

BDF Security Analysis

Threat Model, Trust Assumptions, and Formal Reductions for the Bio-Digital Footprint Protocol

Fair Food Data (FFD)

Version 1.0 · 24 September 2026

Abstract

This paper formalises the security guarantees of the Bio-Digital Footprint (BDF) protocol [Fair Food Data, 2026a]. It develops the threat model with explicit adversary typology, states the trust assumptions on which the protocol depends, and provides per tier security reductions that quantify what the protocol guarantees and against which adversary classes.

Five adversary types of increasing power are considered: a lone fraudster, a k of n coalition with internal coordination, a network adversary controlling a fraction of the underlying storage and routing layer, a polynomial time bounded computational adversary, and a coordinated cartel that compromises all attestors and the laboratory simultaneously. Eight trust assumptions are stated formally and grouped into cryptographic, network, social, and operational categories. Each tier of the protocol's evidence stack is then reduced to specific assumptions, with cryptographic guarantees at T_1 bounded by collision and signature security, statistical guarantees at T_2 derived from independence between attestors and calibrated detection thresholds, identifiability guarantees at T_3 conditional on reference base integrity, and Bayesian forensic bounds at T_4 conditional on laboratory accreditation.

The DREAM revocation mechanism is analysed quantitatively: a path of n hops in a network with honesty rate h admits successful revocation with probability bounded by $1 - (1 - h)^n$. The composition theorem establishes that per tier guarantees compose additively without weakening, and the residual cartel risk (Type 5) is shown to shrink under the Hardening Pathways through explicit reductions to hardware attestation, cryptoeconomic stake calibration, and zero knowledge soundness. The paper closes with a quantitative discussion of limits, including pre genesis attacks, side channels, and the adversarial assumptions that the protocol does not protect against.

The intended audience is registry assessors evaluating BDF based methodologies for ICVCM CCP labelling, peer reviewers, and protocol implementers seeking a rigorous foundation on which to build deployments.



Contents

1	Introduction and Scope	5
1.1	Relation to the BDF Main Paper	5
1.2	What This Paper Proves and What It Does Not	5
1.3	Notation	6
1.4	Audience and Intended Use	7
2	Adversary Typology	8
2.1	Type 1: Lone Fraudster	8
2.2	Type 2: k of n Coalition	8
2.3	Type 3: Network Adversary	8
2.4	Type 4: Computational Adversary	9
2.5	Type 5: Coordinated Cartel	9
2.6	Adversary Capability Matrix	10
3	Trust Assumptions	10
3.1	A1: Cryptographic Primitives at 128 bit Security	11
3.2	A2: Deterministic Canonicalisation	11
3.3	A3: Routing Diversity in the Storage Network	11
3.4	A4: Honesty Threshold Among Attestors	12
3.5	A5: Postage Economics Functional	12
3.6	A6: Anchoring Chain Security	13
3.7	A7: ISO 17025 Accreditation Integrity	13
3.8	A8: External Time Source Consensus	14
3.9	Assumption Summary Table	14
4	Tier 1 Cryptographic Floor: Security Reductions	15
4.1	Object Identity Uniqueness	15
4.2	Signature Unforgeability	15
4.3	Ownership Chain Consistency	15
4.4	Burn Non Reusability	16
4.5	Distribution Conservation	16
4.6	Probabilistic Fork Resistance	17
4.7	Summary of Tier 1 Guarantees	17
5	Tier 2 Telemetry Anomaly Detection: Statistical Guarantees	18
5.1	Conservation Invariant: Deterministic Detection under Independence	18
5.2	Envelope Tests: False Positive and False Negative Rates	19
5.3	Cross Vector Coherence under Independence	19
5.4	Adversarial Robustness: Adaptive vs Non Adaptive	20
5.5	Empirical Calibration: Pilot Data Reference	21
5.6	Summary of Tier 2 Guarantees	21
6	Tier 3 Spectroscopic Classification: Identifiability under Reference Base Integrity	22
6.1	Identifiability Guarantee	22

6.2	Reference Base Contamination Model	22
6.3	Slow Rolling Adversarial Training	23
6.4	Classifier Robustness under Cultivar Drift	24
6.5	Summary of Tier 3 Guarantees	24
7	Tier 4 Laboratory Forensic: Bayesian Posterior Bounds	25
7.1	Posterior Threshold	25
7.2	Dependence on Laboratory Integrity	25
7.3	Multi Laboratory Cross Checking	26
7.4	Sensitivity to Prior	26
7.5	Treatment of Ambiguous Verdicts	27
7.6	Summary of Tier 4 Guarantees	27
8	DREAM Revocation: Quantitative Guarantee	28
8.1	DREAM Construction Recap	28
8.2	Revocation Success Probability	28
8.3	Limits and Failure Modes	29
8.4	Auditable Erasure Events	29
8.5	Composition with the Cryptographic Floor	30
9	Composition Theorem	30
9.1	Statement of the Composition Theorem	30
9.2	Cartel Residual Risk Reduction	31
9.3	Adversary Coverage Matrix Update	32
9.4	Composition Bound Tightness	33
10	Hardening Pathway Reductions	33
10.1	Proplet Hardware Attestation	33
10.2	Cryptoeconomic Stake Calibration	34
10.3	Zero Knowledge Selective Verification	35
10.4	Composition of Hardening Pathways	36
11	Quantitative Limits	36
11.1	Cryptographic Security Level	37
11.2	Side Channels and Implementation Attacks	37
11.3	Pre Genesis Attacks	38
11.4	Network Majority Adversary	39
11.5	Economic and Legal Constraints	39
11.6	Summary of Limits	39
12	Conclusion	40
A	Proof Sketches	42
A.1	Auxiliary Lemmas	42
A.2	Full Sketch: Object Identity Uniqueness (Lemma 1)	42
A.3	Full Sketch: Signature Unforgeability (Lemma 2)	43

A.4	Full Sketch: Ownership Chain Consistency (Lemma 3)	43
A.5	Full Sketch: Distribution Conservation (Theorem 1)	43
A.6	Full Sketch: Composition Theorem (Theorem 3)	44
A.7	Full Sketch: Cartel Risk Bound (Theorem 6)	45
A.8	On Mechanisation	45
B	Adversary Capability Matrix	46
B.1	Adversary versus Trust Assumption Matrix	46
B.2	Adversary versus Tier Guarantee Matrix	47
B.3	Hardening Pathway Coverage Matrix	47
B.4	Quick Reference: Bounds at a Glance	47
	Document History	49

1 Introduction and Scope

The Bio-Digital Footprint (BDF) protocol [Fair Food Data, 2026a] provides verifiable records for transformed bioassets through a typed digital twin and a tiered attestation model. The main paper introduces the protocol’s structural elements (Bio-Digital Twin, attestation vectors, four detection tiers, three layer architecture) and gives an informal summary of the security guarantees. This paper develops the formal foundation: it states the threat model, the trust assumptions, and the per tier security reductions in a form suitable for registry assessment, peer review, and rigorous implementation.

1.1 Relation to the BDF Main Paper

The BDF main paper [Fair Food Data, 2026a] is the structural reference. This paper does not redefine the protocol; it formalises what the main paper describes. References of the form “Section X of the main paper” in this document point to [Fair Food Data, 2026a]. Companion papers covering deployment methodology [Fair Food Data, 2026b] and the open standards stack [Fair Food Data, 2026c] are referenced where their content interacts with the security guarantees.

The reader is assumed to be familiar with the structural elements of the main paper, in particular: the canonical object identity (H_O derived from beeson serialisation), the four attestation vectors $\vec{A}$, $\vec{L}$, $\vec{P}$, $\vec{R}$, the four tiers T_1 through T_4 , the Single Owner Chunk semantics on Swarm, and the burn and distribution operations on the Bio-Digital Twin lifecycle.

1.2 What This Paper Proves and What It Does Not

In scope. This paper develops:

- A typology of adversaries against the BDF protocol, ordered by increasing power (Section 2).
- A formal statement of eight trust assumptions on which the protocol depends (Section 3).
- Per tier security reductions, with proof sketches in the body and full sketches in Appendix A, for T_1 (Section 4), T_2 (Section 5), T_3 (Section 6), and T_4 (Section 7).

- A quantitative bound on DREAM revocation success in terms of network honesty rate and path length (Section 8).
- A composition theorem stating that per tier guarantees compose additively without weakening (Section 9).
- Reductions for the Hardening Pathways: Proplet hardware attestation, cryptoeconomic stake calibration, and zero knowledge soundness (Section 10).
- A discussion of quantitative limits and assumption boundaries (Section 11).

Out of scope. The following are explicitly outside the scope of this paper:

- Side channel attacks against signing devices or laboratory instruments. These depend on hardware specifics not fixed by the protocol.
- Implementation bugs in reference software. The reductions assume a faithful implementation of the canonical primitives.
- Adversaries who have broken the underlying cryptographic primitives below 128 bit security. The protocol inherits the security of its primitives and does not strengthen them.
- Pre genesis attacks where the adversary corrupts the physical reality before any signed evidence exists. This residual gap is acknowledged and bounded in Section 11.
- Economic attacks against the underlying anchoring chains beyond the threshold honesty fraction.
- Legal and contractual mechanisms; the paper analyses cryptographic and statistical guarantees, not contractual remedies.

Proof depth. Proof sketches in the body present the structure of each reduction with sufficient detail for a cryptographically literate reader to reconstruct the full argument. Full sketches with auxiliary lemmas appear in Appendix A. Fully formal proofs are deferred to a successor paper or to peer review feedback.

1.3 Notation

The following notation is used throughout. Symbols are reused from the BDF main paper where applicable.

O

A Bio-Digital Twin (BDT), a typed canonical object aggregating attestation vectors and external attestations.

H_O

The canonical hash of the BDT, derived as $H_O = \mathcal{H}_{\text{BMT}}(\text{Serialise}(O))$ under beeson canonicalisation [Tron et al., 2026].

$\vec{A}, \vec{L}, \vec{P}, \vec{R}$

The four attestation vectors: Agronomic, Logistic, Process, Recipe.

$\vec{E}$

An External Attestation issued by an ISO 17025 accredited laboratory.

$T_i, i \in \{1, 2, 3, 4\}$

The protocol’s evidence tiers.

σ_i

Signature by signer A_i in the ownership chain.

$\text{Sign}_a, \text{Verify}_a$

Signature and verification operations under the keypair of party a .

$\mathcal{A}$

An adversary against the protocol; subscripted by type when needed ($\mathcal{A}_1$ to $\mathcal{A}_5$).

λ

The cryptographic security parameter (typically $\lambda = 128$).

$\text{negl}(\lambda)$

A function that vanishes faster than any inverse polynomial in λ .

h

Honest neighbourhood fraction of the underlying storage network.

n

Number of hops in a DREAM path.

θ_i

Detection threshold or posterior bound at tier T_i .

$\text{Adv}_{\mathcal{A}}^{\text{prim}}$

The adversary’s advantage against primitive prim (collision resistance, EUF-CMA, etc.).

1.4 Audience and Intended Use

This paper has three primary audiences.

Registry assessors. Assessors evaluating BDF based methodologies for ICVCM Core Carbon Principles labelling [Integrity Council for the Voluntary Carbon Market, 2024] and for cross registry recognition (Verra, Gold Standard, ICR) require explicit threat models and quantitative guarantees against the assessment criteria. This paper provides the formal grounds on which an assessor can map BDF protocol guarantees to specific principles and benchmark them against alternative approaches.

Peer reviewers. Cryptographers, distributed systems researchers, and food traceability researchers reviewing the BDF protocol require formal statements with proof sketches in standard idiom. The structure of this paper follows the convention of cryptographic security analyses: assumption tables, lemma proof reduction chains, and explicit composition arguments.

Protocol implementers. Engineers building deployments of the BDF protocol require a clear list of trust assumptions (with the consequences of violating each), the sensitivity of the guarantees to parameter choices (cohort sizes, honesty fractions, detection thresholds), and the trade offs between simpler operational profiles and stronger cryptographic profiles. This paper provides those.

The paper is self contained for these audiences with the exception of the structural definitions in the main paper and the bibliography of cryptographic primitives, which are cited rather than restated.

2 Adversary Typology

The BDF protocol is analysed against five adversary types of increasing power. Each type captures a distinct attack surface and corresponds to a class of real world threats observed in supply chain fraud, voluntary carbon market manipulation, and adversarial trade in the integrated systems the protocol is designed to support. The types are not strictly nested; they overlap on capabilities but differ in coordination scope and computational constraints.

2.1 Type 1: Lone Fraudster

Definition 1 ($\mathcal{A}_1$ Lone Fraudster). *A single corrupted attester with control of one signing key and the ability to issue arbitrary signed payloads under that key. The adversary has no coordination with other attestors and no privileged access to the storage or routing layer.*

The lone fraudster is the canonical low effort adversary. Examples include: a farm operator inflating origin masses, a transport operator forging delivery times, a processing plant fabricating cycle telemetry. The adversary's capability is bounded to issuing signed objects under one key. Detection by other attestors at independent stages is the primary defence (Section 5).

The adversary's success against the protocol depends on whether the fabricated payload escapes detection by the cross referencing tests at T_2 and the spectroscopic and forensic checks at T_3 and T_4 . The reductions in Section 5 establish that $\mathcal{A}_1$ is detected with overwhelming probability under standard independence assumptions among the remaining attestors.

2.2 Type 2: k of n Coalition

Definition 2 ($\mathcal{A}_2(k, n)$ Coalition). *A coalition of k corrupted attestors out of n total attestors with internal coordination. The coalition members can produce coherent narratives across their respective vectors but cannot influence the remaining $n - k$ honest attestors.*

The coalition adversary captures multi party fraud where stage operators conspire to fabricate a coherent supply chain story. Examples include: a vertically integrated operator forging $\vec{A}$ at the farm, $\vec{L}$ in transit, and $\vec{P}$ at the plant simultaneously; or a third party logistics operator colluding with a corrupted laboratory to produce false External Attestations.

The protocol's T_2 detection capability against $\mathcal{A}_2$ depends on the size of the honest fraction $n - k$ and the independence of their vectors. Where the cross referencing tests of Section 5 require at least one independent honest attester in the relevant cross product, $\mathcal{A}_2$ is detected when $k < n$ and the honest attester sits in the intersected stage. The protocol does not in general protect against $\mathcal{A}_2(n, n)$, which is captured by Type 5.

2.3 Type 3: Network Adversary

Definition 3 ($\mathcal{A}_3(\rho)$ Network Adversary). *An adversary controlling a fraction $\rho \in [0, 1]$ of the underlying content addressed storage and routing layer (typically Swarm). The adversary can selectively*

deny service to retrieval requests routed through compromised neighbourhoods, attempt to censor or delay specific chunks, or attempt Sybil takeover of a neighbourhood.

The network adversary attacks the protocol’s data availability layer rather than its cryptographic content. Examples include: a state actor attempting to suppress evidence of a fraud after issuance, a competitor attempting to deny a verifier access to the public record at a critical moment, a Sybil attack inflating the apparent honesty fraction of a neighbourhood.

