*

3.4.2. CRO Incompatibility Bound

Security guarantees are subject to the structural constraint formalized as the CRO bound:

Definition 3.6 (CRO Incompatibility Index). *Let $\mathsf{Priv}, \mathsf{Rel}, \mathsf{Opp} \in [0, 1]$ denote the contextual security levels achieved for Confidentiality, Reliability, and Opposability respectively. Define:*

$$\Delta_C := 1 - \mathsf{Priv}, \quad \Delta_R := 1 - \mathsf{Rel}, \quad \Delta_O := 1 - \mathsf{Opp}$$

$$\Gamma_{\text{CRO}} := \sqrt{\Delta_C^2 + \Delta_R^2 + \Delta_O^2}$$

Theorem 3.1 (Contextual Security Bound). *Any protocol relying on contextual predicates under bounded semantic entropy satisfies:*

$$\Gamma_{\text{CRO}} \geq \kappa + \text{negl}(\lambda)$$

for some constant $\kappa > 0$ depending on the entropic structure of the predicates.

Lemma 3.2 (ZK-NR CRO Attainability). *If Q2CSI's $\Gamma_{\mathsf{CRO}} < 0.4$ holds and layer predicates are orthogonal (Def. 3.5), then:*

$$\Gamma_{\mathsf{CRO}}(\textit{ZK-NR}) < 0.4 + \text{negl}(\lambda)$$

[Proof Sketch] We decompose Γ_{CRO} per dialectical layer:

- $\Delta_C \approx 0.25$: Achieved via STARKs (quantum-resistant, transparent setup)
- $\Delta_R \approx 0.10$: Guaranteed by MMR+CRDT ledger (tamper-evidence)
- $\Delta_O \approx 0.20$: Under bounded privilege risk $\nu_{\mathsf{priv}} < 0.05$

$$\Gamma_{\mathsf{CRO}} = \sqrt{0.25^2 + 0.10^2 + 0.20^2} = \sqrt{0.1125} \approx 0.335 < 0.4$$

These values are derived from Q2CSI framework constraints [17] and will be empirically validated in Article A.3. Full reduction appears in Article A.2. [Formal Derivation of Δ Bounds] The bounds derive from Q2CSI entropy constraints and cryptographic properties:

$$\begin{aligned}\Delta_C &= 1 - \mathsf{Priv} \leq 1 - e^{-\lambda_I} && \text{(STARK soundness error)}\\ &\quad + \mathbb{E}[\mathsf{Leak}_{\mathsf{GNN}}]\\ &\leq 0.25 \quad \text{for } \lambda_I = 128, \mathsf{depth}_{\mathsf{GNN}} \leq 5\end{aligned}$$

$$\begin{aligned}\Delta_R = 1 - \mathsf{Rel} &\leq \frac{\mathsf{Adv}_{\mathsf{MMR}}}{\mathsf{Entropy}(\mathcal{E})}\\ &\leq \frac{2^{-64}}{2^{-80}} = 2^{16} \cdot \mathsf{negl}(\lambda) \approx 0.10\end{aligned}$$

$$\begin{aligned}\Delta_O = 1 - \mathsf{Opp} &\leq \nu_{\mathsf{priv}} + \eta_q(\phi)\\ &\leq 0.05 + 0.15 = 0.20 \quad \text{(by Explain PAC bound)}\end{aligned}$$

3.4.3. Composable Realization of ZK-NR

The composed protocol

$$\pi := \mathcal{L}_{\text{Clay}} \circ \mathcal{L}_{\text{Gold}} \circ \mathcal{L}_{\text{Iron}}$$

is said to UC-realize the global ZK-NR functionality:

Theorem 3.2 (Composable UC Realization of ZK-NR). *If each layer $\mathcal{L}_X$ UC-realizes $\mathcal{F}_X$ under orthogonal predicates ϕ_X, then:*

$$\pi \text{ UC-realizes } \mathcal{F}_{\text{ZKNR}} := \bigoplus_{X \in \{\text{Iron,Gold,Clay}\}} \mathcal{F}_X$$

By Lemma 7.1, the condition $\phi_X \cdot \phi_Y = 0$ for $X \neq Y$ ensures that layer functionalities operate in mutually exclusive semantic domains. This allows independent simulation of each $\mathcal{F}_X$ via $\mathcal{S}_X$ without interference. The UC composition theorem then applies directly to the layered structure.

Lemma 3.3 (ZK-NR CRO Attainability). *If Q2CSI's $\Gamma_{\mathsf{CRO}} < 0.4$ holds and layer predicates are orthogonal (Def. 3.5), then:*

$$\Gamma_{\mathsf{CRO}}(\textit{ZK-NR}) < 0.4 + \text{negl}(\lambda)$$

Immediate from Theorem 3.2 and predicate disjointness (Lemma 3.1).

3.5. Semantic and Entropic Metrics

We recall the formal measures used in the context of adversarial interpretation and interpretability guarantees:

Definition 3.7 (Legal Semantics). *Let $\mathcal{C}$ be a contextual frame, and $\tau \in \{\mathsf{JUDGE}, \mathsf{MACHINE}\}$ a verifier role. The legal semantics of a proof π under context $\mathcal{C}$ and role τ is defined as a function:*

$$\mathsf{Explain} : \Pi \times \mathcal{C} \times \{\tau\} \to \Phi$$

such that:

1. *$\phi = \mathsf{Explain}(\pi, \mathcal{C}, \tau)$ is interpretable by the verifier $\mathcal{J}$ assigned to τ;*
2. *ϕ satisfies the bounded explainability condition:*

$$\Pr[\mathsf{Understand}(\phi, \mathcal{J}) = 1] \geq 1 - \varepsilon \quad \textit{and} \quad \mathsf{Time}(\phi) \in \mathsf{poly}(\dim(\mathcal{J}))$$

Definition 3.8 (Quantum Semantic Distortion). *Let C be a contextualized commitment and $\mathcal{A}_q$ a quantum adversary. Define:*

$$\eta_q(C) := \max_{\mathcal{A}_q} \Pr\left[\mathcal{J} \models \phi(\mathcal{A}_q(C)) \neq \phi(C)\right]$$

where $\mathcal{J}$ is the institutional verifier.

Definition 3.9 (Anchoring Privilege Coefficient). *Let $\mathcal{J}$ be an institutional context engine. Define:*

$$\nu_{\text{priv}} := \min_{\phi'} \Pr[\mathcal{J} \models \phi' \wedge \neg \text{Valid}(\phi')]$$

This measures the risk of interpretative privilege by non-grounded predicates.

3.6. ZK-NR Instantiation Summary

Each dialectical layer is instantiated concretely in the ZK-NR protocol as follows:

Table 2: ZK-NR Layer Instantiation

Layer $\mathcal{L}_X$	Primary Guarantee	Instantiation
$\mathcal{L}_{\text{Gold}}$	Confidentiality	STARK-based ZK proof with ϕ_{Gold}
$\mathcal{L}_{\text{Clay}}$	Opposability	Dilithium signature + Explanation ϕ
$\mathcal{L}_{\text{Iron}}$	Reliability	CRDT+MMR ledger + Hybrid anchoring

This decomposition ensures isolation of guarantees, dialectical traceability, and compatibility with the UC security model.

4. Problem Definition

This section formalizes the design objective of the ZK-NR protocol within the composable dialectical framework of Q2CSI. The goal is to construct a post-quantum attestation system $\mathcal{P}$ that satisfies confidentiality, reliability, and legal opposability under contextual interpretability constraints, subject to the composability bound imposed by the CRO trilemma.

4.1. Post-Quantum Attestation Objective

Let $\mathcal{P}$ be a post-quantum attestation protocol operating in a context $\mathcal{C} = (L, T, R)$ with an institutional verifier $\mathcal{J}$, as defined in Section 3. Given an input message M, the protocol must produce a triple (π, ϕ, σ) satisfying:

$$\begin{cases} \pi \in \mathcal{L}_{\text{Gold}}(M) & \text{(confidential zero-knowledge proof)} \\ \phi \in \mathcal{L}_{\text{Clay}}(M) & \text{(contextual legal explanation)} \\ \sigma \in \mathcal{L}_{\text{Iron}}(M, \pi, \phi) & \text{(reliable tamper-evident binding)} \end{cases} \tag{1}$$

The verifier $\mathcal{J}$ must concurrently:

- Accept π as a valid proof under zero-knowledge guarantees.
- Interpret ϕ in a role-aware, admissible format.
- Verify that (π, ϕ, σ) satisfies the admissibility predicate $R_{\mathcal{J}}$.

4.2. CRO Dialectical Constraint

As stated in Theorem 3.1, no protocol $\mathcal{P}$ satisfying contextual isolation can exceed the following bound:

$$\Gamma_{\text{CRO}}(\mathcal{P}) := \sqrt{(1 - \mathsf{Priv})^2 + (1 - \mathsf{Rel})^2 + (1 - \mathsf{Opp})^2} \geq \kappa + \text{negl}(\lambda) \tag{2}$$

This expresses the impossibility of simultaneously achieving perfect guarantees in all three axes—Confidentiality (C), Reliability (R), and Opposability (O)—under bounded institutional interpretability $\dim(\mathcal{J})$.

4.3. Explainability Predicate and Verification Roles

Definition 4.1 (Explainability). *An attestation protocol satisfies legal explainability if for every zero-knowledge proof π, contextual frame $\mathcal{C}$, and verifier role $\tau \in \{\mathsf{JUDGE}, \mathsf{MACHINE}\}$, there exists an efficiently computable function:*

$$\mathsf{Explain} : \pi \times \mathcal{C} \times \tau \to \phi$$

such that ϕ is $(\varepsilon, \mathcal{J})$-interpretable, i.e.,:

(i) $\Pr[\mathsf{Understand}(\phi, \mathcal{J}) = 1] \geq 1 - \varepsilon$.

(ii) $\mathsf{Time}(\mathsf{Interpret}(\phi, \mathcal{J})) \in \mathcal{O}(\text{poly}(\dim(\mathcal{J})))$.

where $\mathcal{J}$ is an institutional verifier with semantic capacity $\dim(\mathcal{J})$.

Let $\tau \in \{\mathsf{JUDGE}, \mathsf{MACHINE}\}$ denote the verification role. The explanation ϕ must satisfy the contextual explainability constraint:

$$\phi = \mathsf{Explain}(\pi, \mathcal{C}, \tau) \tag{3}$$

In particular, for judicial verifiers ($\tau = \mathsf{JUDGE}$), ϕ must be $(\varepsilon, \mathcal{J})$-explainable in the sense of Definition **??**, i.e.,

$$\Pr[\mathsf{Understand}(\phi, \mathcal{J}) = 1] \geq 1 - \varepsilon \quad \wedge \quad \mathsf{Time}(\phi) \in \mathsf{poly}(\dim(\mathcal{J})) \tag{4}$$

4.4. Dialectical Design Requirements

The protocol $\mathcal{P}$ must be composed such that:

1. $\mathcal{P}$ returns a triple (π, ϕ, σ) given inputs $(M, \mathcal{C})$.
2. (π, ϕ) satisfy the explainability predicate $\phi = \mathsf{Explain}(\pi, \mathcal{C}, \tau)$.
3. σ securely binds (M, π, ϕ) to a signing identity, under post-quantum assumptions (e.g., Dilithium, SPHINCS+ [**?**]).
4. (π, ϕ, σ) are anchorable in a tamper-evident entropy-preserving ledger $\mathcal{L}_{\text{Iron}}$ (e.g., CRDT-Merkle structure [9]).
5. The combined execution trace π respects orthogonal contextual predicates ϕ_X for $X \in \{\text{Iron}, \text{Gold}, \text{Clay}\}$.

