GenLayer: The Intelligent Contract Network

Technical Whitepaper

Albert Castellana José María Lago Edgars Nemše

GenLayer

Version 2.3

Supersedes the 2024 vision paper (v1) as the canonical technical document.

Note. Citations [P1]–[P4] resolve to arXiv identifiers upon publication of the research series. Token supply, distribution, and epoch-level economics are specified in the Tokenomics companion document (forthcoming); this document describes protocol mechanics and contains no offering language. All plots in this document are generated by the published validation pipelines — they are measurements, not illustrations.

Abstract

GenLayer is a decentralized network for *Intelligent Contracts*: programs that combine deterministic on-chain state with natural-language rules, LLM calls, and live web data. GenVM exposes every non-deterministic operation to validator review, while each contract defines an *Equivalence Principle* specifying when different outputs should count as equivalent. Optimistic Democracy turns those judgments into protocol decisions through stake-weighted committee selection, commit–reveal voting, and bonded appeals. This whitepaper describes the implemented execution environment, consensus state machine, fee and staking mechanisms, and verification stack. It also separates protocol mechanics from model-dependent guarantees: finite-panel results hold under declared evaluator and sampling assumptions and certify reproducibility relative to a chosen evaluator population, not semantic truth. State-machine and fee-accounting properties are checked against pinned formal and executable specifications, while full history-dependent sampling security and appeal-pricing calibration remain open research questions. The resulting platform supports intelligent oracles, natural-language governance, trustless world data, and agentic commerce.

Contents

1 Introduction **3**

2 GenVM and Intelligent Contracts **4**

3 The decision layer: what is provable **5**

4 Optimistic Democracy: the implemented protocol **6**

5 Design margins and a possible contentious outcome **9**

6 Economic design: what is machine-checked **9**

6.1 Who pays is different from whether to continue . . . 10
6.2 Validators are paid for round-outcome alignment . . . 10
6.3 Escalation is self-financing . . . 11
6.4 Covered-path financial griefing bounds . . . 11
6.5 Honest, open item: the appeal-pricing upgrade . . . 11
6.6 Staking . . . 12

7 Security: characterized failure modes and countermeasures **12**

8 Verification: don't trust — run it **13**

9 Use cases **14**

10 Token **14**

11 Conclusion **14**

A Protocol parameters (pinned revision) **15**

1 Introduction

Smart contracts made agreements self-executing, but they only speak code. The moment an agreement requires judgment — *is this deliverable of professional quality? did the news actually confirm this event?* — current platforms hand the question to trusted oracles or human committees. Large language models can render such judgments, but they are non-deterministic and disagree with one another, which is precisely what classical Byzantine fault-tolerant consensus cannot absorb: in classical BFT, disagreement among honest nodes is impossible by assumption, so any disagreement is evidence of a fault.

GenLayer's position is that this is not an engineering inconvenience but a new consensus problem — **consensus over ambiguous specifications** — and that it deserved a theory before it deserved a mainnet. A four-paper research series [P1–P4] now formalizes the statistical target, characterizes adaptive evaluator spending and an idealized appeal ladder, audits a pinned implementation baseline, and analyzes whether private incentives implement the social stopping rule. The series deliberately does not claim semantic truth, optimality of threshold voting, equilibrium optimality of deployed appeals, or automatic transfer from a one-round sampling theorem to the deployed history-dependent stake-weighted process. This whitepaper is the protocol-level distillation of that work. Its structure mirrors the system's layers (Figure 1):

- **GenVM** executes Intelligent Contracts and mediates their non-deterministic operations (Section 2);
- the **decision layer** turns validator judgments into decisions with model-conditional guarantees (Section 3);
- **Optimistic Democracy** is the implemented protocol around that decision rule — committees, appeals, timeouts, tribunal (Section 4);
- the **economic layer** pays for round-outcome alignment and prices escalation (Section 6);
- the **routing layer** defines and exposes population choices that must be measured against the theory's assumptions (Section 7);
- a three-tier **verification stack** pins all of the above to pinned code baselines (Section 8).

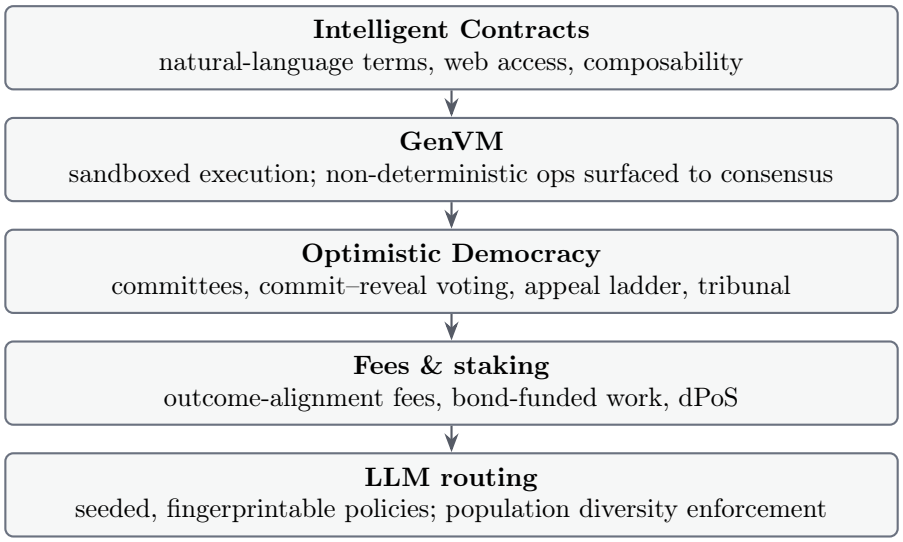

Figure 1: The GenLayer stack. Each layer of this document corresponds to a layer of the system; the verification stack (Section 8) cuts across all of them.