Resistance against $\mathcal{A}_3$ depends on the network’s routing diversity and on the honest neighbourhood fraction h . Section 8 quantifies the trade off for the DREAM revocation case explicitly. The protocol does not protect against an $\mathcal{A}_3$ that captures a majority of the network ($\rho > 1 - h$ for the relevant operations); this limit is documented in Section 11.

2.4 Type 4: Computational Adversary

Definition 4 ($\mathcal{A}_4$ Computational Adversary). *A polynomial time bounded adversary attempting to break the cryptographic primitives directly: collision finding on the hash function, signature forgery without the private key, predictability of nonces or randomness, or breaking the discrete logarithm assumption underlying the signature scheme.*

The computational adversary models the worst case attack on the protocol’s cryptographic floor. The protocol’s resistance is inherited from the standard 128 bit security level of its primitives (SHA–256 or equivalent BMT, ECDSA on secp256k1, AES based authenticated encryption for payload confidentiality). The reductions in Section 4 bound $\mathcal{A}_4$ ’s advantage by the corresponding primitive advantages.

The protocol does not strengthen the underlying primitives. If a primitive falls (post quantum break, novel cryptanalytic discovery), the protocol’s T_1 guarantees fall with it. The protocol does provide a clean migration path through the canonical hash type field of beeson, allowing a future deployment to migrate to post quantum signatures without breaking the structural integrity of historical records.

2.5 Type 5: Coordinated Cartel

Definition 5 ($\mathcal{A}_5$ Coordinated Cartel). *The simultaneous compromise of all n attestors and the laboratory issuing External Attestations, plus a sufficient fraction of the underlying network to suppress dispute investigations. $\mathcal{A}_5$ subsumes all five corruption types but is constrained to remain polynomial time bounded.*

The coordinated cartel is the residual adversary against which the BDF protocol does not provide cryptographic guarantees in its base configuration. If every attestor signing a vector is corrupt and the laboratory is corrupt, no signed evidence on the public record contradicts the cartel’s narrative. The cartel can produce a coherent fabricated supply chain whose every signed payload verifies, every cross referencing test passes, and every laboratory attestation supports the false claim.

The Hardening Pathways of Section 10 progressively shrink this residual risk. Proplet hardware attestation shifts the trust principal from operator to hardware, requiring physical compromise of every device. Cryptoeconomic mechanisms make coordinated fraud expected to lose money under stake

calibration. Zero knowledge proofs allow public verification of properties without revealing the full payload, removing some of the cartel’s information asymmetry.

The protocol explicitly acknowledges $\mathcal{A}_5$ as outside the base guarantees. A deployment that requires resistance against $\mathcal{A}_5$ shall implement the Hardening Pathways in proportion to the value at stake.

2.6 Adversary Capability Matrix

Table 1 summarises the adversary capabilities against the four protocol tiers. The full matrix with per assumption breakdowns appears in Appendix B.

Table 1: Adversary capabilities against BDF tiers, summary form. P indicates the tier provides a guarantee against the adversary, p a partial or conditional guarantee, N no guarantee in the base configuration.

Adversary	T_1	T_2	T_3	T_4
$\mathcal{A}_1$ Lone fraudster	P	P	P	P
$\mathcal{A}_2$ Coalition $k < n$	P	p	p	P
$\mathcal{A}_2$ Coalition $k = n$	P	N	N	p
$\mathcal{A}_3$ Network $\rho < 1 - h$	P	P	P	P
$\mathcal{A}_3$ Network $\rho > 1 - h$	p	p	p	p
$\mathcal{A}_4$ Computational	P	P	P	P
$\mathcal{A}_5$ Coordinated cartel	p	N	N	N

Reading the matrix. The lone fraudster is fully covered at every tier. Coalitions short of full capture are detected at T_2 and T_3 statistically and forensically at T_4 when the laboratory is honest. The network adversary below the network majority is fully resisted at the data availability layer. The computational adversary is fully resisted under the standard 128 bit primitive assumptions. The coordinated cartel is the principal residual risk and is the target of the Hardening Pathway reductions in Section 10.

The full breakdown in Appendix B maps each cell to the specific trust assumptions invoked, providing a self contained reference for assessment.

3 Trust Assumptions

The protocol’s security guarantees rest on eight trust assumptions, formally stated in this section. The assumptions are grouped into four categories: cryptographic (A1, A2, A8), network (A3), social (A4), and operational (A5, A6, A7). Each assumption is stated with its consequence: the guarantees that follow if it holds and the guarantees that fall if it is violated.

A deployment shall declare which assumptions it accepts in its baseline configuration and which it strengthens through Hardening Pathways. A registry assessor evaluating a deployment shall confirm the assumptions and trace any deployment specific deviation.

3.1 A1: Cryptographic Primitives at 128 bit Security

Assumption 1 (A1). *The cryptographic primitives instantiated by the protocol provide at least 128 bit security against a polynomial time adversary $\mathcal{A}_4$. Specifically:*

- *The hash function $\mathcal{H}_{\text{BMT}}$ is collision resistant, with the adversary’s advantage $\text{Adv}_{\mathcal{A}_4}^{\text{coll}}(\lambda) = \text{negl}(\lambda)$ for $\lambda = 128$.*
- *The signature scheme is existentially unforgeable under chosen message attack (EUF-CMA), with $\text{Adv}_{\mathcal{A}_4}^{\text{euf-cma}}(\lambda) = \text{negl}(\lambda)$ for $\lambda = 128$.*
- *The authenticated encryption used for payload confidentiality is IND-CCA secure with $\text{Adv}_{\mathcal{A}_4}^{\text{ind-cca}}(\lambda) = \text{negl}(\lambda)$ for $\lambda = 128$.*

Default instantiation. The reference deployment uses SHA–256 within a binary Merkle tree for $\mathcal{H}_{\text{BMT}}$ [Merkle, 1987, NIST, 2015], ECDSA on secp256k1 for signatures [Johnson et al., 2001, NIST, 2023], and AES-256 in GCM mode for authenticated encryption.

If A1 fails. The cryptographic floor at T_1 collapses entirely. Object identities can be aliased (collision), signatures can be forged (EUF-CMA break), or payloads can be selectively decrypted without keys (IND-CCA break). The protocol provides no remedy at the protocol layer if A1 fails; remediation is by primitive migration and re-anchoring of the historical record under the new primitives.

3.2 A2: Deterministic Canonicalisation

Assumption 2 (A2). *The canonical serialisation function Serialise defined by beeson [Tron et al., 2026] is deterministic: for any typed object O , $\text{Serialise}(O)$ produces a unique byte sequence, and any two implementations of Serialise that pass the conformance suite produce identical output for the same input.*

Justification. Beeson’s canonical encoding embeds the schema in the representation, fixes the field ordering by name lexicographic order, normalises numeric encodings, and bans non canonical UTF-8 sequences. The conformance suite covers edge cases (negative zero, NaN representations, leading zeros in big integers).

If A2 fails. The canonical hash H_O becomes implementation dependent. Two parties with the same logical object compute different hashes; the public record fragments. The mitigation in this case is to pin the implementation version in the canonical hash type field and require all implementers to ship the same conformance suite. The protocol does not currently provide a graceful degradation path for A2 failure.

3.3 A3: Routing Diversity in the Storage Network

Assumption 3 (A3). *The underlying content addressed storage network exhibits sufficient routing diversity such that any retrieval request from a benign node has at least one path through honest neighbourhoods*

to the destination chunk. Quantitatively: the honest neighbourhood fraction h satisfies $h > h_{\min}$, where $h_{\min}$ depends on the operation analysed and the adversary's budget.

Justification. Swarm [Trón, 2024] achieves routing diversity through Kademlia style XOR distance [Maymounkov and Mazières, 2002] with neighbourhood replication. The honesty rate h is empirically estimated from the network operator base; for production Swarm at the time of writing, $h \geq 0.9$ is a reasonable default for the analyses in this paper. Sybil resistance is provided by the proof of work admission to the network and by the postage stamp economic constraint [Trón, 2024].

If A3 fails. The protocol's data availability layer fails. Specifically, DREAM revocation (Section 8) loses its quantitative bound. Retrieval of older chunks becomes unreliable. The protocol's T_1 structural guarantees on signed objects already retrieved remain intact (they are cryptographic, not network), but new evidence cannot be reliably published or fetched.

3.4 A4: Honesty Threshold Among Attestors

Assumption 4 (A4). For any attestation vector $\vec{V} \in \{\vec{A}, \vec{L}, \vec{P}, \vec{R}\}$ involved in a cross referencing test, at least one of the cross referenced vectors is signed by an honest attestor independent of the others.

Justification. The protocol does not assume a fixed k of n honesty bound across all attestors. It assumes pairwise independence between cross referenced stages: the field operator who signs $\vec{A}$ is independent of the transport operator who signs $\vec{L}$ is independent of the plant SCADA which signs $\vec{P}$. Where the independence is structural (different organisations, different jurisdictions, different liability), A4 is empirical. Where the independence is contractual (vertically integrated supply chains), A4 reduces to a coalition adversary problem and Section 5 treats the corresponding cases.

If A4 fails. The protocol's statistical detection at T_2 fails for the affected vectors. The cross referencing tests that depend on independence cease to provide guarantees against $\mathcal{A}_2(n, n)$. The Hardening Pathways (Section 10) provide cryptographic and economic substitutes for the lost statistical independence: hardware attestation gives independence at the device level, cryptoeconomic stakes give independence through diverging financial interests.

3.5 A5: Postage Economics Functional

Assumption 5 (A5). The postage stamp mechanism of the underlying storage network [Trón, 2024] is functional: chunks with valid stamps remain available until their stamp expires, and chunks with expired stamps are eligible for garbage collection. The economic cost of pinning a chunk grows linearly with the chunk count and time, ensuring sustainable storage.

Justification. Swarm's postage stamps are economically calibrated to the cost of node storage and bandwidth. A deployment paying postage for its anchored data secures availability for the postage period. A deployment that ceases to renew postage signals retention non interest, equivalent to deletion

at the storage layer. This is the structural mechanism the BDF protocol leverages for right to erasure compliance under GDPR Article 17.

If A5 fails. Persistence of payloads becomes unreliable. The protocol can no longer rely on retrieving older $\vec{V}$ vectors, breaking continuity of the verification record. Mitigation: a deployment may mirror anchored data to a centrally administered store for the duration of its crediting period, bridging the postage failure case at the cost of additional centralisation.

3.6 A6: Anchoring Chain Security

Assumption 6 (A6). *The public chain to which ownership transfers and burn events are committed satisfies the standard blockchain liveness and persistence properties: an honest fraction $f > \frac{1}{2}$ of the chain's consensus participants ensures that valid transactions are eventually included and remain in the canonical chain.*

Justification. The reference deployment anchors to Ethereum or to Gnosis Chain, both of which satisfy A6 under their respective consensus mechanisms. A deployment may select alternative anchoring chains; the assumption transfers as long as the chosen chain provides equivalent guarantees.

If A6 fails. Ownership transitions can be censored or reordered, double signing becomes hard to detect, and burn events can be retroactively erased from the chain's view. The protocol's T_1 guarantees on already finalised transactions remain (they are cryptographic, not consensus dependent), but new transactions become unreliable. Mitigation: a deployment may anchor to multiple independent chains and require co-anchoring as a precondition for finality, at the cost of additional latency.

3.7 A7: ISO 17025 Accreditation Integrity

Assumption 7 (A7). *Laboratories issuing External Attestations are accredited under ISO/IEC 17025:2017 [International Organization for Standardization, 2017] by an ILAC MRA member accreditation body [International Laboratory Accreditation Cooperation, 2024]. The accreditation process provides reasonable assurance that the laboratory's measurement processes are competent, traceable, and free of fraud.*

Justification. ILAC MRA mutual recognition extends across most jurisdictions where the protocol is deployed. Accreditation includes periodic on site assessment, proficiency testing, traceability of measurement standards, and mandatory disclosure of impartiality conflicts. This is a strong but imperfect external check.

If A7 fails. T_4 External Attestations from the failing laboratory carry no juridical weight. Detection: the verifier can request external proficiency testing or cross check against a second laboratory. Where multiple laboratories are independently accredited and concur, the residual risk that all are simultaneously corrupted is part of the cartel adversary $\mathcal{A}_5$.

3.8 A8: External Time Source Consensus

Assumption 8 (A8). *A consensus on absolute time is available within a bounded skew Δ_t (typically $\Delta_t \leq 60$ seconds) across all signing entities, provided by a combination of GPS time, NTP, and the timestamp of the anchoring chain.*

Justification. The protocol uses absolute time in attestation vectors (recovery date, segment start, cycle start, etc.). A skew below the cadence of T_2 envelope tests is acceptable. Skew above the cadence allows a fraudster to backdate or postdate signed payloads, undermining the cross referencing tests.

If A8 fails. The temporal anchors of attestation vectors lose meaning. Mitigation: a deployment may require at minimum two independent time sources per signed payload (e.g. GPS plus the anchoring chain’s block timestamp) and reject payloads whose two sources diverge beyond Δ_t . This degrades the protocol gracefully under partial time source failure.

3.9 Assumption Summary Table

Table 2 summarises the assumptions, their categories, and the tiers they support.

Table 2: Summary of trust assumptions and the tiers they support. A check indicates that the assumption is required for that tier’s guarantees in the base configuration.

Id	Assumption	Category	T_1	T_2	T_3	T_4
A1	Cryptographic primitives 128 bit	Cryptographic	✓	✓	✓	✓
A2	Deterministic canonicalisation	Cryptographic	✓	✓	✓	✓
A3	Routing diversity in storage	Network	✓	✓	✓	✓
A4	Honesty among attestors	Social		✓	✓	✓
A5	Postage economics functional	Operational	✓	✓	✓	✓
A6	Anchoring chain security	Operational	✓	✓	✓	✓
A7	ISO 17025 accreditation	Operational				✓
A8	External time source consensus	Cryptographic	✓	✓	✓	✓

Reading the table. A1, A2, A3, A5, A6, and A8 are required at every tier, reflecting their structural role in the protocol’s foundation. A4 is required from T_2 upward, where statistical detection depends on independence among attestors. A7 is required only at T_4 , where it underpins the juridical weight of External Attestations.

Hardening pathway substitutes. The Hardening Pathways of Section 10 provide substitutes for A4 (and partially for A7). Proplet hardware attestation substitutes hardware level independence for organisational independence among attestors, mitigating A4 violations. Multiple independent laboratories at T_4 mitigate A7 violations through redundancy.

4 Tier 1 Cryptographic Floor: Security Reductions

The Tier 1 guarantees of the BDF protocol are structural: they depend only on the cryptographic primitives and the canonical encoding, not on the truthfulness of underlying physical measurements. This section reduces each T_1 guarantee to a primitive assumption from Section 3, with proof sketches for the lemmas. Full sketches are in Appendix A.