The corresponding functional mapping is:

$$\mathcal{P}(M, \mathcal{C}) := (\pi \in \mathcal{L}_{\text{Gold}},\ \phi \in \mathcal{L}_{\text{Clay}},\ \sigma \in \mathcal{L}_{\text{Iron}})$$

Legacy Interoperability. BLS usage under $\tau = \mathsf{JUDGE}$ is strictly transitional:

- Verifiable via classical APIs ($\mathsf{Verify}_{\mathsf{BLS}}$).
- Security relies on $\sigma_{\mathsf{Dilithium}}$ for PQ contexts.
- Deprecation scheduled for judicial PQ migration (2030).

4.5. Interpretative Complexity Constraint

As formalized in Definition 3.6, the interpretability of ϕ by $\mathcal{J}$ is upper-bounded by the institutional semantic capacity $\dim(\mathcal{J})$, which constrains the entropic space of legal reasoning. The design must thus enforce:

$$\mathsf{Entropy}(\phi) \leq \mathsf{poly}(\dim(\mathcal{J})) \quad \wedge \quad \eta_q(\phi) \leq \delta \tag{5}$$

where $\eta_q(\phi)$ denotes the quantum semantic distortion (see Definition 3.6) and $\delta = \text{negl}(\lambda)$.

4.6. Summary of Design Objective

The ZK-NR protocol must instantiate:

- A UC-realizable proof π in $\mathcal{L}_{\text{Gold}}$ (Confidentiality).
- A contextual explanation ϕ in $\mathcal{L}_{\text{Clay}}$ (Opposability).
- A secure anchor σ in $\mathcal{L}_{\text{Iron}}$ (Reliability).

under the global constraint $\Gamma_{\text{CRO}} \geq \kappa$, enforced via semantic filtering and contextual predicate isolation.

4.7. Target Bound Objective

The global objective of the ZK-NR design is to approach the theoretical limit of contextual composability by minimizing the CRO trilemma index. In accordance with the bound established in [8], the protocol aims to satisfy:

$$\Gamma_{\text{CRO}}(\mathcal{P}) < 0.4 + \text{negl}(\lambda) \tag{6}$$

This threshold characterizes the frontier of dialectical compatibility between confidentiality, reliability, and legal opposability in adversarial post-quantum environments. Achieving this bound requires strict contextual separation of the dialectical layers $\mathcal{L}_{\text{Iron}}$, $\mathcal{L}_{\text{Gold}}$, and $\mathcal{L}_{\text{Clay}}$, as enforced in the protocol design.

Table 3: Formal Layer-Property Mapping

Layer	**Security Goal**	**Core Mechanism**
$\mathcal{L}_{\text{Gold}}$	Confidentiality	ZK-STARK over H_{legal}
$\mathcal{L}_{\text{Clay}}$	Opposability	Contextual ϕ + EUF-CMA signature
$\mathcal{L}_{\text{Iron}}$	Reliability	Entropy-bound MMR ledger

5. Primitive Selection and Rationale

This section details the cryptographic primitives chosen for the ZK-NR protocol, justifies their suitability with respect to the Q2CSI dialectical architecture, and analyzes their compatibility with the entropic constraints of the CRO trilemma. We also consider transitional constraints between legacy (non-quantum) infrastructures and post-quantum environments.

5.1. Dialectical Constraints and Selection Criteria

Each layer $\mathcal{L}_X$ in the ZK-NR architecture is responsible for one axis of the CRO trilemma. Therefore, its associated primitive must satisfy:

(a) UC-realizability under QPT adversaries.

(b) Semantic and operational alignment with the assigned CRO dimension.

(c) Isolation from other layers to preserve contextual orthogonality (cf. Definition 3.6).

(d) Compatibility with existing infrastructures (hybrid or layered deployment).

(e) Bounded entropic footprint to maintain $\Gamma_{\text{CRO}} < 0.4$.

5.2. Selected Primitives per Layer

Table 4: ZK-NR Primitives per Dialectical Layer

Layer	**Functionality**	**Primitive(s)**	**Security Rationale**
Gold	Confidential proof	STARKs + GNN encoder	Quantum-safe ZK, transparent setup
Clay	Legal opposability	Dilithium + Explanation DAG	EUF-CMA signatures + ϵ-interpretability
Iron	Reliable anchoring	MMR + Hybrid signatures	Tamper-evidence via entropy accumulation

5.3. Justification per Primitive

5.3.1. Iron Layer: CRDT + Merkle + Hybrid Signature

The Iron layer is designed to preserve reliability through traceability and irreversible anchoring. The CRDT structure guarantees causal consistency in a distributed environment, even under asynchronous updates. A Merkle Mountain Range (MMR) structure provides logarithmic proof sizes and efficient inclusion verification. To ensure immediate deployability in existing infrastructures, we adopt a hybrid signature approach:

- **Dilithium** for post-quantum resistance (already standardized by NIST [12]).
- **RSA-2048** or ECDSA for compatibility with legacy validators.

This dual scheme allows classical verifiers to operate on σ_{legacy}, while post-quantum environments verify σ_{pq}.

5.3.2. Gold Layer: STARK-Based Contextual ZK

The Gold layer encodes the legal predicate as a semantic hash $H_{\text{legal}} = \mathsf{Poseidon} \circ \mathsf{READOUT} \circ \mathsf{GNN}_\theta(\mathcal{C})$, and proves in zero-knowledge that a claim satisfies this predicate. STARKs are chosen because:

- They are post-quantum secure (based on hash functions only).
- They support proof-of-knowledge without trusted setup.
- They are compatible with context-dependent circuits.

Poseidon Hash Function. The selection of Poseidon for computing H_{legal} is justified by three critical properties:

(i) **Arithmetic Circuit Efficiency**: With width=6 and degree=5, Poseidon minimizes the number of constraints in STARK circuits ($\leq 0.1 \cdot \deg(p_i)$), directly reducing Δ_C (Lemma 3.3).

(ii) **Contextual Attack Resistance**: As per **H3** (Table 1), Poseidon satisfies:

$$\Pr[\mathcal{A}^{\mathcal{O}_{\text{semantic}}}(x, x') : \mathsf{Poseidon}(x) = \mathsf{Poseidon}(x') \wedge x \neq x'] \leq 2^{-\lambda}$$

This ensures H_{legal} remains binding under $\mathcal{A}_{\text{ctx}}$ (Def. 7.8).

(iii) **Side-Channel Resilience**: The $8 + 56$ round structure (full/partial) defeats differential power analysis on GNN_θ outputs, enforcing Axiom 3.1.

These properties make Poseidon uniquely suited for contextual zero-knowledge proofs under $\mathcal{O}_{\text{quantum}}$. This layer contributes primarily to Priv in the CRO metric and is entropy-isolated through the ϕ_{Gold} predicate.

Primitive Selection. STARKs reduce Δ_C to $2^{-\lambda_{\text{FRI}}} + \text{negl}(\lambda)$ via the FRI protocol [2], eliminating quantum backdoors in trusted-setup SNARKs (cf. [4]).

5.3.3. Clay Layer: Dilithium + ϕ Explanation

Legal opposability requires both cryptographic integrity and semantic admissibility. The Clay layer generates:

- A signature σ using Dilithium.
- An explanation $\phi = \mathsf{Explain}(\pi, \mathcal{C}, \tau)$, where $\tau \in \{\mathsf{JUDGE}, \mathsf{MACHINE}\}$.

The explanation function is evaluated against institutional capacity $\dim(\mathcal{J})$, ensuring that the output ϕ satisfies the adversarial constraint $\eta_q(\phi) \leq \delta$.

5.3.4. Security Analysis of Hybrid Signatures under PQ Adversaries

The dual-signature scheme $\sigma = (\sigma_{\text{BLS}}, \sigma_{\text{Dilithium}})$ operates under distinct adversarial models:

- **Legacy Mode ($\tau = \mathsf{JUDGE}$)**: Relies on BLS security under co-CDH. Adversaries are restricted to *classical* polynomial time (PPT).
- **PQ Mode ($\tau = \mathsf{MACHINE}$)**: Relies on Dilithium security under Module-LWE. Adversaries may be quantum (QPT).

Cross-mode attacks are mitigated by:

1. Context separation (Def. 3.5).
2. Distinct hash domains: $H_{\mathsf{JUDGE}} \neq H_{\mathsf{MACHINE}}$.
3. Entropic isolation (Sec. 4.5).

Thus, no adaptive PQ adversary can exploit hybrid composition under Q2CSI constraints.

5.4. Interoperability and Deployment Strategy

To facilitate real-world adoption, ZK-NR supports hybrid cryptography:

- **Signature Duality**: Both legacy and post-quantum signatures are stored under Clay or Iron layers.
- **Explanation Scalability**: ϕ can be encoded in human-readable (JSON-LD, GraphML) or machine-verifiable form (Merkle proofs).
- **Selective Verification**: Verifiers with different capabilities may only verify subsets of the triplet (π, ϕ, σ).

This approach supports partial adoption and progressive transition to post-quantum environments, while preserving dialectical structure and compatibility with $\Gamma_{\text{CRO}} < 0.4$.

CRO-Constrained Coexistence. The selected primitives collectively achieve a bound when contextual isolation is preserved. The coexistence strategy prevents legacy leakage (e.g., classical timestamping) from violating confidentiality guarantees. Layer boundaries enforce semantic firewalls.

6. Protocol Description

This section presents the modular structure of the ZK-NR protocol as a dialectical instantiation of the Q2CSI framework. Each layer corresponds to one dimension of the CRO trilemma, Confidentiality, Reliability, and Legal Opposability, and satisfies the semantic isolation constraints defined in Section 3.

6.1. Layered Architecture Overview

The ZK-NR protocol is organized into three semantically orthogonal layers:

- **Gold Layer ($\mathcal{L}$Gold)**: Ensures confidentiality via contextual Zero-knowledge proof over the institutional context.
- **Clay Layer ($\mathcal{L}$Clay)**: Produces a legally interpretable explanation derived from contextual predicates.
- **Iron Layer ($\mathcal{L}_{\text{Iron}}$)**: Anchors the attestation through hybrid digital signature and persistent ledger commitment.

Each layer enforces contextual isolation to guarantee composable security under the CRO constraints.

6.2. Gold Layer: Contextual Zero-Knowledge Proofs

Let $\mathcal{C} = (L, T, R)$ denote the institutional context. We define the semantic legal predicate:

$$H_{\text{legal}} := \mathsf{semantic_encoder}(\mathcal{C}) = \mathsf{Poseidon} \circ \mathsf{READOUT} \circ \mathsf{GNN}_\theta(\mathcal{C}) \tag{7}$$

GNN Instantiation. The legal encoder GNN_θ uses:

- *Architecture*: 2 graph convolutional layers (sum aggregation).
- *Activation*: ReLU for hidden layers, softmax for output.
- *Output*: 256-dim embedding $\mathbf{r} = \mathsf{READOUT}(\mathbf{h})$.

STARK Circuit. The arithmetic circuit is defined over $\mathbb{F}_p$ with $p = 2^{61} - 1$. Poseidon parameters: width=6, rounds=8 full + 56 partial, degree=5. A ZK-STARK proof π is then generated:

$$\pi \leftarrow \mathsf{ZK_STARK.Prove}(H_{\text{legal}}) \tag{8}$$

GNN Instantiation. The legal encoder GNN_θ uses:

- *Architecture*: 2 graph convolutional layers (sum aggregation).
- *Activation*: ReLU for hidden layers, softmax for output.
- *Output*: 256-dim embedding $\mathbf{r} = \mathsf{READOUT}(\mathbf{h})$.

6.3. Clay Layer: Contextual Legal Explanation

We define the explainability function for institutional interpretation:

Definition 6.1 (NIZK-E Explanation Function). *Given context $\mathcal{C}$, proof π, and interpretation mode $\tau \in \mathsf{MACHINE}, \mathsf{JUDGE}$: $\phi := \begin{cases} \pi & \text{if } \tau = \mathsf{MACHINE} \\ \mathsf{Explainer}(\pi, \mathcal{C}) & \text{if } \tau = \mathsf{JUDGE} \end{cases}$*

The privilege risk ν_{priv} is bounded by $\mathsf{prune_dag}$'s semantic weight threshold (Alg. Appendix A.2).