Relation to the vision paper. The 2024 whitepaper [v1] introduced Intelligent Contracts, GenVM, the Equivalence Principle, and the optimistic-consensus outline. The companion divergence audit now classifies each material v1 claim as retained, superseded, model-dependent, or

open. A number of mechanisms changed during implementation and analysis. This document is self-contained; where v1 described a mechanism that changed, the change and its reason are stated.

How to read this document. Throughout, protocol descriptions refer to the pinned implementation baseline; mathematical guarantees hold under their stated models and sampling assumptions; numerical studies measure the behavior of those declared models rather than live-network performance; and unresolved deployment questions are identified explicitly. Detailed limitations are stated once, in the section where the corresponding result lives.

2 GenVM and Intelligent Contracts

GenVM is a restricted, sandboxed Python-based execution environment. It is account-based (externally owned accounts and contract accounts holding GEN balances); contracts are classes with persistent state; transactions invoke methods with parameters and pre-funded fee budgets. Deterministic execution behaves exactly as in classical smart-contract VMs, including atomic deterministic contract-to-contract calls.

Two operations make contracts *intelligent*, and both are non-deterministic: **LLM calls** (natural-language instructions whose output depends on the model and seed) and **web reads** (live data that changes over time). GenVM surfaces every such operation to the consensus layer with its exact input, so that other validators can re-evaluate it.

The Equivalence Principle. Different validators will produce different-but-acceptable outputs for the same non-deterministic call (two correct summaries of the same article differ textually). The contract developer therefore specifies, per call, an *equivalence principle*: the criterion under which a validator judges whether the leader's output is an acceptable answer. Operationally the validator evaluates a binary query — *given criterion E, is output O acceptable (possibly against my own re-execution O')?* — and votes accordingly. The research series formalizes exactly this object: a task is a binary evaluation query $t = Q(EP, w_1, w_2)$, and every guarantee in Section 3 is stated per task [P1].

Dependent transactions and messages. Appeals can revise a transaction's outcome within its finality window, so dependent computation must be bounded. As in v1: contracts read other contracts' *final* state (or explicitly accept non-final reads as final-at-time-of-reading), and non-deterministic cross-contract effects travel as **messages** — non-revertible, emitted either at first acceptance or at finality, at the developer's choice. This breaks recomputation cascades by design.

Fees at the VM level. Transactions carry pre-funded time-unit budgets (L for leader work, W per validator; Section 6) covering the worst case of the rounds they may trigger, with unspent budget refunded at settlement. This replaces the gas-auction sketch of v1 with the mechanism that is implemented and machine-checked today.

One transaction, end to end. Before the theory, the whole lifecycle once, in ordinary language. A contract promises payment if a submitted report meets a natural-language quality standard. When the transaction arrives, a stake-weighted lottery selects a leader and a small committee of validators. The leader executes the contract in GenVM — the deterministic parts run as on any smart-contract platform, and the quality judgment is an LLM call whose input and output are published in the receipt — and proposes an outcome. Each validator independently re-executes the contract, re-derives its own answer, and judges the leader's output under the

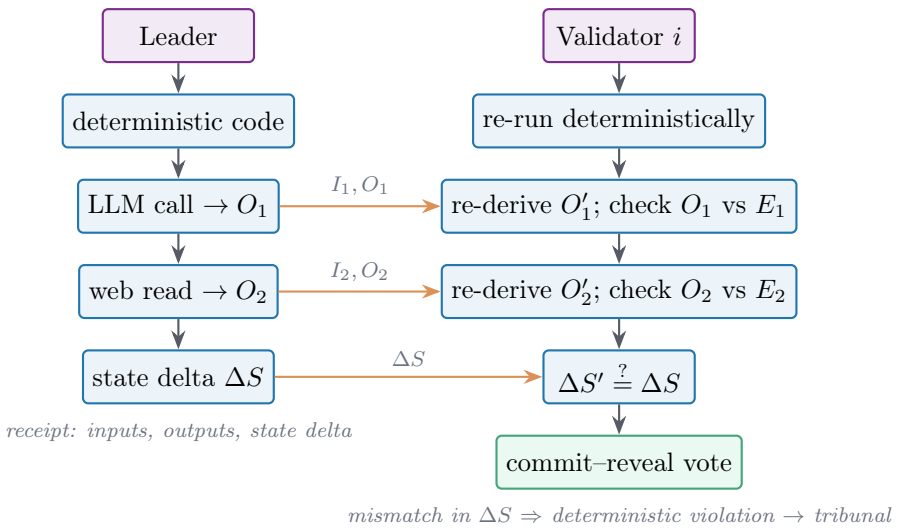

Figure 2: Execution and validation of one transaction. The leader publishes the inputs and outputs of every non-deterministic call plus the final state delta; validators re-execute deterministically with the leader's outputs, independently re-derive their own outputs, and judge the leader's under the per-call Equivalence Principle. A state-delta mismatch is a *deterministic violation* — provably wrong execution, handled by the tribunal, not by voting.