4.1 Object Identity Uniqueness

Lemma 1 (Object Identity Uniqueness). *Under assumptions A1 (collision resistance of $\mathcal{H}_{\text{BMT}}$) and A2 (deterministic canonicalisation), for any two distinct typed objects $O_1 \neq O_2$ that are well formed under their declared schemas, the probability that $H_{O_1} = H_{O_2}$ is bounded by $\text{Adv}_{\mathcal{A}_4}^{\text{coll}}(\lambda) = \text{negl}(\lambda)$ for $\lambda = 128$.*

Proof sketch. By A2, $\text{Serialise}(O_1) \neq \text{Serialise}(O_2)$ since canonicalisation is injective on distinct typed objects. By A1, the probability that a polynomial time adversary finds two distinct byte sequences with the same hash is bounded by the primitive's collision advantage. The composition is straightforward: two distinct objects produce two distinct serialisations produce two distinct hashes except with probability $\text{Adv}_{\mathcal{A}_4}^{\text{coll}}(\lambda)$. $\square$

Consequence. Object identities are unique with overwhelming probability against $\mathcal{A}_4$. The lemma underpins all higher tier guarantees that reference H_O as the unique identifier of a BDT.

4.2 Signature Unforgeability

Lemma 2 (Signature Unforgeability on BDT Operations). *Under assumption A1 (EUF-CMA security of the signature scheme), for any party a with public key pk_a and private key sk_a , no polynomial time adversary $\mathcal{A}_4$ without access to sk_a produces a tuple (m, σ) with $\text{Verify}_{pk_a}(m, \sigma) = \text{accept}$ for any message m not previously signed by a , except with probability $\text{Adv}_{\mathcal{A}_4}^{\text{euf-cma}}(\lambda) = \text{negl}(\lambda)$.*

Proof sketch. Direct reduction to the EUF-CMA security of the underlying signature scheme [Johnson et al., 2001, Bellare and Rogaway, 1996]. The protocol does not modify the signing algorithm; it embeds canonical messages into the EUF-CMA experiment. An adversary forging a BDT operation signature is a forger in the standard EUF-CMA game. $\square$

Consequence. An adversary $\mathcal{A}_1$ who corrupts party b cannot impersonate a different party a to produce signed payloads under a 's key. The tampering surface is bounded to the keys the adversary controls.

4.3 Ownership Chain Consistency

Lemma 3 (Ownership Chain Consistency). *Under A1 (signature unforgeability) and A6 (anchoring chain liveness), for any BDT O with ownership sequence $A_0, A_1, \dots, A_n$ committed to the public chain*

via signatures $\sigma_0, \sigma_1, \dots, \sigma_{n-1}$, the chain is internally consistent (every σ_i verifies against A_i 's public key on the message $H_O \parallel A_{i+1}$) except with probability $n \cdot \text{Adv}_{\mathcal{A}_4}^{\text{euf-cma}}(\lambda) = \text{negl}(\lambda)$.

Furthermore, any conflicting signature σ'_i from A_i that signs $H_O \parallel A'_{i+1}$ with $A'_{i+1} \neq A_{i+1}$ is publicly detectable as a double signing event by any verifier inspecting the public chain.

Proof sketch. The first part follows by union bound from Lemma 2 applied to each step. Each σ_i is unforgeable, so no adversary without sk_{A_i} can produce a valid alternative successor. The second part follows from A6: both σ_i and σ'_i are anchored to the public chain, both verify, and both bear A_i 's signature on conflicting successor commitments. Any verifier reading the chain observes both and detects the double signing publicly. $\square$

Consequence. The ownership chain is unforgeable along its honest segments. A corrupted owner who attempts to fork the chain by signing two different successors is publicly detected.

4.4 Burn Non Reusability

Lemma 4 (Burn Non Reusability). *Under A1 and A6, once a BDT O has been burned (its ownership chain terminates at the designated transformation address TransformAddr), no further valid ownership transfer of O is accepted. Any attempted post burn signature $\sigma' = \text{Sign}_{A_n}(H_O \parallel A')$ for $A' \neq \text{TransformAddr}$ verifies as a double signing by A_n and is rejected by verifiers.*

Proof sketch. The post burn transfer is a conflicting signature by the same owner A_n on the same object H_O , exactly the case covered by the second part of Lemma 3. The protocol's verifier rejects double signings by construction. The cryptographic floor inherits the rejection from the ownership chain consistency. $\square$

Consequence. A consumed input cannot be re-spent. Combined with the transformation event semantics of the main paper [Fair Food Data, 2026a], this prevents the inflation attack where a single input is claimed by multiple downstream products.

4.5 Distribution Conservation

Theorem 1 (Distribution Conservation). *Under A1, A2, and A6, for any finalised BDT O_p subjected to a distribution event that instantiates package BDTs $\{O_{p,1}, \dots, O_{p,M}\}$ and an optional remainder twin R , the metadata quantity invariant*

$$\sum_{k=1}^M q(O_{p,k}) + q(R) = q(O_p)$$

is enforced at the protocol layer. Any distribution event whose outputs violate the invariant is rejected by verifiers.

Proof sketch. The distribution event is itself a signed object whose payload includes the parent BDT identifier, the package outputs, and the remainder. The verifier reading the event from the public chain (A6) recomputes the sum on the right hand side from the signed canonical representations of the outputs (A2) and compares to the parent's declared $q(O_p)$. Equality is structural; inequality is a deterministic predicate that the verifier evaluates and rejects. The signing party cannot fabricate a coherent distribution event with imbalanced outputs because their own signature commits to the canonical representation of each output, and the canonical representation includes the metadata $q(\cdot)$. $\square$

Consequence. Mass (or volume, or count, depending on the schema's declared quantity) is conserved through distribution events. The non fungible token identity of each output is preserved while the metadata quantity invariant is enforced through the verifier's mechanical check.

4.6 Probabilistic Fork Resistance

Proposition 1 (Probabilistic Fork Resistance). *Under A1, A6, and A3 with honest neighbourhood fraction h , the probability that a fork in the ownership chain (two conflicting branches with valid signatures) escapes detection by an arbitrary verifier inspecting the public record decreases exponentially in the number of independent verifiers.*

Specifically, if V verifiers each independently inspect the chain with verification probability p_v , the probability that all fail to observe a fork is bounded by $(1 - p_v)^V$ where p_v is determined by the verifier's reach into the storage network.

Proof sketch. A fork is a pair of conflicting signatures by the same party. By Lemma 3, the fork is structurally observable from the public chain. A verifier observes the fork with probability p_v that depends on its reach into the storage layer (a function of A3). Independent verifiers' detection events are independent under the network topology assumption, so the probability that all V verifiers fail is at most $(1 - p_v)^V$. As V grows, this probability vanishes. $\square$

Consequence. Forks are detected with high probability when the protocol is verified by a diverse set of independent verifiers. The protocol does not rely on a single verifier; its security against forks improves with the number of independent observers.

4.7 Summary of Tier 1 Guarantees

Table 3 summarises the T_1 reductions.

Composition. The T_1 guarantees compose: chain consistency depends on signature unforgeability, burn non reusability inherits from chain consistency, distribution conservation depends on canonicalisation and chain anchoring. The composition theorem of Section 9 formalises this composition and shows it does not amplify the per primitive advantages beyond a polynomial factor.

Table 3: Summary of Tier 1 guarantees and their reductions.

Guarantee	Assumptions	Bound
Object identity uniqueness	A1, A2	$\text{Adv}_{\mathcal{A}_4}^{\text{coll}} = \text{negl}(\lambda)$
Signature unforgeability	A1	$\text{Adv}_{\mathcal{A}_4}^{\text{euf-cma}} = \text{negl}(\lambda)$
Ownership chain consistency	A1, A6	$n \cdot \text{Adv}_{\mathcal{A}_4}^{\text{euf-cma}}$ for chain length n
Burn non reusability	A1, A6	Inherited from chain consistency
Distribution conservation	A1, A2, A6	Structural; deterministic verifier check
Fork resistance (probabilistic)	A1, A3, A6	$(1 - p_v)^V$ in V independent verifiers

Adversary coverage. All T_1 guarantees hold against $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3$ with $\rho < 1 - h$, and $\mathcal{A}_4$. The cartel adversary $\mathcal{A}_5$ does not violate T_1 : even a fully corrupted set of attestors cannot produce a coherent record that violates structural integrity. The cartel’s success is in fabricating a structurally valid but factually false record; T_1 does not detect this, but the higher tiers do under specific assumption regimes.

5 Tier 2 Telemetry Anomaly Detection: Statistical Guarantees

The Tier 2 detectors operate over the canonicalised payloads of signed attestation vectors. Their guarantees are statistical rather than cryptographic: they bound the probability that an adversary’s fraudulent payload escapes detection given the deployment’s calibrated thresholds. This section formalises the per detector guarantees, characterises adversarial robustness, and quantifies how cohort sizes and independence assumptions translate to detection power.

5.1 Conservation Invariant: Deterministic Detection under Independence

Lemma 5 (Conservation Invariant Determinism). *Let $\vec{V}_a$ and $\vec{V}_b$ be two independent attestation vectors signed by independent attestors a and b , jointly committing to a shared metadata quantity q (e.g. mass at origin and mass at destination of a logistic segment). Under A4 (at least one of a, b is honest), the conservation predicate $|q_a - q_b|/q_a \leq \tau$ for threshold τ either holds (verdict *pass*) or is violated (verdict *warn* or *fail*) deterministically. A coalition of all signers committing to a coherent fabricated q is detectable only by a vector outside the coalition.*

Proof sketch. The signed quantities q_a and q_b are committed values in canonical form (Lemma 1). The verifier reads both, computes the predicate, and outputs the deterministic verdict. The detection failure mode is precisely the case where both signers are corrupt and commit to the same fabricated q ; this is excluded by A4. $\square$

Consequence against $\mathcal{A}_1$. The lone fraudster who signs only $\vec{V}_a$ cannot fabricate q_a that matches an honest q_b except by chance. The detector flags any genuine discrepancy.

Consequence against $\mathcal{A}_2(k, n)$. A coalition that includes both signers of a cross referenced pair can fabricate matching $q_a = q_b$ at coalition coordinated values. The conservation invariant then passes, but the absolute value of q may still be detected by a vector outside the coalition (e.g. the Process Vector signed by the SCADA system referencing the same lot's expected input mass).

5.2 Envelope Tests: False Positive and False Negative Rates

Definition 6 (Envelope Test). *For a parameter X with declared envelope $[X_{\min}, X_{\max}]$ in the Recipe Vector $\vec{R}$, the envelope test on a Process Vector observation X_{obs} returns *pass* if $X_{\min} \leq X_{\text{obs}} \leq X_{\max}$ and *fail* otherwise. A windowed variant (*warn band*) admits at most w consecutive readings outside the envelope before escalation.*

Proposition 2 (Envelope Test False Positive Rate). *Assume the parameter X in honest cycles follows a distribution $\mathcal{D}_X$ with cumulative distribution function F_X . The false positive rate (probability of a *warn* or *fail* verdict on an honest cycle) is*

$$\text{FPR} = 1 - F_X(X_{\max}) + F_X(X_{\min}^-).$$

Proof sketch. Direct computation from the distribution's tail probabilities. The envelope test fails iff the observation is in the tail. $\square$

Proposition 3 (Envelope Test False Negative Rate). *Assume the adversary's fabricated payload follows a distribution $\mathcal{D}_X^{\text{adv}}$ with c.d.f. F_X^{adv} . The false negative rate (probability of a *pass* verdict on an adversarial cycle) is*

$$\text{FNR} = F_X^{\text{adv}}(X_{\max}) - F_X^{\text{adv}}(X_{\min}^-).$$

Proof sketch. Adversary succeeds iff its fabricated value lies inside the envelope. The probability is the c.d.f. mass within the envelope under $\mathcal{D}_X^{\text{adv}}$. $\square$

Calibration cohort sizing. The envelope $[X_{\min}, X_{\max}]$ is calibrated against an empirical cohort. Standard concentration arguments (Hoeffding, Chernoff) bound the deviation of the empirical envelope from the true distribution. For an honest cohort of size N and target FPR α , the empirical envelope at level α has confidence band of order $O(\sqrt{\log(1/\delta)/N})$ at confidence $1 - \delta$. Thus a calibration cohort of $N \geq 100$ honest cycles provides envelope confidence within a few percent of the true distribution at standard confidence levels.

Adversarial calibration risk. If the adversary participates in the calibration cohort, the empirical envelope shifts toward the adversarial distribution. The protocol mitigates this through reference cohort isolation (Section 6) and through external validation of the cohort.

5.3 Cross Vector Coherence under Independence

Definition 7 (Cross Vector Coherence Test). *Given two attestation vectors $\vec{V}_a$ and $\vec{V}_b$ committing to overlapping or related quantities (e.g. logistic latency Δt_f in $\vec{L}$ and observed mass loss Δm in $\vec{P}$), the*

cross coherence test compares a derived quantity $f(\vec{V}_a, \vec{V}_b)$ against an expected value or band derived from physical or operational constraints.

Proposition 4 (Cross Vector Detection under Independence). *Under A4 (signer a and signer b are independent), an adversary corrupting only one of the two signers cannot guarantee that their fabricated value agrees with the honest signer’s value on the cross referenced quantity, except with probability bounded by the natural variability of the honest distribution.*

Specifically, if the honest cross quantity follows $\mathcal{D}_f$ with variance σ_f^2 around the predicted mean, the adversary’s success probability against a band test of width $\pm k\sigma_f$ is bounded by $\Phi(k) - \Phi(-k) \leq 1 - 2\text{negl}(k)$ for $k \rightarrow \infty$ where Φ is the standard normal CDF.

Proof sketch. The adversary controls one of the two values. The other is sampled from $\mathcal{D}_f$ by the honest signer. The adversary’s best strategy is to commit to the predicted mean of the honest distribution. The honest sample then concentrates around the mean by $\mathcal{D}_f$ ’s tails; the band test passes with probability bounded by the distribution’s mass within $\pm k\sigma_f$. $\square$

Consequence. The cross vector test gains its detection power from the independence of the two signers and the natural variability of the cross referenced quantity. As the band width k tightens or as the distribution’s tails are heavier, the adversary’s success probability decreases.

5.4 Adversarial Robustness: Adaptive vs Non Adaptive

The Tier 2 detectors are designed against non adaptive adversaries who do not observe the detection thresholds before committing their fabricated payloads. In practice, the thresholds are public after deployment (Section 5.2), making adaptive adversaries a realistic concern.