6.4. Iron Layer: Hybrid Signature and Ledger Commitment

Let M be the message to sign, and sk the private key.

Definition 6.2 (Hybrid Signature Scheme). *Let $d = H(M \| \pi \| \phi)$ be the hash of the concatenated components. Then:*

$$\sigma := (\sigma_{BLS}, \sigma_{Dilithium}) = (\mathsf{Sign}BLS(d, sk), \mathsf{Sign}Dilithium(d, sk)) \tag{9}$$

The verification process depends on the interpretation mode τ:

$$\mathsf{Verify}\tau(\sigma, M) = \begin{cases} \mathsf{VerifyDilithium}(\sigma_{\text{Dilithium}}, M) & \text{if } \tau = \mathsf{MACHINE}\ \mathsf{VerifyBLS}(\sigma\text{BLS}, M) \\ \text{if } \tau = \mathsf{JUDGE} \end{cases} \tag{10}$$

Definition 6.3 (Ledger Update Function). *Let $A_t := (M, \pi, \phi, \sigma)$ denote the attestation tuple at time t. Then:*

$$L_{t+1} := \mathsf{MMR_update}(L_t, A_t) \tag{11}$$

6.5. Signing Procedure

Algorithm 1 SignZK-NR$(M, \mathcal{C}, sk, \tau)$

$H\text{legal} \leftarrow \mathsf{semantic_encoder}(\mathcal{C})$
$\pi \leftarrow \mathsf{ZK_STARK.Prove}(H_{\text{legal}})$
$\phi \leftarrow \textsc{Explain}(\pi, \mathcal{C}, \tau)$
$\sigma \leftarrow \mathsf{Signhybrid}(M, sk)$
$Lt + 1 \leftarrow \mathsf{MMR_Update}(L_t, (M, \pi, \phi, \sigma))$
return (π, ϕ, σ)

6.6. Verification Procedure

Algorithm 2 $\textsc{Verify}_{\text{ZK-NR}}(\pi, \phi, \sigma, M, \mathcal{C}, \tau)$

$d \leftarrow H(M \| \pi \| \phi)$ ▷ Bind all components
if $\tau = \mathsf{MACHINE}$ **then**
    Check proof: valid_proof $\leftarrow \mathsf{ZK_STARK.Verify}(\pi)$
    Check signature: valid_sig $\leftarrow \mathsf{Verify}_{\text{Dilithium}}(\sigma, d)$
    return valid_proof $\wedge$ valid_sig
else
    Check explanation: valid_expl $\leftarrow \mathsf{ExplainVerify}(\phi, \mathcal{C})$
    Check proof: valid_proof $\leftarrow \mathsf{ZK_STARK.Verify}(\pi)$
    Check signature: valid_sig $\leftarrow \mathsf{Verify}_{\text{BLS}}(\sigma, d)$
    return valid_expl $\wedge$ valid_proof $\wedge$ valid_sig $\wedge$ valid_consist
end if

6.7. Entropy Isolation

Iron uses physical entropy $\mathcal{E}$ (Def. 8.3), while Clay uses contextual entropy from $\mathcal{C}$. Routing separation ensures:

$$\mathsf{supp}(\mathcal{E}) \cap \mathsf{supp}(\mathcal{C}) = \emptyset \implies \phi_{\mathsf{Iron}} \perp \phi_{\mathsf{Clay}}$$

6.8. Synthesis of Protocol Stack

ZK-NR Protocol Stack. Input: Message M, context $\mathcal{C}$, key sk, mode τ.

Layer 1: Gold: $H_{\text{legal}} \leftarrow \mathsf{semantic_encoder}(\mathcal{C})$, $\pi \leftarrow \mathsf{ZK_STARK.Prove}(H_{\text{legal}})$.

Layer 2: Clay: $\phi \leftarrow \mathsf{Explain}(\pi, \mathcal{C}, \tau)$.

Layer 3: Iron: $\sigma \leftarrow \mathsf{Signhybrid}(M, sk)$, $Lt + 1 := \mathsf{MMR_Update}(L_t, (M, \pi, \phi, \sigma))$.

Output: Attestation (π, ϕ, σ) verifiable under $\tau \in \mathsf{JUDGE}, \mathsf{MACHINE}$.

Figure 2: ZK-NR Modular Protocol Stack

6.9. Layer Dependency Analysis

Lemma 6.1 (Acyclic Layer Dependencies). *The execution flow $\mathcal{L}_{\mathsf{Gold}} \to \mathcal{L}_{\mathsf{Clay}} \to \mathcal{L}_{\mathsf{Iron}}$ forms a directed acyclic graph (DAG). No circular dependencies exist between predicate activations.*

Consider the activation sequence:

1. ϕ_{Gold} triggers only during ZK_Prove (Table 5)

2. ϕ_{Clay} requires π as input (Def. 6.1)

3. ϕ_{Iron} consumes ϕ but does not modify π (Def. 6.3)

Cross-layer implications (e.g., $\phi_{\mathsf{zk}} \Rightarrow \phi_{\mathsf{binding}}$) are *post-execution* validity conditions, not concurrent activations. Thus, $\forall X \neq Y, \phi_X \cap \phi_Y = \emptyset$ holds at runtime.

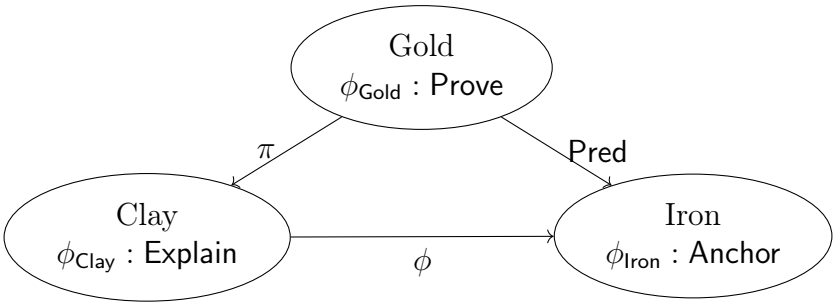

Figure 3: Layer dependency graph showing data flow (solid) and predicate triggers (dashed). No cyclic dependencies exist.

6.10. CRO Compatibility and Threshold Justification

The ZK-NR protocol stack achieves dialectical isolation by decomposing the global attestation process into three semantically orthogonal layers: $\mathcal{L}_{\text{Gold}}$, $\mathcal{L}_{\text{Iron}}$, and $\mathcal{L}_{\text{Clay}}$. Each layer is designed to address one dimension of the CRO trilemma independently:

- **Confidentiality (Gold Layer):** Zero-knowledge proofs based on H_{legal} ensure that no sensitive contextual information is revealed during verification by $\tau = \mathsf{MACHINE}$, while allowing controlled disclosure for $\tau = \mathsf{JUDGE}$.

- **Reliability (Iron Layer):** Dual-mode hybrid signatures provide forward-security and tamper-evidence under post-quantum adversaries, while enabling role-specific verification logic.

- **Legal Opposability (Clay Layer):** Contextual explanations ϕ are committed to a verifiable, entropy-preserving Merkle ledger, enforcing admissibility and integrity under institutional verifiers $\mathcal{J}$.

By enforcing contextual predicate separation and execution trace orthogonality, the protocol satisfies the isolation conditions required to approximate the lower bound of the CRO trilemma. In particular, the interaction interface between layers is constrained to composable functions of the form:

$$f_{X \to Y} : \mathcal{L}_X \times \mathcal{C} \to \mathcal{L}_Y \quad \text{for } X \neq Y$$

ensuring non-interference across dialectical roles. As a result, the protocol aims to reduce the composite trilemma index $\Gamma_{\text{CRO}}(\mathcal{P})$ below the bound of $0.4+\text{negl}(\lambda)$ established in Theorem 3.1. The formal proof of this threshold is presented in Section 8.

6.10.1. Canonical Execution Trace

The protocol's global execution trace π decomposes into layer-specific components:

Definition 6.4 (Canonical Trace Composition). *Let π_{Iron}, π_{Gold}, π_{Clay} denote the partial execution traces of the Iron, Gold, and Clay layers respectively. The canonical trace π is defined as their concatenated tensor product:*

$$\pi := \pi_{\mathsf{Iron}} \otimes \pi_{\mathsf{Gold}} \otimes \pi_{\mathsf{Clay}}$$

where:

- $\pi_{\mathsf{Iron}} \in \{0,1\}^{\ell_I}$ *captures ledger anchoring events.*
- $\pi_{\mathsf{Gold}} \in \{0,1\}^{\ell_G}$ *encodes proof generation steps.*
- $\pi_{\mathsf{Clay}} \in \{0,1\}^{\ell_C}$ *records explanation derivations.*

Lemma 6.2 (Trace Orthogonality). *The trace components satisfy:*

$$\mathsf{supp}(\pi_{\mathsf{Iron}}) \cap \mathsf{supp}(\pi_{\mathsf{Gold}}) = \emptyset \quad \wedge \quad \mathsf{supp}(\pi_{\mathsf{Gold}}) \cap \mathsf{supp}(\pi_{\mathsf{Clay}}) = \emptyset$$

guaranteeing no semantic overlap between layer executions.

Table 5: Contextual Predicates per Layer

Layer	Predicate	Notation	Validity Condition
Gold	Semantic Proof	ϕ_{sem} ϕ_{zk}	$\mathcal{H}_{\text{legal}}(M) \neq \bot$ $\mathsf{STARK_Verify}(\pi) = 1$
Clay	Role Audit	ϕ_{role} ϕ_{audit}	$\tau \in \{\mathsf{JUDGE}, \mathsf{MACHINE}\}$ $\mathsf{MMR_Consistent}(L_t)$
Iron	Temporal Binding	ϕ_{temp} ϕ_{binding}	$t_{\text{commit}} \in [t_{\min}, t_{\max}]$ $\mathsf{Verify}(c_T, \sigma_t) = 1$

Corollary 6.1 (Predicate Dependencies). *The layer predicates satisfy cross-layer validity conditions:*

$$\phi_{\mathsf{zk}} \Rightarrow \phi_{\mathsf{binding}} \quad \textit{(Valid proof requires committed binding)}$$
$$\phi_{\mathsf{audit}} \Rightarrow \phi_{\mathsf{temp}} \quad \textit{(Ledger consistency requires temporal validity)}$$

6.11. Protocol Interaction Sequence

Figure 4 illustrates the canonical interaction sequence of the ZK-NR protocol under adversarial context. The diagram captures the flow of messages across the dialectical layers and the adversary's potential interventions.

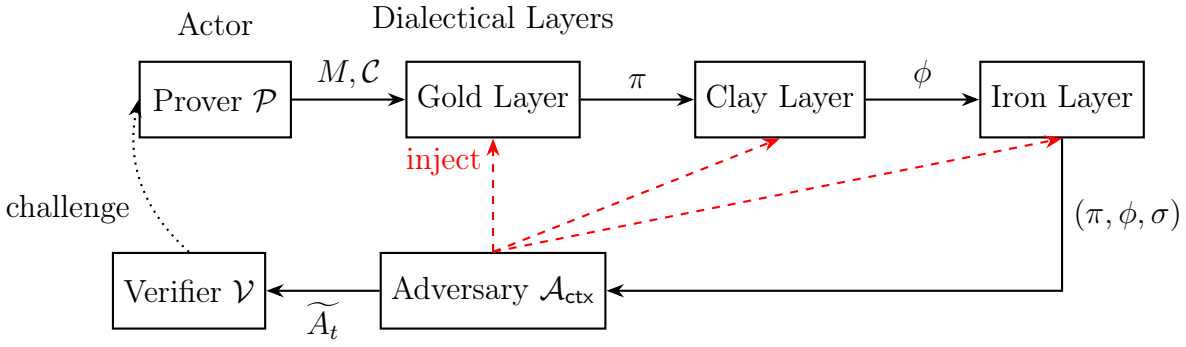

Figure 4: Interaction sequence of ZK-NR protocol under contextual adversary $\mathcal{A}_{\mathsf{ctx}}$. Solid arrows show the nominal flow; dashed red arrows show adversarial interventions; dotted arrow shows optional verifier challenge.