Proposition 5 (Adaptive Adversary Margin). *An adaptive adversary $\mathcal{A}_{1,\text{adp}}$ who knows the detection thresholds $[X_{\min}, X_{\max}]$ in advance can guarantee passage of the envelope test by committing to any value within the envelope. The protocol’s defence against adaptive adversaries shifts to:*

1. *Cross vector coherence tests, which the adversary cannot trivially satisfy without controlling all referenced vectors.*
2. *External Attestations at T_4 , which compare independent laboratory measurements against the signed payload.*
3. *Cohort level pattern detection: an adaptive adversary fabricating envelope passing values across many lots may produce a degenerate distribution within the envelope (e.g. all values clustered at the envelope mean) that is itself anomalous against the honest distribution’s natural variance.*

Calibration cohort detection of degenerate distributions. The protocol’s reference distribution $\mathcal{D}_X^{\text{honest}}$ is established from an honest calibration cohort. Subsequent observations are tested both against the envelope and against the distributional shape via a goodness of fit test (Kolmogorov-Smirnov or equivalent). An adaptive adversary clustering values at the mean produces a degenerate distribution that fails the goodness of fit test even when individual values pass the envelope. This is the protocol’s principal defence against adaptive adversaries within T_2 ; deeper defence is provided

by T_3 (which trains classifiers on the joint distribution of multiple parameters) and by T_4 (which is independent of the signed payload).

5.5 Empirical Calibration: Pilot Data Reference

The pilot reported in the BDF main paper [Fair Food Data, 2026a] provides empirical calibration data for the conservation invariant. Across fifteen baskets, the cohort mean of the field versus facility weight discrepancy is $\bar{\delta} = +11\%$, classified as `warn` under the methodology’s 5% to 15% band threshold (Section 5).

The verdict is signed and anchored to the pilot’s BDF record. A second verifier replaying the conservation invariant detector against the same signed inputs arrives at the identical `warn` verdict. The empirical illustration grounds the abstract statistical guarantees in observable behaviour: the detector is operational, the verdict is reproducible, and the corrective action (calibration scaling factor $\kappa = 1.110$) is documented in the methodology’s worked example [Fair Food Data, 2026b].

5.6 Summary of Tier 2 Guarantees

Table 4 summarises the T_2 statistical guarantees.

Table 4: Summary of Tier 2 statistical guarantees and their assumption dependencies.

Guarantee	Assumptions	Bound
Conservation invariant determinism	A1, A2, A4	Deterministic detection iff one signer honest
Envelope FPR / FNR	A4	Bounded by $\mathcal{D}_X$ tails (FPR) and $\mathcal{D}_X^{\text{adv}}$ band mass (FNR)
Cross vector coherence	A4	Bounded by $\mathcal{D}_f$ band probability $\Phi(k) - \Phi(-k)$
Adaptive adversary detection	A4	Goodness of fit on cohort distribution; degraded for individual lot

Adversary coverage. The T_2 detectors cover $\mathcal{A}_1$ fully, cover $\mathcal{A}_2(k, n)$ for $k < n$ in the cross referenced positions (statistical strength depends on the band width and the distribution’s variance), and do not protect against $\mathcal{A}_2(n, n)$ which is captured by Type 5. The detectors are independent of the network adversary $\mathcal{A}_3$ (they operate on signed payloads already retrieved) and independent of the computational adversary $\mathcal{A}_4$ in the statistical layer (the cryptographic floor protects against $\mathcal{A}_4$ at T_1).

6 Tier 3 Spectroscopic Classification: Identifiability under Reference Base Integrity

The Tier 3 spectroscopic classifier maps an observed spectrum of the finished product to a class label drawn from a reference base. The reference base is constructed at deployment time from a calibration cohort spanning at least three vintages and three production batches per identity class. This section formalises the identifiability guarantee, characterises the reference base contamination model, and bounds the slow rolling adversarial training risk.

6.1 Identifiability Guarantee

Proposition 6 (Tier 3 Identifiability). *Let $\mathcal{R}$ be a reference base of K identity classes with within class spectral variance σ_w^2 and between class minimum separation $d_{\min}$ in the reduced spectroscopic space (after dimensionality reduction or feature selection). For an observed spectrum s from class k with measurement noise σ_n^2 , the probability of correct classification by a maximum likelihood classifier trained on $\mathcal{R}$ is bounded below by*

$$\Pr[\hat{k} = k] \geq 1 - (K - 1) \cdot \exp\left(-\frac{d_{\min}^2}{8(\sigma_w^2 + \sigma_n^2)}\right),$$

in the Gaussian noise regime under standard concentration arguments.

Proof sketch. Standard application of the Bhattacharyya bound (or the union bound over class confusion errors) for Gaussian discrimination. The misclassification probability is dominated by confusion with the nearest neighbouring class; the union bound over $K - 1$ alternative classes yields the stated bound. Tighter bounds are available for specific classifier families (LDA, SVM, deep networks) but the qualitative shape is preserved. $\square$

Consequence. Identifiability scales with the spectral signal to noise ratio and inversely with the number of competing classes. For deployments where $d_{\min} \gg \sigma_w + \sigma_n$, classification is reliable. For matrices where between class separation is poor (e.g. closely related cultivars under similar growing conditions), additional features or external context (Agronomic Vector data) is required to disambiguate.

6.2 Reference Base Contamination Model

Definition 8 (Reference Base Contamination). *A reference base $\mathcal{R}$ is contaminated if a fraction ϵ of the samples in some class k have been mislabeled or fabricated by an adversary. The contamination is silent if it is not detected at calibration time and only manifests through downstream classification errors.*

Proposition 7 (Effect of Contamination on Class Centroid). *If a fraction ϵ of the samples in class k 's reference set are replaced by adversarial samples drawn from class k' , the empirical centroid of class k in the reference base shifts by approximately $\epsilon(\mu_{k'} - \mu_k)$ where $\mu_k, \mu_{k'}$ are the true class centroids. The within class variance also inflates.*

Proof sketch. Linearity of expectation on the contaminated mixture distribution. The centroid is the weighted mean of the honest and adversarial subsets. $\square$

Consequence. A small contamination fraction shifts the class boundary toward the adversary’s preferred class. Subsequent classification of a genuinely class k sample may be misclassified as k' with probability that grows with ϵ .

6.3 Slow Rolling Adversarial Training

The most dangerous attack against the reference base is a slow rolling adversary that submits a small fraction of contaminated samples to each calibration cycle, gradually shifting the reference distribution over multiple recalibration periods.

Proposition 8 (Slow Rolling Drift Bound). *Suppose the adversary submits a fraction ϵ_t of contaminated samples to the calibration cohort at recalibration cycle t , with the cumulative drift in centroid given by*

$$\Delta\mu_t = \sum_{s \leq t} \epsilon_s \cdot (\mu_{k'} - \mu_k) \cdot w_s,$$

where w_s is the weight assigned to cycle s in the empirical centroid.

For uniform weighting $w_s = 1/T$ over T cycles and constant $\epsilon_s = \epsilon$, the cumulative drift is bounded by $\epsilon \cdot (\mu_{k'} - \mu_k)$.

For drift to exceed half the between class minimum separation (the threshold at which classification flips), the adversary must sustain $\epsilon \geq d_{\min}/(2|\mu_{k'} - \mu_k|)$ across the cohort, often requiring cooperation from a substantial fraction of the calibration sources.

Proof sketch. Direct expectation calculation on the cumulative empirical centroid. The classification flip threshold follows from the linear discriminant geometry. $\square$

Mitigations.

1. **Reference base isolation.** Calibration samples are drawn from a separate stream than production samples, with provenance signed by independent attestors. Contamination requires corrupting the calibration provenance, not just submitting fabricated samples.
2. **External validation of recalibration.** Each recalibration is validated against a third party reference set (e.g. a national food composition laboratory’s archived samples). Drift between the deployment’s reference and the external reference is monitored and bounded.
3. **Reputation slashing.** A deployment whose reference base is later proven contaminated by an external audit has its reputation publicly slashed (Section 10), creating an economic deterrent against silent contamination.
4. **Cohort size discipline.** The methodology requires at least 50 samples per identity class across at least three vintages [Fair Food Data, 2026b]. Contamination of more than a small fraction of this cohort requires sustained adversarial effort that is detectable through provenance audit.

6.4 Classifier Robustness under Cultivar Drift

Proposition 9 (Cultivar Drift Robustness). *Let the natural drift in within class distribution between vintages be $\Delta\mu_{\text{vintage}}$ with variance $\sigma_{\text{vintage}}^2$. The classifier trained on vintages 1 through V achieves vintage $V + 1$ accuracy approximately*

$$\Pr[\hat{k} = k]_{V+1} \geq 1 - (K - 1) \cdot \exp\left(-\frac{(d_{\min} - \Delta\mu_{\text{vintage}})^2}{8(\sigma_w^2 + \sigma_n^2 + \sigma_{\text{vintage}}^2)}\right),$$

provided $d_{\min} > \Delta\mu_{\text{vintage}}$ (the classes remain distinguishable across natural drift).

Proof sketch. Same shape as Proposition 6 with the within class variance inflated by the vintage drift variance, and the effective minimum separation reduced by the drift distance. $\square$

Consequence. Recalibration at semestral cadence (Section 6.5) is required to track natural drift. The methodology’s requirement of three vintages in the reference base provides a margin against the drift; deployments operating on a single vintage are vulnerable to misclassification when the drift exceeds the within class variance.

6.5 Summary of Tier 3 Guarantees

Table 5 summarises the T_3 identifiability guarantees.

Table 5: Summary of Tier 3 identifiability guarantees.

Guarantee	Assumptions	Bound
Identifiability under uncontaminated reference	A1, A2, A4 (cohort)	$1 - (K - 1)e^{-d_{\min}^2/8(\sigma_w^2 + \sigma_n^2)}$
Resistance to slow rolling contamination	A4 (calibration cohort), reference base isolation	Drift exceeds class flip threshold only at $\epsilon \geq d_{\min}/2 \mu_{k'} - \mu_k $
Cross vintage robustness	A4 (cohort spans vintages)	$1 - (K - 1)e^{-(d_{\min} - \Delta\mu_{\text{vintage}})^2/8(\sigma_w^2 + \sigma_n^2 + \sigma_{\text{vintage}}^2)}$

Adversary coverage. T_3 identifiability holds against $\mathcal{A}_1$ and $\mathcal{A}_2(k, n)$ for k small relative to the calibration cohort size, conditional on reference base integrity (uncontaminated calibration). $\mathcal{A}_3$ does not affect T_3 at the classification step (the reference base is local). $\mathcal{A}_4$ is bounded by the cryptographic primitives at the signature layer, leaving the statistical classification’s $\mathcal{A}_4$ surface essentially zero.

The cartel adversary $\mathcal{A}_5$ that simultaneously corrupts the calibration provenance and the production samples can in principle defeat T_3 ; the mitigation is that calibration provenance corruption is detectable through external validation against ISO 17025 accredited laboratories at T_4 , restoring a coherent multi tier defence.

7 Tier 4 Laboratory Forensic: Bayesian Posterior Bounds

The Tier 4 forensic verdict is a Bayesian posterior over hypotheses about the lot's true identity, composition, or origin given the laboratory measurements committed in External Attestations. This section formalises the posterior threshold, the dependence on laboratory integrity (A7), and the multi laboratory cross checking that bounds residual risk.

7.1 Posterior Threshold

Definition 9 (Tier 4 Forensic Verdict). *Let H_0 be the null hypothesis (the lot's claimed identity is true) and H_1 the alternative (the claim is false). Given laboratory measurement E committed in an External Attestation, the Tier 4 verdict is *accept* if the posterior $\Pr[H_0 \mid E] \geq \theta_4$ for posterior threshold θ_4 , and *reject* otherwise. The default $\theta_4 = 0.99$ corresponds to forensic confidence at standard juridical levels.*

Proposition 10 (Bayesian Posterior under Independent Likelihood). *Under prior $\Pr[H_0]$ and likelihood ratio $\Lambda(E) = \Pr[E \mid H_0] / \Pr[E \mid H_1]$, the posterior is*

$$\Pr[H_0 \mid E] = \frac{\Pr[H_0] \cdot \Lambda(E)}{\Pr[H_0] \cdot \Lambda(E) + (1 - \Pr[H_0])}.$$

The verdict crosses the threshold θ_4 when

$$\Lambda(E) \geq \frac{\theta_4(1 - \Pr[H_0])}{(1 - \theta_4) \Pr[H_0]}.$$

Proof sketch. Direct application of Bayes' rule [Robert, 2007]. The threshold condition is algebra on the posterior expression. $\square$

Numerical example. For a uniform prior $\Pr[H_0] = 0.5$ and threshold $\theta_4 = 0.99$, the required likelihood ratio is $\Lambda \geq 99$. For laboratory measurements with high specificity (typical for chromatographic or mass spectrometric methods), Λ values in the hundreds to thousands are achievable, comfortably crossing the threshold. For lower specificity methods, multiple independent measurements are combined (Section 7.3).

7.2 Dependence on Laboratory Integrity

Lemma 6 (Conditional on A7). *The Bayesian posterior of Definition 9 is valid only conditional on assumption A7 (ISO 17025 accreditation and operational integrity of the laboratory). If the laboratory is corrupted, the likelihood $\Pr[E \mid H_0]$ and $\Pr[E \mid H_1]$ are no longer the laboratory's measurement uncertainty but the adversary's choice of fabricated value, and the posterior is not meaningful.*

Proof sketch. The likelihoods are determined by the laboratory's measurement process. Corruption replaces the process with an adversarial choice. The Bayesian formula remains formally valid but its inputs are no longer the laboratory's calibrated uncertainty. $\square$

Consequence. A single laboratory’s External Attestation is meaningful only to the extent that A7 holds. The protocol’s T_4 guarantee against single laboratory corruption requires multi laboratory cross checking (Section 7.3).

7.3 Multi Laboratory Cross Checking

Proposition 11 (Independent Multi Laboratory Posterior). *Given L independent External Attestations from L ISO 17025 accredited laboratories, each providing measurement E_i with likelihood ratio Λ_i , and assuming the laboratories’ measurements are conditionally independent given the underlying truth, the combined posterior is*

$$\Pr[H_0 \mid E_1, \dots, E_L] = \frac{\Pr[H_0] \cdot \prod_{i=1}^L \Lambda_i}{\Pr[H_0] \cdot \prod_{i=1}^L \Lambda_i + (1 - \Pr[H_0])}.$$

The aggregate likelihood ratio $\Lambda^* = \prod_i \Lambda_i$ multiplies, providing exponential confidence growth in L .

Proof sketch. Conditional independence of $E_1, \dots, E_L$ given the truth implies factoring of the joint likelihood. The posterior is then the standard Bayesian update with the product of likelihood ratios. $\square$

Robustness to single laboratory corruption. If one of the L laboratories is corrupted and produces fabricated E_i , the corrupted likelihood replaces Λ_i with an adversarial value. The remaining $L - 1$ honest measurements still contribute their honest likelihoods. Where the honest measurements concentrate around one hypothesis and the corrupted one diverges, the posterior is dominated by the honest majority for $L \geq 3$ in typical regimes.