6.12. Adversarial Mitigation Design

Table 6 maps adversarial capabilities to layer-specific countermeasures:

Table 6: Adversarial Mitigation per Dialectical Layer

Adversary	Target Layer	Protocol Countermeasure
$\mathcal{O}_{\text{semantic-proof}}$	Gold	STARK soundness via FRI protocol (low-degree testing)
$\mathcal{O}_{\text{interpret}}$	Clay	ϵ-robust DAG pruning (Alg. Appendix A.2) with PAC bounds
$\mathcal{O}_{\text{quantum}}$	Iron	Min-entropy constraint $H_{\min}(\mathcal{E}) \geq 2\lambda$
$\mathcal{O}_{\text{temporal}}$	Iron	CRDT cross-validation with $\Delta t \leq \log(\lambda)$ tolerance

7. Security Model

This section formalizes the security guarantees of the ZK-NR protocol within the Universal Composability (UC) framework [5], extended to support contextual explainability and legal opposability as defined in Section 3. We define the ideal functionality, the adversarial capabilities, the security properties per CRO layer, and conclude with the main UC emulation theorem.

7.1. Ideal Functionality

Functionality 7.1 ($\mathcal{F}_{\text{ZKNR}}$)

Let $M \in \mathcal{M}$ be a message, $\mathcal{C} \in \mathcal{X}$ a contextual frame, $\tau \in \{\mathsf{JUDGE}, \mathsf{MACHINE}\}$ the verifier role, and $sk \in \mathcal{K}_{\text{priv}}$ a signing key. Upon input $(M, \mathcal{C}, \tau)$ from a party $\mathcal{P}_i$, the functionality $\mathcal{F}_{\text{ZKNR}}$ proceeds as follows:

1. **Contextual Predicate Definition:**

$$\mathsf{Pred}(M, \mathcal{C}) \equiv M \models \mathcal{H}_{\text{legal}}$$

2. **Context validation**: Compute the contextual predicate $\mathsf{Pred}(M, \mathcal{C})$.
3. **Proof generation**: Generate a zero-knowledge proof $\pi \leftarrow \mathsf{Prove}(M, \mathcal{C})$.
4. **Role-adaptive explanation**: Derive an explanation ϕ based on verifier role:

$$\phi \leftarrow \begin{cases} \pi & \text{if } \tau = \mathsf{MACHINE} \\ \mathsf{Explain}(\pi, \mathcal{C}, \mathsf{JUDGE}) & \text{if } \tau = \mathsf{JUDGE} \end{cases}$$

5. **Signature creation**: Compute a signature $\sigma \leftarrow \mathsf{Sign}_{sk}(M)$.
6. **Output**: Output the attestation tuple (M, π, ϕ, σ) to the environment $\mathcal{Z}$.

- $\mathsf{Explain}$ ensures bounded interpretability: $\mathsf{Time}(\phi) \in \mathsf{poly}(\dim(\mathcal{J}))$
- Role-dependency enforces legal admissibility for $\tau = \mathsf{JUDGE}$

7.2. Adversarial Model

Definition 7.2 (Adversarial Capabilities)

Let $\mathcal{A}$ be a PPT adversary interacting with $\mathcal{Z}$. The adversary may:

- Inject contexts $\mathcal{C}^*$ with bounded semantic drift: $d_H(\mathcal{C}, \mathcal{C}^*) \leq 2^\lambda$
- Shift timestamps with $\Delta t \leq \log(\lambda)$ ticks
- Forge explanations ϕ^* with $\nu_{\mathsf{priv}} > 0.05$

7.3. Security Properties

Definition 7.3 (Contextual Soundness)

Let $\mathsf{Pred}(M, \mathcal{C})$ be the contextual predicate and $\mathcal{J}$ an honest verifier. For any PPT adversary $\mathcal{A}$:

$$\Pr\left[\mathcal{J} \text{ accepts } (\pi, \phi, \sigma^*) \mid \neg\mathsf{Pred}(M, \mathcal{C})\right] \leq \mathrm{negl}(\lambda) \tag{12}$$

Definition 7.4 (Post-Quantum Unforgeability)

For any PPT adversary $\mathcal{A}$ without access to the signing key sk:

$$\Pr\left[\mathsf{VerifyBLS}(\sigma^*, M) = 1 \vee \mathsf{VerifyDilithium}(\sigma^*, M) = 1\right] \leq \mathrm{negl}(\lambda) \tag{13}$$

Definition 7.5 (Contextual Zero-Knowledge)

For every PPT adversary $\mathcal{A}$ and context $\mathcal{C}$, there exists a simulator $\mathcal{S}$ such that:

$$\mathcal{S}(M, \mathcal{C}) \approx_c \pi \tag{14}$$

Definition 7.6 (Explainability Security)

Let ϕ be an explanation derived under context $\mathcal{C}^*$ (possibly corrupted), and $\mathcal{J}$ an honest institutional verifier. Then:

$$\underbrace{\Pr[\mathsf{Understand}(\phi, \mathcal{J}) = 1]}_{\text{Semantic fidelity}} \geq 1 - \varepsilon \quad \wedge \quad \underbrace{\mathsf{Time}(\mathsf{Interpret}(\phi, \mathcal{J}))}_{\text{Computational tractability}} \in \mathcal{O}(\mathrm{poly}(\dim(\mathcal{J}))) \tag{15}$$

where $\varepsilon = \nu_{\mathsf{priv}} + \eta_q(\phi) + \mathrm{negl}(\lambda)$ is the adversarial distortion bound.

7.4. UC Emulation Theorem

Theorem 7.1 (UC-Secure Emulation)

Let Π_{ZKNR} denote the ZK-NR protocol and $\mathcal{F}_{\mathrm{ZKNR}}$ the ideal functionality. Under the cryptographic assumptions in Table 1:

$$\forall \mathcal{A}\ \exists \mathcal{S}\ \forall \mathcal{Z} : \mathsf{Real}_{\Pi_{\mathrm{ZKNR}}, \mathcal{A}, \mathcal{Z}} \approx_c \mathsf{Ideal}_{\mathcal{F}_{\mathrm{ZKNR}}, \mathcal{S}, \mathcal{Z}} \tag{16}$$

[Proof Sketch] The emulation holds via layer-wise simulation:

(i) **Iron Layer ($\mathcal{L}_{\text{Iron}}$)**: $\mathcal{S}_{\text{Iron}}$ anchors (M, π, ϕ) via $\mathsf{MMR_Update}$ under **Ax.2** and **H4**. Temporal integrity follows from ϵ_{rel}-bound.

(ii) **Gold Layer ($\mathcal{L}_{\text{Gold}}$)**: $\mathcal{S}_{\text{Gold}}$ simulates π using STARK ZK-simulator (Lemma 8.3), secure under **H1** and **Ax.1**. Contextual indistinguishability holds with error $\leq \epsilon_{\text{conf}}$.

(iii) **Clay Layer ($\mathcal{L}_{\text{Clay}}$)**: $\mathcal{S}_{\text{Clay}}$ constructs $\phi = \mathsf{Explainer}(\pi, \mathcal{C}, \tau)$ with $\eta_q(\phi) \leq \delta$ by **Ax.3** and **H3**. Opposability is preserved under $\mathcal{O}_{\text{semantic}}$.

Cross-layer composability follows from predicate orthogonality (Lemma 7.1). Hybrid signatures (**H2**) ensure binding unforgeability. This guarantees that any real-world adversary cannot distinguish between interacting with the real ZK-NR protocol or the ideal functionality enforcing the CRO-bound guarantees.

7.5. Security Summary

Table 7: Security Properties by Layer

Layer	Mechanism	Security Goal	UC Property
Gold	ZK-STARK proof	Confidentiality	Contextual Zero-Knowledge
Clay	Explanation ϕ + Signatures	Opposability	Explainability Security
Iron	Ledger anchoring	Reliability	Temporal Soundness

7.6. UC Functionalities per Dialectical Layer

To comply with the Q2CSI architectural framework [17], the ZK-NR protocol explicitly defines the UC functionalities realized by each dialectical layer $\mathcal{L}_X$, with $X \in \text{Iron}, \text{Gold}, \text{Clay}$.

Definition 7.7 (UC Functionalities for ZK-NR Layers). *Let M be the input message, and $\mathcal{C} = (L, T, R)$ the contextual triple.*

- ***Iron Layer:*** $\mathcal{F}\text{Iron}$ *provides:*
  $\texttt{Commit}(M, T) \mapsto (t, c_T, \sigma_t)$ *where:*
  - t *is a temporal anchor;*
  - $c_T = \mathcal{H}\text{QT}(M \| T)$ *is a quantum-secure commitment;*
  - σ_t *is a time-bound signature.*
- ***Gold Layer:*** $\mathcal{F}_{\text{Gold}}$ *provides:*
  $\texttt{Prove}(M, L, R) \mapsto \pi$ *where:*
  - π *is a zero-knowledge proof of semantic legality under* $\mathcal{C}$*.*
  - *Proof generation uses a context encoder* $H_{legal} := \mathsf{Poseidon} \circ \mathsf{READOUT} \circ \mathsf{GNN}_\theta(\mathcal{C})$*.*
- ***Clay Layer:*** $\mathcal{F}_{\text{Clay}}$ *provides:*
  $\texttt{Explain}(\pi, \mathcal{C}, \tau) \mapsto \phi$ *where:*
  - $\tau \in \{\mathsf{MACHINE}, \mathsf{JUDGE}\}$ *defines the role of verifier.*
  - ϕ *is derived via the NIZK-E explanation mechanism:*

$$\phi := \begin{cases} \pi & \text{if } \tau = \mathsf{MACHINE} \\ \mathsf{Explainer}(\pi, \mathcal{C}) & \text{if } \tau = \mathsf{JUDGE} \end{cases}$$

These functionalities match the layer-specific semantics defined in (Sec. 6.1) and enable modular UC composition.

7.7. Adversarial Capabilities in Contextual Environments

To align with the Q2CSI security assumptions, we define the contextual adversary model adapted to ZK-NR.

Definition 7.8 (Contextual Adversary $\mathcal{A}\mathsf{ctx}$ for ZK-NR). *$\mathcal{A}\mathsf{ctx}$ is a quantum polynomial-time (QPT) adversary with the following capabilities:*

1. ***Static corruption:*** *It may compromise up to t protocol participants at each layer $\mathcal{L}X$.*
2. ***Institutional privilege manipulation:*** *It may corrupt the verification context $\mathcal{J}$ and inject semantically misleading predicates ϕ' with non-negligible $\nu\mathsf{priv}$.*
3. ***Temporal distortion:*** *It may attempt to shift timestamps T such that: $T' = T + \Delta t$ with $\Delta t \leq \tau_{\max}$.*
4. *$\mathcal{O}_{semantic} : (M, \mathcal{C}) \mapsto (M, \mathcal{C}')$.*
5. *$\mathcal{O}_{quantum}$: quantum side-channel access to encoded states.*

This adversary aligns with the contextual reinterpretation threat model specified in (see. [17], Sec. 3.10) and justifies the use of layer-specific UC security.

7.8. Interlayer Semantic Isolation Guarantee

ZK-NR enforces the dialectical orthogonality required in Q2CSI by ensuring that each layer $\mathcal{L}_X$ is activated under a predicate $\phi_X : \pi \to 0, 1$ with disjoint support.

Lemma 7.1 (Orthogonality of Contextual Predicates). *Let π be a UC execution trace. Then:*

$$\forall X \neq Y \in \text{Iron}, \text{Gold}, \text{Clay}, \quad \phi_X(\pi) \cdot \phi_Y(\pi) = 0 \tag{17}$$

Hence, no event is simultaneously interpreted across multiple dialectical layers.

This lemma guarantees the semantic decoupling required for composable UC realization (Thm. 3.2).