Proposition 12 (Cartel Resistance Threshold for Multi Lab T_4). *For L laboratories with cohort majority honest ($\lceil L/2 \rceil$ honest), and likelihoods Λ_i approximately log normal with mean μ_Λ and standard deviation σ_Λ , the multi laboratory posterior crosses the threshold θ_4 even with up to $\lfloor L/2 \rfloor$ corrupted laboratories, provided the honest signal exceeds an adversarial bias by at least $2\sigma_\Lambda$ in the log likelihood scale.*

Proof sketch. Standard concentration argument on the sum of log likelihoods. The honest contribution dominates the corrupted contribution by a factor exponential in $\sqrt{L_{\text{honest}}}$ in the standard regime. $\square$

Consequence. Multi laboratory T_4 provides a quantifiable resistance threshold against laboratory corruption. The protocol’s recommended deployment uses $L \geq 2$ laboratories for high stakes claims, with $L \geq 3$ where the posterior depends critically on the verdict. A single laboratory deployment carries the residual risk of A7 violation as a single point of failure.

7.4 Sensitivity to Prior

Proposition 13 (Robustness to Prior Choice). *The Tier 4 verdict at threshold $\theta_4 = 0.99$ requires likelihood ratio $\Lambda \geq 99 \cdot (1 - \Pr[H_0]) / \Pr[H_0]$. For priors in the range $\Pr[H_0] \in [0.1, 0.9]$, the required Λ ranges from approximately 11 (favourable prior) to 891 (unfavourable prior). High specificity laboratory measurements with $\Lambda \geq 1000$ achieve the threshold across the entire prior range.*

Consequence. For laboratories with high specificity methods, the posterior threshold is robust to prior choice. The prior matters when the laboratory’s discrimination capability is weak relative to the threshold; in those cases, the deployment shall provide a documented prior derived from independent evidence (e.g. supply chain history, regional baseline rates of fraud).

7.5 Treatment of Ambiguous Verdicts

When the posterior falls between $1 - \theta_4 = 0.01$ and $\theta_4 = 0.99$, the verdict is ambiguous: neither H_0 nor H_1 has decisive evidence. The protocol records the ambiguity and triggers either:

1. Additional laboratory analysis (e.g. a second independent laboratory or a different analytical method).
2. Cross referencing with T_2 and T_3 verdicts to combine evidence across tiers.
3. Suspended issuance pending additional evidence.

The composition theorem of Section 9 formalises how Tier 4 evidence combines with lower tier evidence to resolve ambiguity.

7.6 Summary of Tier 4 Guarantees

Table 6 summarises the T_4 Bayesian guarantees.

Table 6: Summary of Tier 4 Bayesian forensic guarantees.

Guarantee	Assumptions	Bound
Single laboratory posterior	A1, A2, A6, A7	$\Pr[H_0 \mid E] \geq \theta_4$ when $\Lambda \geq \theta_4(1 - \Pr[H_0]) / ((1 - \theta_4) \Pr[H_0])$
Multi laboratory cross checking	A1, A2, A6, A7 (per lab)	Aggregate $\Lambda^* = \prod \Lambda_i$; exponential confidence in L
Cartel resistance ($L \geq 2$ labs)	A7 honesty for $\lceil L/2 \rceil$ labs	Posterior crosses threshold provided honest signal exceeds adversarial bias by $2\sigma_\Lambda$
Robustness to prior	A1, A2, A6, A7	For $\Pr[H_0] \in [0.1, 0.9]$ and high specificity methods, threshold maintained

Adversary coverage. T_4 holds against $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3$ within their respective bounds at the cryptographic floor and against $\mathcal{A}_4$ inheriting the floor. Resistance against $\mathcal{A}_5$ at T_4 depends on multi laboratory deployment with at least one honest laboratory. A single laboratory T_4 is vulnerable to $\mathcal{A}_5$; the recommended deployment uses $L \geq 2$ for high stakes claims.

Composition with lower tiers. Tier 4 evidence is most valuable when it is combined with lower tier signals. A T_4 verdict that confirms T_3 classification or that resolves a T_2 ambiguity carries higher confidence than a T_4 verdict standing alone. The composition theorem of Section 9 formalises the additivity property.

8 DREAM Revocation: Quantitative Guarantee

DREAM (Deletable References in Encrypted Append only Memory) is the construction by Tron and Diedrich [Trón and Diedrich, 2025] that provides retrospective access revocation in content addressed storage. The BDF protocol uses DREAM as the substrate for the right to erasure compliance described in the main paper [Fair Food Data, 2026a]. This section quantifies the revocation guarantee in terms of the network honesty fraction and the path length.

8.1 DREAM Construction Recap

A DREAM reference distributes the chunk identity across multiple network neighbourhoods. Granting access requires the grantee to follow a path through these neighbourhoods to assemble the decryption key. To revoke access, the uploader publishes new chunks to a designated batch that disrupts the path. As long as at least one neighbourhood on the path is honest and applies the disruption, the revocation succeeds: the grantee can no longer assemble the key.

The protocol does not modify DREAM; it embeds DREAM references into the BDF record where right to erasure is exercised. The cryptographic and economic properties of DREAM are inherited from [Trón and Diedrich, 2025].

8.2 Revocation Success Probability

Theorem 2 (DREAM Revocation Probability). *Let h be the honest neighbourhood fraction of the underlying storage network (assumption A3) and n the number of hops in the DREAM path. The probability that revocation succeeds (at least one honest neighbourhood disrupts the path) is bounded below by*

$$\Pr[\text{revoke succeeds}] \geq 1 - (1 - h)^n.$$

For $h = 0.9$ and $n = 6$, this yields $\Pr[\text{revoke succeeds}] \geq 1 - 0.1^6 = 1 - 10^{-6}$, exceeding 99.9999 percent.

Proof sketch. Each neighbourhood on the path is independently honest with probability h under the network model. Revocation fails only if every neighbourhood on the path is dishonest. The probability of this event is $(1 - h)^n$ under independence. The complement gives the revocation success probability.

The independence assumption is empirically justified for Swarm [Trón, 2024], where neighbourhoods are constructed via XOR distance over peer addresses with proof of work admission. Sybil attacks attempting to correlate neighbourhood honesty are bounded by the admission cost. $\square$

Sensitivity to honesty rate. Table 7 tabulates the revocation success probability for representative parameter combinations.

Operational implication. The protocol’s right to erasure compliance is grounded in DREAM with quantitative cryptographic backing rather than in policy based revocation. A user exercising the right

Table 7: DREAM revocation success probability $1 - (1 - h)^n$ for representative honesty rates h and path lengths n .

Honesty h	$n = 2$	$n = 4$	$n = 6$	$n = 8$	$n = 10$
0.50	0.7500	0.9375	0.9844	0.9961	0.9990
0.70	0.9100	0.9919	0.9993	0.9999	$> 1 - 10^{-5}$
0.80	0.9600	0.9984	0.9999	$> 1 - 10^{-5}$	$> 1 - 10^{-6}$
0.90	0.9900	0.9999	$> 1 - 10^{-5}$	$> 1 - 10^{-7}$	$> 1 - 10^{-9}$
0.95	0.9975	$> 1 - 10^{-5}$	$> 1 - 10^{-7}$	$> 1 - 10^{-9}$	$> 1 - 10^{-12}$

to be forgotten under GDPR Article 17 [Fair Food Data, 2026a] obtains revocation with probability bounded by the table above. Operational deployments target $h \geq 0.8$ and $n \geq 6$, yielding revocation confidence at the $1 - 10^{-5}$ level.

8.3 Limits and Failure Modes

Network majority adversary. If the adversary captures a majority of the network ($h < 0.5$), the revocation guarantee degrades sharply: at $h = 0.4$ and $n = 6$, success probability is $1 - 0.6^6 \approx 0.953$, still acceptable but with notable failure rate. At $h < 0.3$, the guarantee weakens to the point where path length must increase materially. The protocol’s defence in this regime is the underlying network’s economic and Sybil resistance [Trón, 2024]; a network with persistent $h < 0.5$ is operationally compromised regardless of DREAM.

Long retention contradicts revocation. DREAM revocation operates by disrupting paths to chunks that have already been distributed. Chunks that have been retained by parties outside the DREAM scope (e.g. a grantee who downloaded the data and stored it locally) are outside the revocation guarantee. The protocol’s right to erasure addresses the publicly retrievable layer; it does not retrieve copies that have been exfiltrated. This is the standard limit of cryptographic deletion mechanisms and is consistent with GDPR’s distinction between the operator’s obligation and the data’s diffusion.

Incremental revocation cost. Each revocation event requires the uploader to publish a small batch of disruption chunks and pay the corresponding postage. The cost is bounded and predictable; for typical deployments it is a small fraction of the original storage cost. The protocol does not currently optimise for batch revocation of many chunks at once, though such optimisations are documented in [Trón and Diedrich, 2025].

8.4 Auditable Erasure Events

The revocation event is itself a signed object on the BDT record. A verifier inspecting the record observes both the original chunk anchor and the subsequent revocation event. Together these provide auditable evidence that erasure was exercised at a specific time, with known cryptographic strength, by a specific party. The verifier does not retrieve the plaintext (which is no longer accessible after revocation) but can confirm the structural integrity of the audit trail and the existence of the erasure act.

This is a stronger property than policy based revocation, where deletion is the absence of evidence. Cryptographic revocation provides positive evidence of the erasure event itself, supporting compliance audits and stakeholder claims about right to erasure exercise.

8.5 Composition with the Cryptographic Floor

Revocation does not weaken the structural T_1 guarantees on the public chain. The ownership chain, distribution events, and burn semantics remain valid; only the encrypted payload becomes inaccessible. The verifier reading the chain after revocation observes:

- The structural anchor (the canonical hash H_O).
- The ownership chain transitions before and after revocation.
- The signed revocation event itself.
- The absence of accessible plaintext for the affected payload.

A verifier evaluating a claim that depends on the revoked plaintext cannot reproduce the original T_2 detector verdict; the methodology accepts that the original verdict was recorded before the erasure (Section 6.7 of the methodology paper [Fair Food Data, 2026b]) and accepts the revocation as the auditable trail of the right to erasure exercise.

9 Composition Theorem

The per tier reductions of Sections 4 through 7 bound the protocol’s guarantees against specific adversary classes at specific tiers. This section establishes that those per tier guarantees compose: when a deployment runs all four tiers, the guarantees combine without weakening, and the residual cartel risk is precisely the intersection of the tier specific failure regimes.

9.1 Statement of the Composition Theorem

Theorem 3 (Composition of BDF Tier Guarantees). *Let G_i denote the security guarantee provided by tier T_i against an adversary $\mathcal{A}$ under a specified set of trust assumptions. Then the composed guarantee of running tiers T_1, T_2, T_3, T_4 jointly is*

$$G_{\text{composed}}(\mathcal{A}) = G_1(\mathcal{A}) \vee G_2(\mathcal{A}) \vee G_3(\mathcal{A}) \vee G_4(\mathcal{A}),$$

in the sense that the adversary defeats the composed system only by simultaneously defeating every tier whose guarantee applies under the deployment’s trust assumptions. The composition is sound: no tier weakens any other, and the bounds aggregate by union or by intersection as appropriate.

Specifically, against an adversary attempting to fabricate a structurally valid record with false content, the probability of success is bounded by

$$\Pr[\mathcal{A} \text{ succeeds}] \leq \min \left(\text{Adv}_{\mathcal{A}}^{T_1}, \text{FNR}^{T_2}, 1 - \text{Identifiability}^{T_3}, 1 - \theta_4^{T_4} \right),$$

where each term corresponds to the failure mode of the respective tier.

Proof sketch. We argue that detection by any single tier is sufficient to defeat the adversary, hence the composed system requires the adversary to defeat every applicable tier simultaneously.

T_1 structural defeat. The adversary must produce signed payloads that are canonically valid, signature verifying, and consistent with the ownership chain (Lemmas 1, 2, 3). Failure to do so triggers T_1 rejection. The adversary’s success probability at T_1 alone is bounded by $\text{Adv}_{\mathcal{A}_4}^{\text{coll}} + n \cdot \text{Adv}_{\mathcal{A}_4}^{\text{euf-cma}}$ which is negligible under A1.

T_2 statistical defeat. Conditional on T_1 acceptance, the adversary’s payloads are subjected to conservation, envelope, and cross vector tests. By Propositions 3 and 4, the adversary’s success probability against the T_2 test ensemble is bounded by the false negative rate of the most sensitive applicable detector and by the cross vector independence argument.

T_3 identifiability defeat. Conditional on T_2 pass, the spectroscopic classifier evaluates the finished product against the reference base. By Proposition 6, the misclassification probability is bounded by the Bhattacharyya formula.

T_4 Bayesian defeat. Conditional on T_3 pass, the External Attestation provides Bayesian evidence on the lot’s claimed identity. By Proposition 11, multi laboratory cross checking yields exponential confidence.

The adversary defeats the composed system only by defeating all applicable tiers. The composition bound follows by the chain of conditioning: the adversary’s joint success probability is the product (under independence) or the minimum (under correlation) of the per tier failure bounds. $\square$

Soundness. The composition does not weaken any per tier guarantee. Each tier operates over the same signed payloads but applies independent tests. A failure at tier T_i does not cascade to weaken the test at tier T_j for $j \neq i$; the tiers are evaluated independently and the composed verdict is the conjunction of all tier verdicts.

9.2 Cartel Residual Risk Reduction

Corollary 1 (Cartel Risk under Composition). *The coordinated cartel adversary $\mathcal{A}_5$ that defeats all attestors and the laboratory simultaneously is the only adversary class that defeats the composed system at base configuration without invoking the Hardening Pathways. The cartel’s success requires:*

1. *All attestors signing $\vec{A}, \vec{L}, \vec{P}, \vec{R}$ in coordination (T_2 and T_3 defeated).*
2. *All ISO 17025 laboratories issuing External Attestations corrupted (T_4 defeated).*
3. *The structural payload still verifies under T_1 (this is implied by valid signatures from corrupted attestors).*

The per tier failure bounds collapse under the cartel because each failure regime is satisfied simultaneously.

Mitigation through Hardening Pathways. Section 10 demonstrates that:

- **Proplet attestation (substitute for A4).** The trust principal at signing shifts from operator to hardware. The cartel must compromise hardware roots of trust at every signing device, raising

the attack cost by orders of magnitude.

- **Cryptoeconomic mechanisms (substitute partial for A4).** Stakes calibrated above the marginal benefit of fraud convert coordinated attack into expected economic loss. The cartel’s expected utility is negative in expectation.
- **Multi laboratory T_4 (substitute for A7).** Multiple independent ISO 17025 laboratories provide A7 redundancy. The cartel must corrupt all of them simultaneously, with each additional laboratory raising the coordination cost.
- **ZK proofs (substitute for full payload disclosure).** Selective verification reveals less to the verifier without weakening the cryptographic guarantee, reducing the cartel’s information asymmetry.

The composition theorem applies to the hardened protocol with the substitutions: where the base configuration depends on A4 (attestor honesty), the hardened configuration depends on hardware attestation integrity plus stake economic calibration. The substitution is structural, not statistical.