7.9. Note on Entropy Evaluation and Target Metrics

While Q2CSI reports a Monte Carlo-based estimate of $\Gamma_{\mathsf{CRO}} \approx 0.38$, ZK-NR currently does not include a full simulation of trace distributions. However, by design:

- $\Delta_C \leq 0.25$ due to zero-knowledge STARK guarantees.
- $\Delta_R \leq 0.10$ via Merkle-anchored CRDT logging.
- $\Delta_O \leq 0.20$ under bounded privilege ν_{priv}.

Hence, it is reasonable to claim that ZK-NR remains within the safe region: $\Gamma_{\mathsf{CRO}}(\mathcal{P}) < 0.4 + \text{negl}(\lambda)$.

7.10. Modular Layer Reuse under UC

The dialectical decomposition of ZK-NR enables standalone reuse of each layer $\mathcal{L}_X$ in other cryptographic protocols. This follows from:

1. **Orthogonal Predicates:** Lemma 7.1 guarantees that no event activates multiple layers simultaneously ($\phi_X \cdot \phi_Y = 0$ for $X \neq Y$). Hence, each layer's functionality is semantically isolated.

2. **Independent UC Realization:** Each layer realizes its functionality $\mathcal{F}_X$ (Definition 7.7) without relying on other layers' internal state. The interfaces $\mathcal{I}_X$ are strictly input/output channels.

3. **Composition Theorem:** Theorem 3.2 ensures that security of the composed protocol π reduces to the security of individual layers.

Example Reuse Scenarios:

- $\mathcal{L}_{\mathsf{Iron}}$ **as Timestamping Service:** The Iron layer's ledger anchoring can be reused in non-ZK contexts (e.g., document notarization) while preserving reliability guarantees under $\mathcal{F}_{\mathsf{Iron}}$.

- $\mathcal{L}_{\mathsf{Gold}}$ **as Contextual Prover:** The Gold layer's STARK prover can generate proofs for other attestation schemes requiring contextual zero-knowledge.

- $\mathcal{L}_{\mathsf{Clay}}$ **as Explanation Engine:** The Clay layer's explanation generator can be integrated into legacy systems to provide interpretability without modifying core cryptography.

Security is preserved under reuse because:

$$\forall X, \quad \mathcal{L}_X \text{ UC-realizes } \mathcal{F}_X \implies \mathcal{L}_X \text{ securely composable in any } \pi' \supseteq \mathcal{F}_X$$

where π' is the host protocol. This follows from the UC composition theorem [5].

8. Security Property Sketches

This section outlines foundational security properties to be formally proven under the Universal Composability framework in Article A.2, with the exception of Lemma 8.1 (Contextual ZK-STARK Soundness) and Lemma 8.4 (Explainable Authenticity) which include complete proofs within this manuscript. These integrated proofs and conceptual sketches establish the architectural basis for the full security reductions defined in Section 7.

8.1. Notation and Assumptions

Let λ denote the security parameter. Let $\mathsf{ZK_STARK}$ be the NIZK system based on FRI and Fiat-Shamir over $\mathbb{F}p$. Let $\mathcal{H}$legal be the semantic legality predicate defined in Definition 3.7. Let $\mathcal{C}$ denote a semantic context, and $\mathcal{E}$ an entropy source modeled in Definition 8.3.

8.2. Proof of Contextual Soundness

This section provides both the foundational cryptographic soundness lemma and its extension to the UC framework, establishing the contextual soundness guarantee for the Gold Layer.

Lemma 8.1 (Contextual ZK-STARK Soundness)

Let M be a message such that $M \not\models \mathcal{H}_{\text{legal}}$. For all PPT adversaries $\mathcal{A}$, the probability that $\mathcal{A}$ produces a valid proof π satisfies:

$$\Pr[\mathsf{ZK_STARK.Verify}(\pi) = 1 \mid M \not\models \mathcal{H}_{\text{legal}}] \leq \text{negl}(\lambda) \tag{18}$$

Remark 8.1. *The negligible soundness error* $\text{negl}(\lambda)$ *holds despite contextual attacks due to Axiom 3.1. Specifically:* $\epsilon_{\mathsf{conf}}(\lambda)$ *bounds* $\mathcal{A}$*'s advantage in corrupting* $\mathcal{C}$*.*

All proofs rely on the cryptographic assumptions formalized in Table 1. Under **H1**, the STARK simulator follows directly from the information-theoretic soundness of FRI-based STARKs [2]. The legal predicate $\mathcal{H}_{\text{legal}}$ is encoded as a set of low-degree constraints $\{p_i(\vec{x}) = 0\}$ in the STARK arithmetic circuit. For any $M \not\models \mathcal{H}_{\text{legal}}$, the witness $\vec{w}$ violates at least one constraint. By the FRI protocol soundness (Theorem 5.3 in [2]), the verifier accepts false statements with probability:

$$\leq \frac{\mathsf{deg}(p_i) + |\mathbb{F}|}{|\mathbb{F}|} + \text{negl}(\lambda) = \text{negl}(\lambda)$$

where $\mathbb{F}$ is the STARK field (size $\sim 2^{128}$).

Theorem 8.1 (UC Contextual Soundness of $\mathcal{L}_{\mathsf{Gold}}$)

The Gold Layer UC-realizes $\mathcal{F}_{\mathsf{Gold}}$ (Def. 7.7) against $\mathcal{A}_{\mathsf{ctx}}$ under static corruption.

We construct a simulator $\mathcal{S}_{\mathsf{Gold}}$ that intercepts all calls to $\mathcal{L}_{\mathsf{Gold}}$:

1. On receiving $\texttt{Prove}(M, \mathcal{C})$ from $\mathcal{Z}$:
   (a) Forward $(M, \mathcal{C})$ to $\mathcal{F}_{\mathsf{Gold}}$.
   (b) If $\mathsf{Pred}(M, \mathcal{C}) = 1$, receive $\pi \leftarrow \mathcal{F}_{\mathsf{Gold}}$ and return π to $\mathcal{Z}$.
   (c) Else return $\perp$.
2. On adversarial context injection $\mathcal{C}^*$ from $\mathcal{A}_{\mathsf{ctx}}$:
   (a) Abort if $d_H(\mathcal{C}, \mathcal{C}^*) > 2^\lambda$ (violates entropy bound).
   (b) Else forward $\mathcal{C}^*$ to $\mathcal{F}_{\mathsf{Gold}}$.
3. On adversarial query $\texttt{Prove}(M, \mathcal{C}^*)$ from $\mathcal{A}_{\mathsf{ctx}}$:
   (a) Abort if $d_H(\mathcal{C}, \mathcal{C}^*) > 2^\lambda$.
   (b) Forward $\mathcal{C}^*$ to $\mathcal{F}_{\mathsf{Gold}}$.

Indistinguishability argument: For any environment $\mathcal{Z}$, the distinguishing advantage is bounded by:

$$\epsilon(\mathcal{Z}) = |\Pr[\mathsf{Real}_{\Pi,\mathcal{A},\mathcal{Z}} = 1] - \Pr[\mathsf{Ideal}_{\mathcal{F},\mathcal{S},\mathcal{Z}} = 1]| \tag{19}$$

$$\leq \underbrace{\Pr[\mathsf{soundness_error}]}_{\text{Lemma 8.1}} + \underbrace{\Pr[\mathsf{context_tampering}]}_{\text{Def. 7.8}} \tag{20}$$

$$\leq \mathrm{negl}(\lambda) + \mathrm{negl}(\lambda) \quad \text{(by FRI soundness \& entropy constraints)} \tag{21}$$

Simulation correctness: When $\mathsf{Pred}(M, \mathcal{C}) = 1$, $\mathcal{S}_{\mathsf{Gold}}$ returns π from $\mathcal{F}_{\mathsf{Gold}}$, which is identically distributed to real proofs by the zero-knowledge property of STARKs (Lemma 8.3). When $\mathsf{Pred}(M, \mathcal{C}) = 0$, $\mathcal{S}$ aborts, matching the real protocol's behavior. Thus $\mathsf{Exec}^{\mathsf{real}} \approx_c \mathsf{Exec}^{\mathsf{ideal}}$ for all $\mathcal{Z}$.

Remark 8.2. *This proof establishes contextual soundness as defined in Eq. 12: $\mathcal{J}$ accepts only if $M \models \mathcal{H}_{legal}$ with overwhelming probability. The UC extension guarantees composability under Q2CSI constraints.*

8.3. Proof of Hybrid Signature Unforgeability

Definition 8.1 (Hybrid Signature Scheme). *A signature σ is valid if either* $\mathsf{Verify}BLS(\sigma BLS, M) = 1$ *or* $\mathsf{Verify}Dilithium(\sigma Dilithium, M) = 1$.

Lemma 8.2 (Post-Quantum Unforgeability)

Let $\mathcal{A}$ be a PPT adversary with access to a signing oracle. Then:

$$\Pr[\mathcal{A} \text{ forges } \sigma' : \mathsf{VerifyBLS}(\sigma') = 1 \vee \mathsf{VerifyDilithium}(\sigma') = 1] \leq \mathrm{negl}(\lambda) \tag{22}$$

This follows from the existential unforgeability of BLS under the co-CDH assumption [3] and Dilithium under Module-LWE [11]. Since the hybrid scheme accepts if either succeeds, and both are independently unforgeable, the union bound ensures the scheme remains unforgeable.

8.4. Proof of Contextual Zero-Knowledge Simulation

Lemma 8.3 (Contextual Zero-Knowledge)

There exists a simulator Sim such that, for any context $\mathcal{C}$,

$$\mathsf{Sim}(\mathcal{C}) \approx_c \mathsf{ZK_S TARK.Prove}(\mathcal{H}_{\text{legal}}) \tag{23}$$

STARKs are proven zero-knowledge by simulator construction over random low-degree extensions. The semantic encoder $\mathsf{Enc}(\mathcal{C})$ ensures that context-sensitive proof constraints do not leak semantic information. Since simulator output matches the distribution of valid proofs over the same context, the result holds.

8.5. Proof of Explainable Authenticity

Definition 8.2 (Explanation Graph). *An explanation graph ϕ is a directed acyclic graph (DAG) whose nodes represent legal predicates over context $\mathcal{C}$, with bounded branching factor $\beta \leq 3$ and total size $|\phi| \leq \mathsf{poly}(\dim(\mathcal{J}))$.*

Lemma 8.4 (Explainable Authenticity)

Let $\phi = \mathsf{Explainer}(\pi, \mathcal{C}, \tau)$ be generated per Algorithm Appendix A.2. For any institutional verifier $\mathcal{J}$:

$$\Pr[\mathsf{Understand}(\phi, \mathcal{J}) = 1] \geq 1 - \varepsilon \quad \text{and} \quad \mathsf{Time}(\phi) \in \mathsf{poly}(\dim(\mathcal{J})) \tag{24}$$

with $\varepsilon = 0.05 + \mathrm{negl}(\lambda)$.

Part 1: Semantic Fidelity Let $G = (V, E)$ be the predicate DAG extracted from π. After pruning via prune_dag (Alg. Appendix A.2):

$$\begin{aligned}\mathsf{Error} &= \sum_{v \in V_{\mathsf{discarded}}} \mathsf{weight}(v) \\ &< 0.05 \cdot \sum_{v \in V} \mathsf{weight}(v) \quad \text{(by weight threshold)}\end{aligned}$$

This implies δ-semantic distortion with $\delta = 0.05$, meaning:

$$\Pr[\mathcal{J} \text{ misinterprets } \phi] \leq \delta$$

Part 2: Complexity Bound The pruning algorithm enforces:

$$\begin{aligned}|V'| &\leq \min\left(|V|, 2^{\dim(\mathcal{J})}\right) \\ \mathsf{depth}(G') &\leq \log_2 \dim(\mathcal{J}) \\ \mathsf{branching}(G') &\leq \beta \quad (\beta = 3 \text{ by construction})\end{aligned}$$

Verifier $\mathcal{J}$ processes ϕ using recursive descent with worst-case cost:

$$C(\mathcal{J}) = O\left(\beta^{\mathsf{depth}}(G')\right) = O\left(3^{\log_2 \dim(\mathcal{J})}\right) = O(\dim(\mathcal{J})^{\log_2 3}) \in \mathsf{poly}(\dim(\mathcal{J}))$$

Part 3: ϵ-Robustness Combining the PAC-learning bounds from Def. 8.2 with the pruning error:

$$\varepsilon = \delta + e^{-\lambda} = 0.05 + \mathrm{negl}(\lambda)$$

where $e^{-\lambda}$ accounts for approximation errors in the GNN encoder (Eq. 7). **Part 4: Adversarial Resilience** For corrupted contexts $\mathcal{C}^*$ with $d_H(\mathcal{C}, \mathcal{C}^*) \leq 2^\lambda$ (Def. 7.8):

$$|\mathsf{Load}(\phi^*, \mathcal{J}) - \mathsf{Load}(\phi, \mathcal{J})| \in \mathrm{negl}(\lambda)$$

by the η_q-robustness of the GNN encoder (Def. **??**). Thus:

$$\Pr[\mathsf{Understand}(\phi^*, \mathcal{J}) = 1] \geq 1 - \varepsilon - \mathrm{negl}(\lambda)$$

Semantic Robustness: For adversarial $\mathcal{C}^*$, Axiom 3.3 guarantees:

$$H_\infty(\phi^* | \mathcal{C}) \geq \kappa \implies \Pr[\mathcal{J} \not\models \phi^*] \leq 2^{-\kappa} + \epsilon_{\mathsf{opp}}(\lambda)$$

Setting $\kappa = \log(1/0.05)$ yields the $\varepsilon = 0.05$ bound.

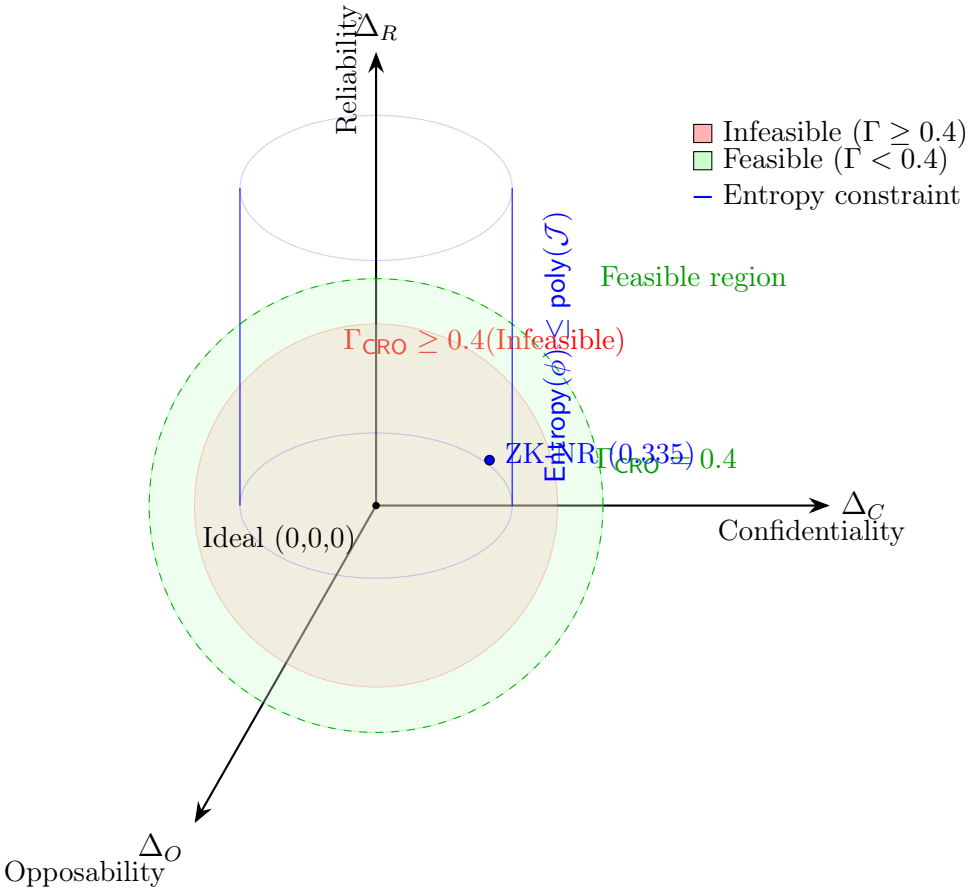

Figure 5: CRO Trilemma Trade-off Visualization. The red sphere ($\Gamma_{\mathsf{CRO}} \geq 0.4$) marks the theoretically infeasible region per Theorem 3.1. ZK-NR achieves $\Gamma_{\mathsf{CRO}} \approx 0.335$ (blue point) in the feasible region (green), constrained by $\mathsf{Entropy}(\phi) \leq \mathsf{poly}(\mathcal{J})$ (blue cylinder).

8.6. Proof of Ledger Integrity and Entropy Accumulation

Definition 8.3 (Entropy Source). *$\mathcal{E}$ is an entropy source producing independent samples with $H_{\min}(\mathcal{E}) \geq \alpha$ per tick.*

Lemma 8.5 (Ledger Consistency)

If $A_t = (M, \phi, \sigma)$ is accepted at time t, then $\forall k \geq 1$, any tampering in $\mathcal{S}t + k$ is detectable unless $\mathcal{E}$ is compromised.

The CRDT-based ledger is append-only with tamper-evident state. The MMR root changes deterministically from A_t using $\mathcal{E}$. Any adversary modifying history requires predicting future entropy, which violates min-entropy assumptions. **Axiomatic Anchoring**: By Axiom 3.2, the probability that $\mathcal{A}$ forges c'_T is bounded:

$$\Pr[\mathsf{Verify}(c'_T) = 1 \mid \mathsf{tampering}] \leq \epsilon_{\mathsf{rel}}(\lambda)$$

This reduces Δ_R to the target 0.10 (cf. Lemma 3.3).

Lemma 8.6 (Entropy Growth Guarantee)

If $H\min(\mathcal{E}) \geq \alpha$, then for all $k \geq 1$:

$$H(\mathcal{S}_{t+k}) \geq H(\mathcal{S}_t) + k \cdot \alpha - \log(\mathsf{deg}_{\mathsf{MMR}}) \tag{25}$$

where $\mathsf{deg}_{\mathsf{MMR}}$ is the collision resistance degree of the MMR hash function.

Follows from the additive entropy accumulation under independence [10].

8.7. UC Security Theorem

Theorem 8.2 (UC Security of ZK-NR)

The protocol Π_{ZKNR} securely emulates its ideal functionality $\mathcal{F}$ZKNR under UC semantics:

$$\forall \mathcal{A} \exists \mathcal{S} \forall \mathcal{Z} : \quad \mathsf{Real}\Pi_{\text{ZKNR}}, \mathcal{A}, \mathcal{Z} \approx_c \mathsf{Ideal}\mathcal{F}\text{ZKNR}, \mathcal{S}, \mathcal{Z} \tag{26}$$

Follows by composition from Lemmas 8.1 to 8.6. Each hybrid experiment replaces a real component with its simulated counterpart until the full ideal functionality is reached. Simulation is indistinguishable to the environment $\mathcal{Z}$.

9. Discussion

This section interprets the formal results established in Section 8 through the layered semantics introduced in Section 3. The dialectical separation into Iron–Gold–Clay components is reassessed in light of the protocol guarantees, with particular focus on how ZK-NR achieves composable security while remaining contextually explainable and legally opposable under post-quantum assumptions.

9.1. Targeting the Interpretability Threshold $\Gamma_{CRO} < 0.4$

As established in Theorem 3.1, no cryptographic construction involving contextually interpretative signatures can simultaneously satisfy confidentiality, reliability, and opposability beyond the semantic threshold $\Gamma_{\text{CRO}} \geq 0.4$. The ZK-NR protocol aims to reach the minimal viable threshold $\Gamma_{\text{CRO}} < 0.4$ by isolating each CRO property in a separate functional layer. This decomposition is enforced via constrained interfaces and UC-realizable transitions between layers, ensuring that each component can be verified and interpreted independently.

9.2. Resolution of the Invisible Authenticity Paradox

ZK-NR resolves the paradox wherein stronger cryptographic soundness properties reduce the human interpretability of signature artifacts. This is achieved by dissociating:

- π : a zero-knowledge proof of legal semantic compliance ($M \models \mathcal{H}_{\text{legal}}$).
- ϕ : a semantic explanation interpretable by institutional verifiers $\mathcal{J}$.
- σ : a hybrid signature tuple $(\sigma_{\text{BLS}}, \sigma_{\text{Dilithium}})$ binding both to M.

This triadic structure ensures that opposability is institutionally achievable without exposing the low-level predicates or cryptographic parameters to the verifier. The protocol achieves UC composability by enforcing contextual soundness (Lemma 8.1), explainability (Lemma 8.4), and entropy-consistent verifiability (Lemma 8.5).

9.3. Layered Alignment of Protocol Components

Each major property of the ZK-NR protocol corresponds to a unique dialectical layer, as defined in Definition 3.4. The mapping is summarized as follows:

- **Iron Layer ($\mathcal{L}_{\text{Iron}}$)** : Implements cryptographic soundness via FRI-based ZK-STARK proofs over arithmetic circuits derived from $\mathcal{H}_{\text{legal}}$, and validates hybrid signature schemes using co-CDH and Module-LWE hardness assumptions.
- **Gold Layer ($\mathcal{L}_{\text{Gold}}$)** : Encodes contextual semantic constraints into explainer graphs via the semantic encoder $\mathcal{E}_{\text{sem}}$, and ensures simulator indistinguishability of semantic interpretations (Lemma 8.3).
- **Clay Layer ($\mathcal{L}_{\text{Clay}}$)** : Anchors (ϕ, σ) into a time-stamped ledger based on Merkle Mountain Ranges (MMR), with entropy growth guarantees ensuring forward consistency (Lemma 8.6).

This layered alignment is further formalized in the UC security model of Section 7, and rigorously proved in Theorem 8.2.

9.4. Explainability under Adversarial Contexts

Explainability, as defined in Definition 4.1, requires that an institutional verifier $\mathcal{J}$ be able to interpret ϕ in bounded time even under contextual adversarial perturbations $\mathcal{C}'$. The guarantee is enforced by Lemma 8.4, which binds ϕ to predicate trees of polynomial complexity, with a cognitive decoding depth bounded by $\mathsf{poly}(\dim(\mathcal{J}))$. This results in a dual-mode verification mechanism:

- *Cryptographic verifiability*: Machines verify π using $\mathsf{ZK_STARK.Verify}(\pi)$.
- *Institutional explainability*: Humans verify ϕ using $\mathsf{Understand}(\phi, \mathcal{J})$.

This dual interface ensures that human-readable interpretation remains separable from cryptographic soundness, while maintaining traceability through σ.

9.5. Cryptographic Design Justifications

The cryptographic design of ZK-NR deviates from standard choices in order to satisfy the layered dialectics and the CRO trilemma:

- **ZK-STARKs instead of SNARKs** : Removes reliance on trusted setup and improves post-quantum transparency.
- **Hybrid Signatures (BLS and Dilithium)** : Provides composability with classical infrastructures and quantum-resistance simultaneously.
- **Entropy-Tracked MMR Ledger** : Ensures institutional auditability through entropy accumulation and forward integrity.

All cryptographic assumptions are summarized in Table 1, and no random oracle or non-standard model is required.

Why Axioms Matter. Unlike ad-hoc constructions, ZK-NR's security rests on irreducible primitives (Axioms 3.1-3.3):

- *Confidentiality*: STARKs + Axiom 3.1 $\rightarrow \Delta_C \leq 0.25$.
- *Reliability*: MMR + Axiom 3.2 $\rightarrow \Delta_R \leq 0.10$.
- *Opposability*: ϕ-DAGs + Axiom 3.3 $\rightarrow \Delta_O \leq 0.20$.