9.3 Adversary Coverage Matrix Update

Table 8 updates the adversary coverage matrix from Section 2 with the composition argument applied. Each cell now reflects the composed guarantee under the listed deployment configuration.

Table 8: Adversary coverage under composition. Configurations: Base = T_1 to T_4 without Hardening Pathways; HW = base plus Proplet hardware attestation; HW+CE = HW plus cryptoeconomic mechanism; HW+CE+ML = HW+CE plus multi laboratory T_4 .

Adversary	Base	HW	HW+CE	HW+CE+ML
$\mathcal{A}_1$ Lone fraudster	P	P	P	P
$\mathcal{A}_2$ Coalition $k < n$	P	P	P	P
$\mathcal{A}_2$ Coalition $k = n$	p	P	P	P
$\mathcal{A}_3$ Network $\rho < 1 - h$	P	P	P	P
$\mathcal{A}_3$ Network $\rho > 1 - h$	p	p	p	p
$\mathcal{A}_4$ Computational	P	P	P	P
$\mathcal{A}_5$ Coordinated cartel	N	p	p	P

Reading the table. The base configuration covers all standard adversaries except the coordinated cartel. Adding Proplet hardware attestation (HW) covers the coalition with full attestor capture and provides partial cartel coverage by raising the hardware compromise cost. Adding cryptoeconomic mechanisms (CE) deepens cartel resistance through stake calibration. Adding multi laboratory T_4 (ML) closes the cartel’s remaining laboratory route to a structural majority requirement, providing full coverage in the operational regime.

The protocol thus offers a tunable security spectrum: deployments with low value at stake operate at Base, deployments with high stakes (carbon credit issuance for high integrity tier, high value provenance claims) operate at HW+CE+ML.

9.4 Composition Bound Tightness

The composition bound presented in Theorem 3 is conservative. In practice, an adversary that defeats T_1 structurally is unlikely to also fabricate values within the T_2 envelope and within the T_3 classifier, because these failure modes are correlated with structural fabrication. Tighter bounds are achievable in specific regimes (e.g. when the adversary is constrained to small payload variations). The conservative bound is sufficient for registry assessment and operational planning; tighter analysis is left to deployment specific case studies.

10 Hardening Pathway Reductions

The base protocol relies on assumption A4 (honesty among attestors) for its T_2 statistical guarantees. The Hardening Pathways introduced in the main paper [Fair Food Data, 2026a] provide cryptographic, economic, and information theoretic substitutes for A4 that progressively close the cartel residual risk. This section formalises the security reductions for each pathway.

10.1 Proplet Hardware Attestation

Following Szabo’s Proplet construction [Szabo, 2001], the protocol can be hardened by replacing the human operator at signing with a tamper resistant device that signs autonomously upon witnessing the physical event. The trust principal shifts from operator to hardware.

Definition 10 (Proplet Compliant Signing). *A Proplet compliant signing device satisfies:*

1. *Hardware root of trust based on a Trusted Platform Module 2.0 [Trusted Computing Group, 2019] or equivalent secure element.*
2. *Remote attestation [Coker et al., 2011] that proves the device’s identity and current firmware state to a verifier.*
3. *GPS based self geofencing that constrains signing to a declared geographic region.*
4. *On chain ownership registry that binds the device’s public key to its physical owner of record.*
5. *Signing operations performed inside the secure boundary; private keys never leave the secure element.*

Theorem 4 (A4 Substitution under Proplet). *Under the Proplet compliant signing of Definition 10, the protocol’s T_2 statistical guarantees that depend on A4 (honesty among attestors) reduce to the following hardware assumptions:*

- **H1.** *Hardware tamper resistance: extracting the private key from the secure element requires physical destruction of the device with cost exceeding C_{tamper} .*
- **H2.** *Remote attestation soundness: a verifier accepting an attestation only if the device’s firmware hash matches a known good value cannot be deceived by a polynomial time adversary except with negligible probability.*
- **H3.** *Sensor calibration integrity: the device’s onboard sensors produce measurements within their calibrated uncertainty, with calibration verified at manufacture and recertified periodically.*

The trust assumption A4 (organisational independence of human attestors) is replaced by H1 plus H2 plus H3 (hardware level independence of physical devices). Coalition attacks now require physical compromise of every Proplet device involved in the supply chain, raising the attack cost to $n \cdot C_{\text{tamper}}$ for n devices.

Proof sketch. The reduction is direct: the signing principal in the cryptographic layer is the device's secure element, not the human operator. An adversary attempting to fabricate a payload must either: (a) Compromise the secure element to extract the private key (cost C_{tamper} per device, bounded below by the secure element's physical security). (b) Compromise the firmware to sign fabricated payloads while passing remote attestation (defeated by H2). (c) Compromise the sensors to produce fabricated readings (defeated by H3 plus the physical sensor's calibration).

The intersection of these compromise paths defines the hardware level adversary; the social adversary that suborns a human operator is no longer the operative threat. $\square$

Operational implication. A deployment using Proplets shifts the trust boundary from people to devices. Coalition attacks remain possible against a fully compromised operator (who could replace devices), but each device replacement is itself a signed event detectable from the on chain ownership registry. The combinatorial cost of executing a coalition attack against n Proplet devices simultaneously, with synchronised firmware compromise, exceeds the value at stake in any deployment the protocol is designed to support.

10.2 Cryptoeconomic Stake Calibration

Definition 11 (Staked Attestor). A staked attestor is an attestor that posts a deposit S before signing payloads. The deposit is locked for a challenge window W . A successful fraud proof against any signed payload during W slashes the deposit and pays a bounty B to the challenger. The staked attestor's expected payoff under fraud is

$$U_{\text{fraud}} = G - p_{\text{detect}} \cdot S - (1 - p_{\text{detect}}) \cdot 0,$$

where G is the gross gain from fraud and p_{detect} is the probability of detection within W .

Theorem 5 (Stake Calibration Theorem). Under a deployment where every attestor is staked with S and the BDF protocol's T_2 , T_3 , and T_4 detection mechanisms produce fraud proofs with probability p_{detect} , the rational attestor refrains from fraud whenever

$$S \geq \frac{G}{p_{\text{detect}}}.$$

For deployments where the detection probability exceeds 0.5 (typical when multiple tiers operate in composition under Theorem 3), a stake exceeding twice the expected gross fraud benefit is sufficient to deter rational adversaries.

Proof sketch. Direct expected utility argument. The attestor compares the expected fraud payoff $G - p_{\text{detect}}S$ against the honest payoff (zero gain, zero risk) and selects the higher utility. Setting the inequality $G - p_{\text{detect}}S \leq 0$ yields the stated condition. This is the standard cryptoeconomic argument from the Casper line of work [Buterin et al., 2020]. $\square$

Coordinated cartel under stake calibration. For the coordinated cartel adversary $\mathcal{A}_5$, the calibrated stake bound applies multiplicatively: the cartel’s joint payoff is bounded by the sum of individual fraud gains, while the joint slashing risk is the sum of individual stakes. The cartel’s expected utility under correlated stakes is negative whenever the per attester stake calibration of Theorem 5 holds individually, providing direct cartel resistance through economic incentives.

Reputation onchain. A complementary mechanism stores attester reputation as a signed object on the public chain. Honest attestation accumulates reputation; a proven fraud destroys it. Reputation provides a long horizon incentive that complements the short horizon stake: an attester with high accumulated reputation has strong incentive to remain honest even when individual stakes are small.

Insurance pools. A further layer collects premiums from attestors and pays claims when fraud is proven against any participant. The pool internalises the risk and provides aggregate coverage that exceeds the individual stake. This is documented in [Fair Food Data, 2026a] as a separate Hardening Pathway component.

10.3 Zero Knowledge Selective Verification

Zero knowledge proofs [Goldwasser et al., 1989] allow a verifier to confirm a property of a payload without seeing the payload. For BDF, this enables verification of specific predicates over signed vectors while preserving commercial confidentiality.

Definition 12 (ZK Predicate Verification). *For a signed payload $\vec{V}$ and a predicate $P(\vec{V}) \in \{0, 1\}$ (e.g. “the GPS coordinate falls within polygon Π ”), the protocol generates a zero knowledge proof π such that the verifier checks π in time bounded by the circuit complexity, accepts $P(\vec{V}) = 1$ with completeness, and rejects $P(\vec{V}) = 0$ with soundness probability $1 - \text{negl}(\lambda)$.*

Proposition 14 (ZK Soundness Reduction). *Under the security of the underlying ZK system (Groth16 [Groth, 2016], Plonk, or equivalent), a polynomial time adversary who does not know a witness w such that $P(w) = 1$ cannot produce a verifying proof π except with probability $\text{Adv}_{\mathcal{A}_4}^{\text{zk-sound}}(\lambda) = \text{negl}(\lambda)$.*

Proof sketch. Direct reduction to the soundness of the underlying ZK system. The protocol does not modify the ZK construction; it embeds the BDF predicate into the proof system’s circuit and inherits the soundness. $\square$

Operational implications.

- **Confidentiality.** The verifier learns only $P(\vec{V}) = 1$, not the payload itself. The protocol’s selective disclosure mechanism (Section 4.7) is augmented: instead of revealing the encrypted payload to the verifier under NDA, the operator generates a ZK proof of the relevant predicate and reveals only the proof.
- **Performance.** ZK proof generation is currently practical for simple circuits (predicates expressed in a few thousand constraints) and impractical for complex multivariate forensic computations. The protocol’s adoption of ZK is incremental, starting with simple predicates (geofencing, threshold tests) and extending as proof generation costs decline.

- **Cartel adversary reduction.** ZK does not eliminate cartel risk but reduces the cartel’s information advantage: the cartel can still fabricate the underlying payload, but cannot compel the verifier to evaluate a property the cartel did not pre commit to via the ZK proof. The verifier’s selection of the predicate is itself a defence against the cartel’s narrative.

10.4 Composition of Hardening Pathways

The three Hardening Pathways are independent and reinforce each other:

- Proplets without economic incentives are tamper resistant but the device owner can still authorise fraud through legitimate device control.
- Cryptoeconomic mechanisms without hardware are bounded by the social adversary (stakes can be socialised across cartel members and the deterrent dilutes).
- ZK proofs without identifiable signers are powerless: the verifier needs to know whose signature is being proven.

Theorem 6 (Hardening Pathway Composition). *The composition of Proplet attestation, cryptoeconomic stake calibration, and ZK selective verification reduces the cartel residual risk to a quantifiable economic parameter:*

$$\text{Cartel Risk} \leq \frac{G}{p_{\text{detect}}^{\text{composed}}} + n \cdot C_{\text{tamper}},$$

where $p_{\text{detect}}^{\text{composed}}$ is the composed detection probability across all four tiers under Theorem 3, and $n \cdot C_{\text{tamper}}$ is the hardware compromise cost across all Proplet devices in the cartel.

Proof sketch. Combining Theorems 4, 5, and Proposition 14: the cartel’s expected fraud cost is the sum of its hardware compromise cost (Proplet) and its expected slashing exposure (stake), against an expected gain G . The composition argument from Theorem 3 establishes the joint detection probability that enters the slashing condition. $\square$

Tunable security. The deployment selects the Hardening Pathway depth proportional to the value at stake. Low stakes deployments operate at base configuration; high stakes deployments at full hardening. The protocol is consistent across the spectrum: the same canonical objects, the same tiers, the same detector logic. Only the substitutes for assumption A4 vary.

11 Quantitative Limits

The protocol’s guarantees are bounded by the assumptions on which they rest. This section enumerates the quantitative limits and the residual gaps, distinguishing between limits that the protocol acknowledges as inherent and limits that are addressable through future work or deployment specific choices.

11.1 Cryptographic Security Level

The protocol inherits the security level of its primitives. At the time of writing, the reference instantiation provides 128 bit security against a polynomial time classical adversary. Specific limits:

- **Hash collision.** SHA-256 within BMT provides 2^{-128} collision security against a generic birthday attack. Cryptanalytic improvements that lower this would degrade T_1 object identity uniqueness proportionally.
- **Signature forgery.** ECDSA on secp256k1 provides 128 bit security against polynomial time forgery. A break of the elliptic curve discrete logarithm assumption would degrade ownership chain consistency to a non guarantee.
- **Authenticated encryption.** AES-256-GCM provides 128 bit security against IND-CCA attacks. A break would degrade payload confidentiality but not structural integrity.

Post quantum exposure. A practical quantum computer running Shor’s algorithm at scale would break ECDSA. The protocol’s migration path is to upgrade the signature scheme through the canonical hash type field of beeson, allowing historical records to be re-anchored under a post quantum signature scheme. The migration is a structural change with operational cost; it is not a transparent upgrade. Proactive deployments may pre commit to a hybrid signature scheme that signs every payload with both ECDSA and a post quantum candidate (e.g. Falcon, Dilithium), accepting the larger signature footprint as the cost of forward security.

Quantum exposure on encryption. AES-256 retains 128 bit security against Grover’s algorithm, providing post quantum confidentiality at the symmetric layer. The asymmetric encryption used for selective disclosure key sharing inherits the post quantum exposure of the signature scheme; the same migration path applies.

11.2 Side Channels and Implementation Attacks

The protocol’s security analysis assumes the reference implementation faithfully realises the canonical primitives. Side channel attacks that extract private keys from implementation timing, power consumption, electromagnetic emanation, or fault injection are out of scope. These attacks are properties of the deployment hardware and the implementation rather than of the protocol.

Mitigations available to deployments.

- **Constant time signing.** Reference implementations use constant time field arithmetic to mitigate timing side channels.
- **Hardware secure elements.** Proplets that use TPM 2.0 or equivalent secure elements benefit from hardware level side channel resistance certified at common criteria EAL4+ levels.
- **Operational hygiene.** Air gapping signing devices, physical access control, and tamper evident enclosures reduce side channel exposure.

The protocol does not mandate any specific side channel mitigation; the choice is a deployment decision driven by the value at stake and the threat environment.

11.3 Pre Genesis Attacks

The protocol provides no guarantees against attacks that occur before any signed evidence exists. If the adversary corrupts the physical reality (substitutes a different cultivar, contaminates the harvest, alters mass through tampering with the basket) before the genesis attestation is signed, the BDT comes into existence with the false content already embedded. The signed payloads will be internally consistent and the cryptographic floor will hold; the lower tier guarantees do not detect the substitution because the substitution predates the evidence.

Lemma 7 (Pre Genesis Gap). *For any attestation vector $\vec{V}$ signed at genesis under a base configuration without Proplet attestation, an adversary who controls the physical reality before the signing event can produce a $\vec{V}$ that is internally consistent (and accepted at all tiers) but factually false. The protocol's T_1 through T_4 guarantees do not constrain pre genesis content.*

Proof sketch. The signing process commits to whatever the operator (or device) records. If the recorded reality is itself fabricated upstream, the cryptographic commitment is to the fabrication. The protocol detects post genesis tampering but not pre genesis substitution. $\square$

Mitigation through Proplet attestation. A Proplet device that signs autonomously upon witnessing the physical event (Section 10.1) closes the pre genesis gap to the device's window of physical access. The adversary must compromise the device or substitute the physical input before the device's sensors observe it. The window of opportunity narrows from "any moment before signing" to "the moment of physical witness", which can be reduced through device tamper evidence and continuous monitoring.