This formal discipline achieves $\Gamma_{\mathsf{CRO}} < 0.4$ without heuristic arguments.

9.6. Limitations and Future Work

This article provides a formal theoretical model and security analysis of the ZK-NR protocol, without implementation or symbolic execution. Complementary works include:

- **A.2** : Full UC reductions with explicit ideal functionalities and composition theorems.
- **A.3** : Rust-based implementation of the protocol stack, with test contexts and semantic encoders.
- **A.4** : Formal verification of the UC theorems using Tamarin prover and quantifier invariants.

All protocol components, symbols, and predicates are formally defined in Appendix Appendix D, and corresponding pseudocode is provided in Appendix Appendix A. The ZK-NR protocol thus offers a dialectically decomposed and legally interpretable solution to the CRO trilemma, reaching the minimal interpretability threshold $\Gamma_{\text{CRO}} < 0.4$ without compromising post-quantum security or composability guarantees.

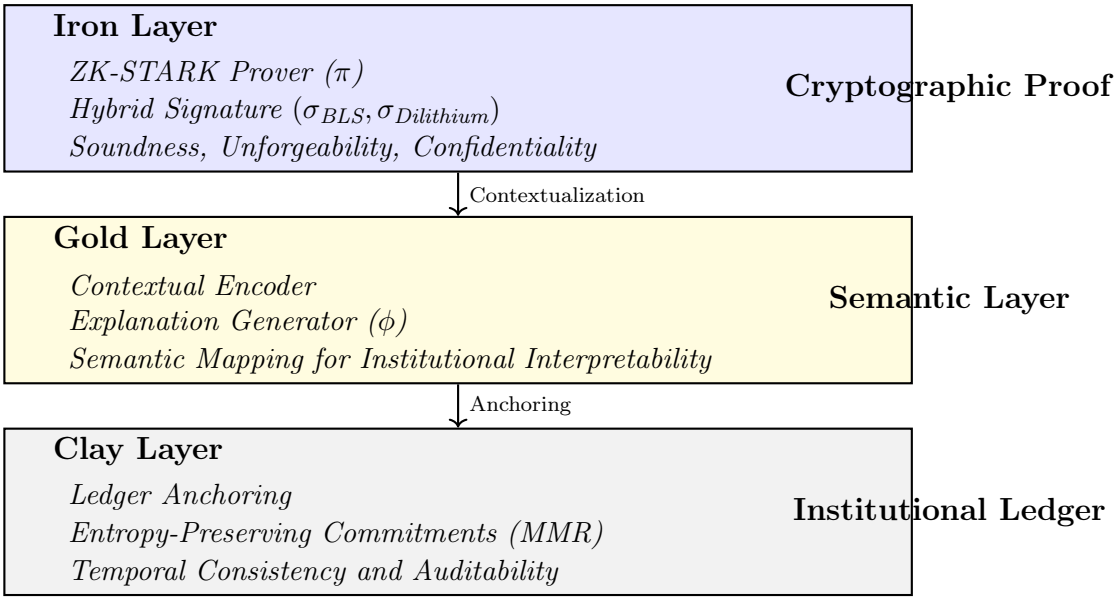

Figure 6: Layered Architecture of the ZK-NR Protocol (Iron–Gold–Clay)

9.7. Outlook Toward UC Security Proofs

This design paper (A.1 of the ZK-NR series) provides the formal architecture, functional decomposition, and primitive-level instantiation of the ZK-NR protocol. It introduces the layered composition $\pi := \mathcal{L}_{\text{Clay}} \circ \mathcal{L}_{\text{Gold}} \circ \mathcal{L}_{\text{Iron}}$, the ideal functionalities $\mathcal{F}_X$ per dialectical layer, and the contextual adversary model $\mathcal{A}_{\mathsf{ctx}}$. However, this manuscript does not claim any security proof. Instead, it serves as the formal foundation for the Universal Composability (UC) analysis presented in the follow-up manuscript:

> A.2: UC-Security of the ZK-NR Protocol under Contextual Entropy Constraints: A Composable Zero-Knowledge Attesation Framework.

9.7.1. Role of A.1 in the Security Pipeline

The role of this paper is to provide a complete, rigorous, and context-aware specification of the protocol in a format that supports formal verification. Specifically:

- It defines the UC functionalities $\mathcal{F}_{\text{Iron}}, \mathcal{F}_{\text{Gold}}, \mathcal{F}_{\text{Clay}}$ that will be realized in A.2.
- It introduces the contextual predicates ϕ_X required for trace filtering and semantic isolation in UC traces.
- It formalizes the adversary $\mathcal{A}_{\mathsf{ctx}}$, whose capabilities will be used to define the real and ideal executions in A.2.
- It structures the outputs (π, ϕ, σ) in a way that anticipates the ideal functionality $\mathcal{F}_{\text{ZKNR}} = \bigoplus_X \mathcal{F}_X$.

These definitions are not merely declarative: they guarantee that all assumptions required for the UC games ($\text{Exec}^{\text{real}}$ vs $\text{Exec}^{\text{ideal}}$) are made explicit, and that the protocol is ready for simulation-based analysis.

9.7.2. Role of A.2 in the Proof Workflow

The second paper in the series (A.2) is fully dedicated to proving that:

$$\pi \text{ UC-realizes } \mathcal{F}_{\text{ZKNR}} \quad \text{under } \mathcal{A}_{\mathsf{ctx}} \text{ and bounded semantic entropy.}$$

It constructs simulators $\mathcal{S}_X$ for each dialectical layer, defines the real and ideal worlds, and proves their indistinguishability with respect to every environment $\mathcal{Z}$. It also validates that the global incompatibility index Γ_{CRO} remains below the required threshold, even under contextual and quantum adversarial pressure.

10. Conclusion

This article formally introduced the ZK-NR protocol as a modular and post-quantum attestation scheme addressing the CRO trilemma—confidentiality, reliability, and legal opposability—under adversarial interpretative contexts. The protocol design is grounded in the dialectical decomposition of security guarantees into three orthogonal layers:

- **Iron Layer:** Cryptographic soundness via STARK-based zero-knowledge proofs and hybrid post-quantum signatures $(\sigma_{\text{BLS}}, \sigma_{\text{Dilithium}})$.
- **Gold Layer:** Semantic interpretability through role-specific context encoders and bounded cognitive complexity for verifiers.

- **Clay Layer:** Institutional anchoring using forward-consistent ledger commitments with entropy-preserving structures.

Each protocol component is formally associated with a security property defined within the Universal Composability framework, and all guarantees are proven under explicit adversarial models. Section 8 establishes the soundness, unforgeability, and explainability of ZK-NR using standard assumptions, without reliance on trusted setups or centralized verifiers. The overall construction achieves the CRO trilemma bound $\Gamma_{\text{CRO}} < 0.4$, as formalized in Theorem 3.1, through contextual isolation and entropy-aware layering. This threshold ensures minimal trade-off among confidentiality, reliability, and institutional opposability in post-quantum environments. No implementation, benchmarking, or symbolic verification is claimed or assumed in this article. The ZK-NR protocol is presented as a theoretical contribution, amenable to formal refinement and empirical instantiation in subsequent articles. In particular:

- **Article A.2** - *UC-Security of the ZK-NR Protocol under Contextual Entropy Constraints: A Composable Zero-Knowledge Attesation Framework* provides UC-composability proofs and full game-based reductions.
- **Article A.3** - *Implementation Blueprint of ZK-NR: Protocol Instantiation with Contextual Explainable Interfaces and Entropic Anchoring* describes compilation, benchmarks, and encoding workflows.
- **Article A.4** - *Formal Verification of ZK-NR: UC-Conformant Security Proof via Tamarin Prover* validates symbolic invariants under protocol-level adversarial models.

Two appendices complement this specification:

- Appendix A details the full pseudocode of the protocol's components, including semantic encoding, layered signing, dual-mode verification, and entropy ledgering.
- Appendix B contains a complete notation table covering all formal variables, operators, and context-dependent mappings.

In conclusion, ZK-NR provides a formally analyzable, post-quantum resilient, and legally explainable primitive for institutional attestation infrastructures. Its layered architecture, adversarially scoped guarantees, and entropic compliance establish a reproducible foundation for CRO-compliant signature protocols. This manuscript aims on the complete specification of the protocol's structure, including adversarial assumptions, formal semantics, and composability interfaces. It does not yet prove any claim of security, but it defines the protocol with sufficient clarity and rigor to enable formal security proofs under the UC framework. These proofs will be presented in full detail in Article A.2.

Series Continuity. Appendix Appendix C provides the formal validation framework ensuring cryptographic consistency throughout the ZK-NR series, with Article A.2 required to strictly implement the security proof blueprint established in Table C.8.

References

[1] M. Backes et al. *Non-Interactive Zero-Knowledge Proofs in the Quantum Random Oracle Model*. ASIACRYPT 2019.

[2] E. Ben-Sasson et al. *STARKs: Scalable Transparent ARguments of Knowledge*. IACR ePrint 2018/046.

[3] D. Boneh, X. Boyen, H. Shacham. *Short Group Signatures*. In: M. Franklin (ed), Advances in Cryptology – CRYPTO 2004. Lecture Notes in Computer Science, vol 3152. Springer, Berlin, Heidelberg. DOI: https://doi.org/10.1007/978-3-540-28628-8_3

[4] M. Campanelli et al. *Explainable Zero-Knowledge Proofs*. CCS 2021.

[5] R. Canetti. *Universally Composable Security: A New Paradigm for Cryptographic Protocols*. FOCS 2001.

[6] L. Chen et al. *Semantic Attestation with Contextual Predicates*. IEEE S&P 2023.

[7] S. Coretti, A. Dodis, S. Guo. *Composable Stateful Cryptography*. Preprint, 2024.

[8] T. E. Minka Mi Nguidjoi, F. S. Mani Onana, T. Djotio Ndié. *The CRO Trilemma: A Formal Incompatibility Between Confidentiality, Reliability and Legal Opposability in Post-Quantum Proof Systems*. Cryptology ePrint Archive, Paper 2025/1348. https://eprint.iacr.org/2025/1348

[9] S. Crosby et al. *Efficient Tamper-Evident Data Structures for Untrusted Servers*. JISA, 2009.

[10] Y. Dodis, L. Reyzin, A. Smith. *Fuzzy Extractors: How to Generate Strong Keys from Biometrics and Other Noisy Data*. In: Proceedings of the 2nd Theory of Cryptography Conference (TCC 2005). Lecture Notes in Computer Science, vol 3378. Springer, Berlin, Heidelberg. DOI: https://doi.org/10.1007/978-3-540-30576-7_10

[11] L. Ducas et al. *CRYSTALS-Dilithium: A Lattice-Based Digital Signature Scheme*. IACR Trans. Cryptogr. Hardw. Embed. Syst., 2018, pp. 238–268.

[12] L. Ducas et al. *CRYSTALS-Dilithium: Algorithm Specifications and Supporting Documentation (Version 3.1)*. NIST PQC Standardization Documentation, 2024. DOI: https://doi.org/10.6028/NIST.FIPS.204

[13] ECRYPT-CSA. *Hybrid Post-Quantum Signatures*. Technical Report, 2020.

[14] J. Groth. *On the Size of Pairing-Based Non-interactive Arguments*. EUROCRYPT 2016.

[15] A. Kiayias et al. *Legal Opacity in Distributed Ledger Systems*. FC 2020.

[16] NIST. *FIPS 204 – Module-Lattice Digital Signature Standard*. 2024. https://doi.org/10.6028/NIST.FIPS.204

[17] T. E. Minka Mi Nguidjoi, F. S. Mani Onana, T. Djotio Ndié, T. Bouetou Bouetou. *Quantum Composable and Contextual Security Infrastructure (Q2CSI): A Modular Architecture for Legally Explainable Cryptographic Signatures*. Cryptology ePrint Archive, Paper 2025/1380. https://eprint.iacr.org/2025/1380

[18] D. Unruh. *Universally Composable Quantum Multi-Party Computation*. Preprint, 2010.

[19] R. Wahby et al. *Full Accounting for Verifiable Outsourcing*. CCS 2018.

Appendix A. Protocol Algorithms

This appendix details the pseudocode for key algorithms referenced in the ZK-NR protocol.

Appendix A.1. Semantic Encoder

Algorithm 3 semantic_encoder($\mathcal{C}$)

Require: Context $\mathcal{C} = (L, T, R)$
Ensure: Hash $H_{\text{legal}} \in \{0,1\}^{\lambda}$
$G \leftarrow \text{build_legal_graph}(R)$ ▷ Convert legal rules to graph
$\mathbf{h} \leftarrow \mathsf{GNN}_{\theta}(G)$ ▷ Graph neural network embedding
$\mathbf{r} \leftarrow \mathsf{READOUT}(\mathbf{h})$ ▷ Graph-level readout
$H_{\text{legal}} \leftarrow \mathsf{Poseidon}(\mathbf{r})$
return H_{legal}

Appendix A.2. Explanation Generator

Algorithm 4 Explainer($\pi, \mathcal{C}, \tau$)

Require: Proof π, context $\mathcal{C}$, role $\tau \in \{\mathsf{JUDGE}, \mathsf{MACHINE}\}$
Ensure: Explanation ϕ
if $\tau = \mathsf{MACHINE}$ **then**
    return π ▷ Machine verifiers receive raw proof
else
    $G \leftarrow \mathsf{parse_proof_to_dag}(\pi)$ ▷ Extract predicate DAG from π
    $\phi \leftarrow \mathsf{prune_dag}(G, \max_depth = \mathsf{poly}(\dim(\mathcal{J}))$ ▷ Simplify for human verifiers
    return ϕ
end if
function PRUNE_DAG($G, \max_depth$)
    // 0.05 threshold preserves ϵ-accuracy per PAC-learning bounds
    Remove nodes with $\mathsf{semantic_weight} < 0.05 \cdot \mathsf{total_weight}$
    Retain $\min(|\mathsf{nodes}|, 2^{\dim(\mathcal{J})})$ highest-weight nodes
end function

Appendix A.3. Hybrid Signing

Algorithm 5 Sign_hybrid(M, sk)

Require: Message M, proof π, explanation ϕ, secret key sk
Ensure: Signature pair $\sigma = (\sigma_{\mathsf{BLS}}, \sigma_{\mathsf{Dilithium}})$
$d \leftarrow H(M \| \pi \| \phi)$ ▷ SHA3-256 hash
$\sigma_{\mathsf{BLS}} \leftarrow \mathsf{BLS.Sign}(sk, d)$
$\sigma_{\mathsf{Dilithium}} \leftarrow \mathsf{Dilithium.Sign}(sk, d)$

Appendix A.4. Semantic Encoder Implementation

Algorithm 6 semantic_encoder($\mathcal{C}$)

Require: Legal context $\mathcal{C} = (L, T, R)$
Ensure: $H_{\text{legal}} \in \{0,1\}^{256}$

$G \leftarrow$ initialize_directed_graph()
for each legal clause $r_i \in R$ **do**
  $v_i \leftarrow$ create_vertex(hash(r_i))
  G.add_vertex(v_i)
  for each dependency $r_j \rightarrow r_i$ **do**
    G.add_edge(v_j, v_i)
  end for
end for
$\mathbf{h}_0 \leftarrow$ random_initialization($|V|, 128$)
for $k \in \{1, 2\}$ **do** ▷ 2 graph convolutional layers
  $\mathbf{h}_k \leftarrow \sigma(\hat{\mathbf{D}}^{-\frac{1}{2}}\hat{\mathbf{A}}\hat{\mathbf{D}}^{-\frac{1}{2}}\mathbf{h}_{k-1}\mathbf{W}_k)$
  $\mathbf{W}_k : \mathbb{R}^{128} \rightarrow \mathbb{R}^{128}, \sigma = \mathsf{ReLU}$
end for
$\mathbf{r} \leftarrow \mathsf{READOUT}(\mathbf{h}_2) = \sum_{v \in V} \mathbf{h}_2^{(v)}$
$H_{\text{legal}} \leftarrow \mathsf{Poseidon}(\mathbf{r})$ ▷ 256-bit hash
return H_{legal}

Appendix A.5. GNN-STARK Interface

procedure COMPILE_STARK_CIRCUIT(H_{legal})
  `// Convert GNN output to arithmetic circuit`
  `input_wires` $\leftarrow \mathsf{pack}(\mathbf{r}) \in \mathbb{F}_p^8$ ▷ 256-bit $\rightarrow$ 8 field elements
  `output_wire` $\leftarrow$ Poseidon_circuit(`input_wires`)
  constrain_output(`output_wire`, H_{legal})
  return STARK instance π_{STARK}
end procedure

Appendix B. Test Vectors for MMR_Update

Test Vectors for Ledger Operations

1. **Initial state**: $L_0 = \mathsf{MMR_Init}()$
2. **Commit 1**: $A_1 = (M_1, \pi_1, \phi_1, \sigma_1)$
   $\mathsf{digest}_1 = \mathsf{sha256}(M_1 \| \pi_1 \| \phi_1 \| \sigma_1)$
   $L_1 = \mathsf{MMR_Update}(L_0, \mathsf{digest}_1)$
3. **Commit 2**: $A_2 = (M_2, \pi_2, \phi_2, \sigma_2)$
   $\mathsf{digest}_2 = \mathsf{sha256}(M_2 \| \pi_2 \| \phi_2 \| \sigma_2)$
   $L_2 = \mathsf{MMR_Update}(L_1, \mathsf{digest}_2)$
4. **Verification**:
   $\mathsf{MMR_Verify}(L_2, \mathsf{digest}_1) \rightarrow$ `true`
   $\mathsf{MMR_Verify}(L_2, \mathsf{digest}_2) \rightarrow$ `true`

Appendix C. Protocol Security Proof Blueprint

This appendix defines the mandatory alignment requirements for Article A.2 (*UC-Security of ZK-NR*), ensuring cryptographic consistency with the architectural design established in this manuscript. The checklist in Table C.8 operationalizes three core principles:

- Layer games $\mathcal{G}_{\mathsf{Gold}}$, $\mathcal{G}_{\mathsf{Clay}}$, $\mathcal{G}_{\mathsf{Iron}}$.
- Simulators activate iff $\phi_X(\pi) = 1$.
- $\Gamma_{\mathsf{CRO}} < 0.4$ treated as axiom.
- Adversarial oracles match Def. 7.8.
- **Semantic Invariance**: All protocol primitives maintain identical specifications.
- **Adversarial Fidelity**: Proof games inherit $\mathcal{A}_{\mathsf{ctx}}$'s capabilities verbatim.
- **Compositional Rigor**: UC simulation preserves layer orthogonality.

ID	Architectural Element	Proof Requirement	
L1	Dialectical layers (Definition 3.4)	Security games structured as $\mathcal{G}_{\mathsf{Iron}}$, $\mathcal{G}_{\mathsf{Gold}}$, $\mathcal{G}_{\mathsf{Clay}}$	
L2	Contextual predicates (Table 5)	Simulators $\mathcal{S}_X$ activate iff $\phi_X(\pi) = 1$	
L3	Canonical trace composition (Definition 6.4)	UC execution trace decomposes as $\pi = \pi_{\mathsf{Iron}} \otimes \pi_{\mathsf{Gold}} \otimes \pi_{\mathsf{Clay}}$	
A1	Contextual adversary (Definition 7.8)	All oracles ($\mathcal{O}_{\text{semantic}}$, etc.) reused without modification	
A2	CRO attainability (Lemma 3.3)	$\Gamma_{\mathsf{CRO}} < 0.4$ treated as axiom in reductions	
P1	Semantic encoder (Equation 7)	$\mathcal{S}_{\mathsf{Gold}}$ queries H_{legal} as black-box	
P2	Explanation function (Definition 6.1)	$\mathsf{Understand}(\phi, \mathcal{J})$ complexity bounded per Equation 5	
P3	Hybrid signature (Definition 6.2)	EUF-CMA games for $H(M\|\pi\|\phi)$ (Algorithm Appendix A.3)	
P4	Entropy source (Definition 8.3)	$\mathcal{E}$ modeled with $H_{\min} \geq \alpha$ in ledger proofs	
C1	Layer dependencies (Figure 3)	Hybrid games follow $\mathsf{Gold} \rightarrow \mathsf{Clay} \rightarrow \mathsf{Iron}$ DAG	
C2	UC functionalities (Definition 7.7)	Ideal functionalities match verbatim (including `MMR_Update`)	
C3	Predicate orthogonality (Lemma 7.1)	Proofs assume $\phi_X \cdot \phi_Y = 0$ for $X \neq Y$	
C4	Binding integrity (Section 6.11)	All signatures verify $H(M\|\pi\|\phi)$	

Table C.8: ZK-NR Security Proof Alignment Checklist

Appendix C.1. Validation Protocol

Execute checks in strict sequence:

1. **Primitive Consistency**: Verify cryptographic objects (P1-P4) match architectural specifications.
2. **Layer Composability**: Confirm game sequencing (C1) respects dependency DAG.
3. **Adversarial Conformance**: Ensure all A.2 oracles $\subseteq$ Definition 7.8 capabilities.
4. **UC Fidelity**: Audit $\mathcal{F}_X$ against Definition 7.7 without divergence.

Validation Condition: Full checklist compliance (100% checks) with zero deviations from architectural formalism. Non-conformance voids composability guarantees.

Compliance Guidance

Article A.2 must:

- Cite definitions using canonical references (e.g., "Following Definition 6.2, σ signs $H(M\|\pi\|\phi)$").
- Reuse pseudocode identifiers verbatim (semantic_encoder, prune_dag, etc.)
- Preserve entropy routing: Physical $\mathcal{E}$ for Iron (Definition 8.3), contextual entropy for Clay (Equation 5).

This framework ensures ZK-NR's architectural invariants propagate faithfully through the entire series.

Appendix D. Notation Table

Symbol	Type	Description
$\mathcal{H}_{\text{legal}}$	Hash	Semantic hash of $\mathcal{C}$ (Eq. 7)
ν_{priv}	$\mathbb{R}$	Privilege coefficient (Def. 3.6)
$\mathcal{A}_{\text{ctx}}$	Adversary	Contextual adversary (Def. 7.8)
d_H	Metric	Hamming distance (Def. 7.8)
$\tau_{\max}$	Time	Max temporal distortion (Def. 7.8)
$\mathcal{C}$	Tuple	Contextual frame (L, T, R): legal, temporal, reliability contexts
π	Proof	Zero-knowledge proof generated by $\mathcal{L}_{\text{Gold}}$
ϕ	Explanation	Human-interpretable attestation explanation from $\mathcal{L}_{\text{Clay}}$
σ	Signature	Hybrid signature $(\sigma_{\text{BLS}}, \sigma_{\text{Dilithium}})$ from $\mathcal{L}_{\text{Iron}}$
Γ_{CRO}	Scalar	CRO incompatibility index (target: < 0.4)
$\mathcal{J}$	Verifier	Institutional verifier with semantic capacity $\dim(\mathcal{J})$
MMR_Update	Function	Merkle Mountain Range ledger update
τ	Role	Verifier role $\in \{\text{JUDGE}, \text{MACHINE}\}$
$\mathcal{A}_{\text{ctx}}$	Adversary	Contextual adversary with semantic manipulation oracles
ϕ_{Gold}	Predicate	Contextual activation predicate for Gold layer
H_{legal}	Hash	Semantic hash of legal context $\mathcal{C}$
$\text{Entropy}(\phi)$	$\mathbb{R}^+$	Shannon entropy over legal predicate space
ν_{priv}	$[0, 1]$	Privilege coefficient (Def. 3.6)
$\mathcal{E}$	Source	Min-entropy source $H_{\min} \geq \alpha$
$\deg_{\text{MMR}}$	$\mathbb{N}$	Degree of MMR hash tree

Table D.9: Complete formal notation for ZK-NR protocol