Mitigation through cross referencing. Where the genesis attestation $\vec{A}$ is at the farm but the lot is also weighed at facility intake (yielding $\vec{L}$), a substitution at genesis that yields a different mass than declared at intake is detected by the conservation invariant. The pre genesis substitution is detected ex post when the substituted lot encounters honest downstream verification. This does not provide pre genesis prevention, but it provides post hoc detection that supports dispute resolution.

Mitigation through statistical baseline. Where the project declares an expected baseline range for the recovered matter (e.g. blueberry from a specific cultivar in a specific region), spectroscopic classification at T_3 detects substitutions that move the matter outside the baseline distribution. Pre genesis substitutions of one cultivar for another are detected at T_3 even though the genesis attestation itself is internally consistent.

The pre genesis gap is acknowledged as the principal limit of the protocol's base configuration. It is structurally addressable by Proplet attestation and operationally mitigated by cross referencing at higher tiers. A deployment for which pre genesis attacks are a material risk shall implement Proplet attestation as part of the operational profile.

11.4 Network Majority Adversary

The protocol's T_1 structural guarantees are network independent: signed objects already retrieved remain valid regardless of the network state. New evidence and DREAM revocation depend on the network. Specifically:

- **Below honesty threshold.** If $h < 0.5$ (network minority adversary controls majority neighbourhoods), the data availability layer fails: new chunks may not propagate, retrievals may be censored, DREAM revocation may not succeed. The protocol's response is structural: the underlying network's economic and Sybil resistance guarantees become the operative defence.
- **Sybil flooding.** An adversary that floods the network with Sybil identities to inflate apparent honesty is bounded by the proof of work admission cost [Trón, 2024]. The cost rises with the number of Sybils; the adversary must balance against the value of the attack.
- **Eclipse attacks.** An adversary that surrounds a target node with adversarial peers to deny it access to honest neighbourhoods is mitigated by Kademlia's structured routing [Maymounkov and Mazières, 2002] and by node level peer diversity heuristics.

Operational implication. A deployment shall monitor the underlying network's honesty estimate and shall not operate at $h < 0.5$. Where network conditions degrade temporarily, the protocol's higher tier guarantees (T_3 and T_4) remain valid and provide a fallback verification path independent of the network's data availability.

11.5 Economic and Legal Constraints

The protocol's cryptographic guarantees do not extend to economic or legal layers. Specifically:

- **Smart contract bugs.** If the protocol is implemented on a chain whose smart contracts have bugs, the protocol's structural integrity may be compromised at the smart contract layer even when the cryptographic layer is sound. Mitigation: smart contract auditing and formal verification are deployment level concerns.
- **Legal recognition.** The protocol's signed evidence has cryptographic weight but its juridical weight depends on legal recognition of digital signatures and ISO 17025 accreditation in the relevant jurisdictions. Legal recognition is uneven across jurisdictions and evolves over time. Deployments shall consult local counsel.
- **Contractual override.** Parties may agree contractually to override the protocol's verdicts (e.g. a buyer may accept a lot despite a T_2 warn). The protocol records the verdict; the contractual treatment is a layer above.
- **Regulatory change.** Future regulatory changes (e.g. a jurisdictional ban on cryptographic deletion under data protection law) may change the operational scope of the protocol. The protocol's cryptographic guarantees are technology level; their operational use is regulated.

11.6 Summary of Limits

Table 9 summarises the limits enumerated in this section.

Table 9: Summary of protocol limits and mitigations.

Limit	Source	Mitigation
128 bit security ceiling	Primitive choice	Hybrid PQ signatures pre-emptively; migration via beeson type field
Side channels	Implementation hardware	TPM 2.0 secure elements; constant time signing; physical hygiene
Pre genesis attacks	Physical reality before signing	Proplet attestation; cross referencing at T_2 to T_3 ex post
Network majority adversary	$h < 0.5$ in storage layer	T_3 and T_4 fallback; underlying network Sybil resistance
Smart contract bugs	Chain implementation	Auditing; formal verification at deployment
Legal recognition	Jurisdictional variation	Local counsel; ISO 17025 plus ILAC MRA for cross border weight

Honest disclosure. The protocol is positioned as opportunistic and partial collusion resistance, not absolute fraud prevention. The honest disclosure of these limits supports registry assessment and stakeholder trust: an assessor who reads this section knows what the protocol does and does not protect against, and can plan accordingly.

12 Conclusion

This paper formalises the security analysis of the Bio-Digital Footprint protocol. It states the threat model with five adversary types of increasing power, articulates eight trust assumptions in formal language, and provides per tier security reductions with proof sketches that bound the adversary’s success probability against quantifiable parameters of the deployment.

The cryptographic floor at T_1 inherits standard 128 bit security from its primitives: object identity uniqueness from collision resistance, signature unforgeability from EUF-CMA, ownership chain consistency and burn non reusability from the union of these, and distribution conservation from the structural verifier check. The statistical layer at T_2 derives detection power from independence among attestors, with conservation invariants providing deterministic detection under any single honest signer and envelope tests bounded by the natural variability of the underlying distributions. The identifiability layer at T_3 provides classification accuracy that scales with spectral signal to noise ratio and inversely with class count, robust to slow rolling adversarial training under reference base isolation. The Bayesian forensic layer at T_4 provides posterior bounds at the standard $\theta_4 = 0.99$ threshold, with multi laboratory cross checking yielding exponential confidence growth under conditional independence.

The DREAM revocation guarantee binds quantitatively to the network honesty rate and path length: at $h = 0.9$ and $n = 6$, revocation succeeds with probability exceeding $1 - 10^{-6}$. This is cryptographic deletion with measurable confidence, structurally distinct from policy based deletion in centralised systems.

The composition theorem establishes that per tier guarantees combine without weakening: the adversary defeats the composed system only by simultaneously defeating every applicable tier under the deployment’s trust assumptions. The coordinated cartel adversary is the principal residual risk against the base configuration. The Hardening Pathways close this risk progressively: Proplet hardware attestation substitutes hardware level independence for organisational independence (closing assumption A4), cryptoeconomic stake calibration converts coordinated fraud into expected economic loss, and zero knowledge proofs reduce the cartel’s information asymmetry. The combined hardened configuration reduces cartel risk to a quantifiable economic parameter.

The protocol’s limits are honestly enumerated: the 128 bit cryptographic ceiling (with a documented post quantum migration path), side channel exposure (deferred to deployment hardware), pre genesis attacks (closed by Proplet attestation), and network majority adversaries (defended by Sybil resistance of the underlying storage). The protocol does not promise absolute fraud prevention; it promises opportunistic and partial collusion detection, with explicit characterisation of the regimes in which each tier provides each guarantee.

The intended consumers of this analysis are registry assessors evaluating BDF based methodologies for ICVCM Core Carbon Principles labelling [Integrity Council for the Voluntary Carbon Market, 2024], peer reviewers evaluating the protocol for publication, and protocol implementers building deployments. For each audience, the paper provides the formal grounding required to assess, validate, or implement the protocol with quantitative confidence.

What this analysis does not replace. The proof sketches in this paper are sufficient for the structural reductions but are not full formal proofs in the sense of an interactive theorem prover or a mechanised verification suite. The protocol’s adoption at high integrity scale will benefit from such mechanisation; the present paper provides the mathematical structure on which that work can build.

Operational readiness. The protocol’s base configuration (without Hardening Pathways) is operationally ready for deployments at low to medium stakes where the trust assumptions A1 through A8 hold. The hardened configurations close the cartel residual risk for deployments at high stakes (carbon credit issuance for the high integrity tier, premium provenance claims, regulatory compliance for high value assets). The protocol scales smoothly across the spectrum without changes to the canonical objects, the tier semantics, or the verifier logic.

Outlook. The successor work in the BDF programme includes: full mechanised proofs of the lemmas presented here in a proof assistant such as Coq or Lean, empirical validation of the T_2 detection thresholds against larger pilot cohorts, extension of the spectroscopic classifier coverage to additional cultivars and matrices, and operational integration of Proplet hardware in the next pilot generation. Each of these strengthens specific components of the analysis without changing the structural form.

The protocol provides a verifiable foundation on which deployments can build. The security analysis presented in this paper is the formal grounding of that foundation.

A Proof Sketches

This appendix expands the proof sketches in the body to include auxiliary lemmas, the cryptographic reduction structure, and the concentration arguments invoked at T_2 and T_3 . The expansion is sufficient for a cryptographically literate reader to reconstruct the full argument; fully formal proofs are deferred to a successor paper or to peer review feedback.

A.1 Auxiliary Lemmas

Lemma 8 (Canonicalisation Injectivity). *Under assumption A2, the canonical serialisation function Serialise is injective on the type universe of well formed beeson objects: for any two distinct typed objects $O_1 \neq O_2$, $\text{Serialise}(O_1) \neq \text{Serialise}(O_2)$.*

Proof sketch. By construction of beeson [Tron et al., 2026]: the serialisation embeds the schema and its instance with a deterministic encoding that uniquely determines the object from the byte sequence. The encoding fixes field ordering, normalises numeric encodings, and rejects ambiguous representations. Two distinct objects differ in at least one field; the encoding of that field differs in the byte sequence; hence the serialisations differ.

The conformance suite that all implementations of beeson must pass tests injectivity on a representative universe of objects. Implementations that fail the conformance suite are non conformant and outside the scope of A2. $\square$

Lemma 9 (BMT Hash Collision Bound). *Under assumption A1 and the structural reduction of $\mathcal{H}_{\text{BMT}}$ to a Merkle tree over SHA-256 block compressions [Merkle, 1987], the collision probability for the BMT hash is bounded by the underlying SHA-256 collision probability, which is $\text{Adv}_{\mathcal{A}_4}^{\text{coll}}(\lambda) = \text{negl}(\lambda)$ for $\lambda = 128$.*

Proof sketch. A BMT hash collision implies a collision at some node of the tree, which implies a collision of the underlying compression function applied to some input pair. By the union bound over tree nodes (polynomial in the tree depth), the total collision probability is bounded by $\text{poly}(\log N) \cdot \text{Adv}_{\mathcal{A}_4}^{\text{coll}}$ where N is the tree’s leaf count. For practical tree sizes this remains negligible. $\square$

A.2 Full Sketch: Object Identity Uniqueness (Lemma 1)

The lemma claims that distinct objects yield distinct hashes except with negligible probability. The argument composes Lemma 8 (canonicalisation is injective) and Lemma 9 (BMT hash is collision resistant).

1. Suppose $O_1 \neq O_2$ are well formed typed objects.
2. By Lemma 8, $\text{Serialise}(O_1) \neq \text{Serialise}(O_2)$.
3. By Lemma 9, $\Pr[\mathcal{H}_{\text{BMT}}(\text{Serialise}(O_1)) = \mathcal{H}_{\text{BMT}}(\text{Serialise}(O_2))] \leq \text{Adv}_{\mathcal{A}_4}^{\text{coll}}(\lambda)$.
4. Hence $\Pr[H_{O_1} = H_{O_2}] \leq \text{Adv}_{\mathcal{A}_4}^{\text{coll}}(\lambda) = \text{negl}(\lambda)$.

The proof is structural; the cryptographic strength comes entirely from the underlying primitive.

A.3 Full Sketch: Signature Unforgeability (Lemma 2)

The lemma asserts that an adversary without a party's private key cannot produce a valid signature on a previously unsigned message except with negligible probability. The reduction is to the standard EUF-CMA experiment.

1. Suppose $\mathcal{A}_4$ is an adversary that wins the BDT signature unforgeability experiment with non negligible probability ϵ .
2. Construct a reduction $\mathcal{B}$ that uses $\mathcal{A}_4$ as a subroutine to win the standard EUF-CMA experiment for the signature scheme: $\mathcal{B}$ receives the public key pk from the EUF-CMA challenger, presents pk to $\mathcal{A}_4$ as the BDT party's public key, forwards $\mathcal{A}_4$'s queries to its own oracle, and outputs $\mathcal{A}_4$'s forgery as its own.
3. $\mathcal{B}$ wins EUF-CMA with the same probability ϵ .
4. By A1, the EUF-CMA advantage is bounded: $\text{Adv}_{\mathcal{A}_4}^{\text{euf-cma}}(\lambda) = \text{negl}(\lambda)$.
5. Hence $\epsilon = \text{negl}(\lambda)$.

The reduction is tight; no security loss occurs.

A.4 Full Sketch: Ownership Chain Consistency (Lemma 3)

The lemma has two parts: chain consistency under unforgeability, and public detectability of double signing.

Part 1: Chain Consistency. The chain $A_0, A_1, \dots, A_n$ is consistent if every $\sigma_i = \text{Sign}_{A_i}(H_O \parallel A_{i+1})$ verifies under A_i 's public key.

By Lemma 2 applied to each σ_i : an adversary without sk_{A_i} cannot produce a valid σ_i except with probability $\text{Adv}_{\mathcal{A}_4}^{\text{euf-cma}}$. Union bound across n signatures yields the chain consistency probability bound.

Part 2: Public Detectability. Suppose A_i has signed two conflicting successors: σ_i on $H_O \parallel A_{i+1}$ and σ'_i on $H_O \parallel A'_{i+1}$ with $A_{i+1} \neq A'_{i+1}$. By A6 (chain liveness), both signatures, having been published to the public chain, are observable to any verifier.

Define the predicate $\text{DoubleSign}(A_i, H_O)$ over the public chain that returns true iff there exist two valid signatures by A_i on conflicting commitments referencing H_O . The predicate is computable in polynomial time by any verifier with chain access. If σ_i and σ'_i both exist on the chain, the predicate returns true and the double signing is publicly detected.

The protocol's verifier rejects ownership chains where DoubleSign holds for any signer.

A.5 Full Sketch: Distribution Conservation (Theorem 1)

The theorem asserts that the verifier rejects distribution events whose outputs violate the metadata quantity invariant.

1. Let $\mathcal{E}$ be a signed distribution event with parent O_p and outputs $\{O_{p,1}, \dots, O_{p,M}, R\}$ (with R optional). $\mathcal{E}$ is committed to the public chain and signed by the burning owner A_n .
2. By Lemma 1, the canonical hashes $H_{O_p}, H_{O_{p,1}}, \dots, H_{O_{p,M}}, H_R$ uniquely identify their respective objects.
3. The metadata quantity $q(\cdot)$ is part of the canonical schema of each output; canonicalisation (Lemma 8) ensures that the verifier reads the same q value as the signing party committed.
4. The verifier computes $\sum_{k=1}^M q(O_{p,k}) + q(R)$ and compares to $q(O_p)$.
5. If the sum does not equal $q(O_p)$, the verifier returns Reject.
6. The signing party who fabricates an imbalanced distribution must commit to canonical outputs whose q values sum to a different value than the parent's q . The structural inequality is observable by any verifier; no signing strategy avoids the predicate.

The theorem establishes that the conservation invariant is structurally enforced.

A.6 Full Sketch: Composition Theorem (Theorem 3)

The theorem asserts that the per tier guarantees compose into a joint guarantee bounded by the minimum of the per tier failure probabilities.

Setup. Let G_i denote the guarantee of tier T_i against an adversary $\mathcal{A}$ under specified assumptions. Let $F_i = 1 - G_i$ be the corresponding failure probability.

Argument. The adversary's joint success against the composed system requires:

- Defeating T_1 : producing structurally valid signed records (fail probability F_1).
- Defeating T_2 : producing payloads that pass conservation, envelope, and cross vector tests (fail probability F_2).
- Defeating T_3 : producing finished product that classifies correctly under the reference base (fail probability F_3).
- Defeating T_4 : producing External Attestations that yield the desired posterior (fail probability F_4).

The events are sequential: the adversary's payloads are evaluated tier by tier. Conditional on T_1 acceptance, the T_2 test is applied, etc.

The joint success probability is therefore at most the minimum of the per tier failure probabilities, with equality when the failures are perfectly correlated:

$$\Pr[\mathcal{A} \text{ succeeds}] \leq \min(F_1, F_2, F_3, F_4).$$

Under independence (or even under correlation that does not perfectly align failures), the bound is conservative. In practice, an adversary that defeats T_1 structurally has different signatures than one that defeats T_2 statistically, and the failure modes do not align perfectly.

Soundness. The composition does not weaken any per tier guarantee: each tier’s verdict depends only on its own inputs, not on the other tiers’ outputs. Tier independence in evaluation (each tier reads the signed payload directly) prevents cascading weakening.

A.7 Full Sketch: Cartel Risk Bound (Theorem 6)

The theorem asserts that under composed Hardening Pathways, the cartel risk reduces to a quantifiable economic parameter.

Setup. The cartel adversary $\mathcal{A}_5$ aims to produce a fabricated supply chain narrative for a target gain G . The protocol’s defences against $\mathcal{A}_5$ under hardening are:

- Hardware: each Proplet device tampered at cost $\geq C_{\text{tamper}}$. For n devices in the supply chain, total hardware cost is $n \cdot C_{\text{tamper}}$.
- Stake: each attester’s stake S_i is slashed if fraud is detected. Total expected slashing exposure is $\sum_i p_{\text{detect},i} \cdot S_i$, where $p_{\text{detect},i}$ is the probability that attester i ’s fraud is detected.
- Composed detection: under Theorem 3, the joint detection probability across all tiers is bounded by the strongest tier; for the typical multi tier deployment, $p_{\text{detect}}^{\text{composed}} \rightarrow 1$ as the tiers compose.

Cartel utility. The cartel’s expected utility is

$$U_{\mathcal{A}_5} = G - n \cdot C_{\text{tamper}} - p_{\text{detect}}^{\text{composed}} \cdot \sum_i S_i.$$

For the cartel to be rational, $U_{\mathcal{A}_5} > 0$, hence

$$G > n \cdot C_{\text{tamper}} + p_{\text{detect}}^{\text{composed}} \cdot \sum_i S_i.$$

Conversely, for cartel deterrence, the deployment calibrates $\sum_i S_i$ such that the inequality is reversed for any economically plausible G .

Quantitative example. For a deployment with $n = 10$ Proplets at $C_{\text{tamper}} = 50,000$ EUR each, ten attestors at $S_i = 100,000$ EUR each, and $p_{\text{detect}}^{\text{composed}} = 0.99$, the cartel’s break even gain is

$$G_{\text{break even}} = 10 \times 50,000 + 0.99 \times 1,000,000 = 1,490,000 \text{ EUR}.$$

A cartel attack must yield over 1.5 million EUR to be rational, before considering legal and reputational costs that are off ledger but real.

The Hardening Pathway composition therefore reduces the cartel risk to a quantifiable economic parameter that the deployment chooses, calibrated against the value at stake.

A.8 On Mechanisation

The proof sketches in this appendix are sufficient for the structural reductions but stop short of mechanised formal verification. A successor work in the BDF programme will mechanise the lemmas in

a proof assistant (Coq, Lean, Isabelle/HOL) under the standard cryptographic primitive abstractions. The mechanisation will validate:

- The injectivity of canonicalisation under the beeson type system.
- The reduction of BDT signature unforgeability to EUF-CMA.
- The structural verifier rejection predicate for distribution conservation.
- The composition theorem under explicit independence assumptions.

The mechanisation is an investment in long term protocol credibility but is not a precondition for ICVCM CCP labelling or for registry adoption at the present stage.

B Adversary Capability Matrix

This appendix presents the full adversary capability matrix mapping each adversary class to each trust assumption and each tier guarantee. The matrix supports registry assessment by making explicit which assumptions are invoked for which guarantees against which adversaries.

B.1 Adversary versus Trust Assumption Matrix

Table 10 maps each adversary class to the trust assumptions it can plausibly violate. A cell marked • indicates that the adversary class can attempt to violate the corresponding assumption directly; a blank cell indicates that the adversary class is not the operative threat for that assumption.

Table 10: Adversary classes and the trust assumptions they target.

Adversary	A1	A2	A3	A4	A5	A6	A7	A8
$\mathcal{A}_1$ Lone fraudster				•				•
$\mathcal{A}_2$ Coalition $k < n$				•				•
$\mathcal{A}_2$ Coalition $k = n$				•				•
$\mathcal{A}_3$ Network adversary			•		•			
$\mathcal{A}_4$ Computational	•							
$\mathcal{A}_5$ Coordinated cartel				•			•	•

Reading the table. Each adversary class targets specific assumptions. The lone fraudster targets attester honesty (A4) and may also exploit time skew (A8) to backdate payloads. Coalitions extend A4 attacks across multiple parties. The network adversary targets routing diversity (A3) and the postage economics (A5) that underpins data availability. The computational adversary targets the cryptographic primitives (A1). The coordinated cartel targets attester honesty (A4), laboratory accreditation (A7), and time consensus (A8) simultaneously.

The cryptographic and structural assumptions A1, A2, A6 do not appear as direct adversary targets in the matrix because no realistic adversary can attack them at the protocol layer; they fail only through primitive break (A1), implementation bug (A2), or chain consensus failure (A6), each of which is outside the adversary typology of Section 2.

B.2 Adversary versus Tier Guarantee Matrix

Table 11 expands the summary matrix from Section 2 with the specific lemma or proposition that bounds the adversary’s success at each tier.

Table 11: Detailed adversary versus tier guarantee matrix. Each cell references the bounding lemma or proposition. Bound notation: negl for negligible probability, FNR for false negative rate, Φ for the standard normal CDF. Empty cells: not applicable to that adversary class at that tier.

Adversary	T_1	T_2	T_3	T_4
$\mathcal{A}_1$ Lone fraudster	negl (Lemma 2)	FNR^{T_2} (Prop. 3)	1 – Identifiability (Prop. 6)	1 – θ_4 (Prop. 10)
$\mathcal{A}_2$ Coalition $k < n$	negl (chain consistency)	Cross vector band $\Phi(k) - \Phi(-k)$	Centroid shift (Prop. 7)	Multi lab posterior
$\mathcal{A}_2$ Coalition $k = n$	negl (structural)	No bound (A4 violated)	No bound (A4 violated)	Multi lab (still A7)
$\mathcal{A}_3$ Network adversary	Inherited from A6	Operates on retrieved data	Reference base local	External attestation
$\mathcal{A}_4$ Computational	negl (A1)	Inherits T1 crypto floor	Inherits T1 crypto floor	Inherits T1 crypto floor
$\mathcal{A}_5$ Coordinated cartel	negl structural only	No bound (cartel)	No bound (cartel)	No bound base config

Reading the table. The cell entries are the analytic bounds from the body sections. The T_1 floor is uniformly negligible against all adversaries except $\mathcal{A}_3$ (where it depends on chain liveness). The T_2 statistical layer holds against $\mathcal{A}_1$ and $\mathcal{A}_2(k < n)$ but cannot bound $\mathcal{A}_2(k = n)$ or $\mathcal{A}_5$ in the base configuration. T_3 identifiability is similar. T_4 holds against single laboratory corruption only when multiple laboratories cross check.

B.3 Hardening Pathway Coverage Matrix

Table 12 maps each Hardening Pathway to the assumption it substitutes for and the adversary class it newly covers.

Combined coverage. A deployment using all four Hardening Pathways simultaneously covers $\mathcal{A}_5$ structurally: hardware cost, economic disincentive, laboratory redundancy, and information asymmetry reduction together require the cartel to overcome multiple independent defences simultaneously. The composition is not provably tight; the practical effect is that $\mathcal{A}_5$ cost grows superlinearly in the number of pathways deployed.

B.4 Quick Reference: Bounds at a Glance

The following compact summary supports rapid assessment:

Table 12: Hardening Pathway coverage. Each row identifies a pathway, the assumption it substitutes for, the new assumption it introduces, and the adversary class it newly covers.

Pathway	Substitutes	Introduces	Newly covers
Proplet attestation	A4	H1 (tamper resistance), H2 (remote attestation), H3 (sensor calibration)	$\mathcal{A}_2(n, n)$ at hardware cost; partial $\mathcal{A}_5$
Cryptoeconomic stake	A4 (partial)	Stake calibration, fraud proof economics	$\mathcal{A}_2$ economically; partial $\mathcal{A}_5$
Multi laboratory T_4	A7 (single)	A7 redundancy across L labs	Coordinated lab corruption up to $\lfloor L/2 \rfloor$
Zero knowledge proofs	N/A	ZK soundness assumption	Information advantage of $\mathcal{A}_5$

T_1 bound.

$\text{Adv}_{\mathcal{A}_4}^{\text{coll}} + n \cdot \text{Adv}_{\mathcal{A}_4}^{\text{euf-cma}}$ for chain length n , both negligible at 128 bit security.

T_2 envelope FNR.

$F_X^{\text{adv}}(X_{\max}) - F_X^{\text{adv}}(X_{\min}^-)$ for adversary distribution $\mathcal{D}_X^{\text{adv}}$.

T_2 cross vector band.

$\Phi(k) - \Phi(-k)$ for band width $k\sigma_f$.

T_3 identifiability.

$1 - (K - 1) \exp(-d_{\min}^2/8(\sigma_w^2 + \sigma_n^2))$.

T_4 single lab.

Posterior $\geq \theta_4$ requires $\Lambda \geq \theta_4(1 - p)/((1 - \theta_4)p)$.

T_4 multi lab.

Aggregate $\Lambda^* = \prod \Lambda_i$, exponential in L honest labs.

DREAM revocation.

$\Pr[\text{succeed}] \geq 1 - (1 - h)^n$ for honesty rate h and path length n .

Cartel break even.

$G \leq n \cdot C_{\text{tamper}} + p_{\text{detect}}^{\text{composed}} \cdot \sum_i S_i$ for cartel deterrence.

These bounds are the principal numerical results of the paper. A registry assessor or implementer can plug deployment specific values into each bound to derive concrete confidence levels for the deployment's stated guarantees.

Document History

Versions follow semantic versioning: a major increment changes protocol semantics or normative requirements, a minor increment adds material without breaking existing conformance, a patch corrects errors. The current version is stated on the title page and in the PDF metadata.

Version	Date	Changes
0.1	May 2026	Initial draft, companion document 3 of the BDF corpus.
1.0	24 September 2026	First public release. Version and date fixed, licensed under CC BY 4.0, document history added. No technical changes from the May draft.

References

- Mihir Bellare and Phillip Rogaway. The exact security of digital signatures: How to sign with RSA and Rabin. *Eurocrypt*, 1996.
- Vitalik Buterin, Daniel Reijnders, Stefanos Leonardos, and Georgios Piliouras. Incentives in Ethereum’s hybrid Casper protocol. *International Journal of Network Management*, 30(5), 2020.
- George Coker, Joshua Guttman, Peter Loscocco, Amy Herzog, Jonathan Millen, Brian O’Hanlon, John Ramsdell, Ariel Segall, Justin Sheehy, and Brian Sniffen. Principles of remote attestation. In *International Journal of Information Security*, volume 10, pages 63–81, 2011.
- Fair Food Data. The Bio-Digital footprint (BDF) protocol: Verifiable records for transformed bioassets via Bio-Digital twins and tiered attestation. 2026a. Companion paper, version 1.0, September 2026. <https://fairfooddata.org>.
- Fair Food Data. BDF methodology for recovered food carbon. 2026b. Companion paper, version 1.0, September 2026. <https://fairfooddata.org>.
- Fair Food Data. BDF standards stack: Implementation reference. 2026c. Companion paper, version 1.0, September 2026. <https://fairfooddata.org>.
- Shafi Goldwasser, Silvio Micali, and Charles Rackoff. The knowledge complexity of interactive proof systems. In *SIAM Journal on Computing*, volume 18, pages 186–208, 1989.
- Jens Groth. On the size of pairing-based non-interactive arguments. In *EUROCRYPT 2016*, pages 305–326, 2016.
- Integrity Council for the Voluntary Carbon Market. Core carbon principles (CCP) assessment framework. Online, 2024. URL <https://icvcm.org/>.
- International Laboratory Accreditation Cooperation. ILAC mutual recognition arrangement. Online, 2024.
- International Organization for Standardization. ISO/IEC 17025:2017 general requirements for the competence of testing and calibration laboratories, 2017.

- Don Johnson, Alfred Menezes, and Scott Vanstone. The elliptic curve digital signature algorithm (ECDSA). *International Journal of Information Security*, 1(1):36–63, 2001.
- Petar Maymounkov and David Mazières. Kademlia: A peer-to-peer information system based on the XOR metric. *Peer-to-Peer Systems*, pages 53–65, 2002.
- Ralph C. Merkle. A digital signature based on a conventional encryption function. *Advances in Cryptology — CRYPTO ’87*, pages 369–378, 1987.
- NIST. FIPS 180-4 secure hash standard (SHS), 2015.
- NIST. FIPS 186-5 digital signature standard (DSS), 2023.
- Christian P. Robert. *The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation*. Springer, 2nd edition, 2007.
- Nick Szabo. Proplets: Devices for controlling property. Online, 2001. URL <https://nakamotoinstitute.org/library/proplets/>.
- Viktor Trón. *The Book of Swarm: Storage and Communication Infrastructure for a Self-Sovereign Digital Society*. Self-published, 2024. ISBN 978-615-01-9983-2.
- Viktor Trón and Henning Diedrich. A dream come true: Deletable content in immutable storage. *ACM (manuscript)*, 2025. Manuscript submitted to ACM, September 2025.
- Viktor Tron et al. Beeson: A canonical object notation for content-addressed data. 2026. Forthcoming.
- Trusted Computing Group. TPM 2.0 library specification, part 1. Online, 2019.