declared Equivalence Principle (Figure 2); the committee then votes by commit–reveal, so no vote is visible before all are cast. If the result is challenged, the appellant posts a bond that funds another, larger evaluation. Once the appeal window closes, the resulting state becomes final and the payment executes. Each section that follows makes one step of this story precise.

3 The decision layer: what is provable

The model [P1]. Each evaluator episode combines an interpreter configuration, private random inputs, and a frozen task. Under a declared superpopulation law, its binary verdict has mean acceptance rate μ; under a frozen finite census, panels instead sample fixed verdicts without replacement. A K-vote panel uses the declared threshold rule, and at threshold 1/2 its superpopulation reference decision is $D_\infty = \mathbb{1}[\mu \geq 1/2]$.

Four ideas [P1]. The research behind this layer is extensive; the protocol reader needs four of its conclusions.

- **Panel decisions are relative to a declared evaluator population.** Every guarantee is a statement about reproducing D_∞ — the decision of a stated population of evaluator episodes — not about an external ground truth.
- **The correct statistical law depends on how evaluators are sampled.** Conditional on a frozen census, panel error is an exact hypergeometric tail; under i.i.d. evaluator episodes it is an exact binomial tail, with the Hoeffding bound $P(\text{decision} \neq D_\infty) \leq e^{-2K\gamma^2}$ away from the threshold. A bound proved for one sampling regime is not automatically valid for another.
- **Clearer tasks need fewer evaluators.** The **clarity** $\gamma = |\mu - 1/2|$ enters the error exponent as γ^2 (Figure 3): clear tasks are decided reliably by small committees, and no committee size repairs a task whose population mean sits at the threshold.
- **Reproducibility is not semantic truth.** Vote-only data certify agreement with the declared population; the independence premise holds only when episodes really use independent private inputs, and no optimality of threshold voting is claimed.

Adversarial profiles, operational certification of workloads, and the routing-layer consequences of these facts are developed where they are used (Sections 5 and 7).

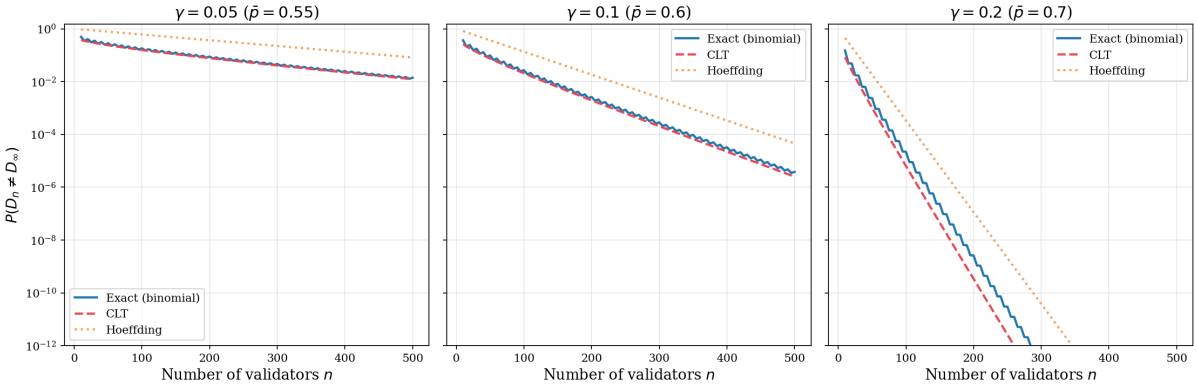

Figure 3: Decision error versus committee size for three clarity levels — exact computation, Monte Carlo, and the Hoeffding upper bound $e^{-2K\gamma^2}$ (measurement from [P1]'s published validation pipeline). Clear tasks are decided reliably by small committees; the exponent is γ^2, which is why clarity, not model perfection, is the governing quantity.

What the real-LLM pilot establishes [P1]. A hash-pinned diagnostic campaign (50 frozen units, 40 routed evaluator rows) demonstrates that the full pipeline — generation, evaluation, certification, publication — runs end to end. Its most instructive result doubles as its central limitation: nine of 13 units constructed to be semantically underdetermined certified at $K = 47$, so evaluator coordination can be strong even when the underlying question has no unique semantic answer. The pilot does not identify a deployment workload law or external truth.

4 Optimistic Democracy: the implemented protocol

The state machine described here is the one formally verified against the implementation (Section 8) — as implemented, not as envisioned. Figure 4 shows the transaction lifecycle.

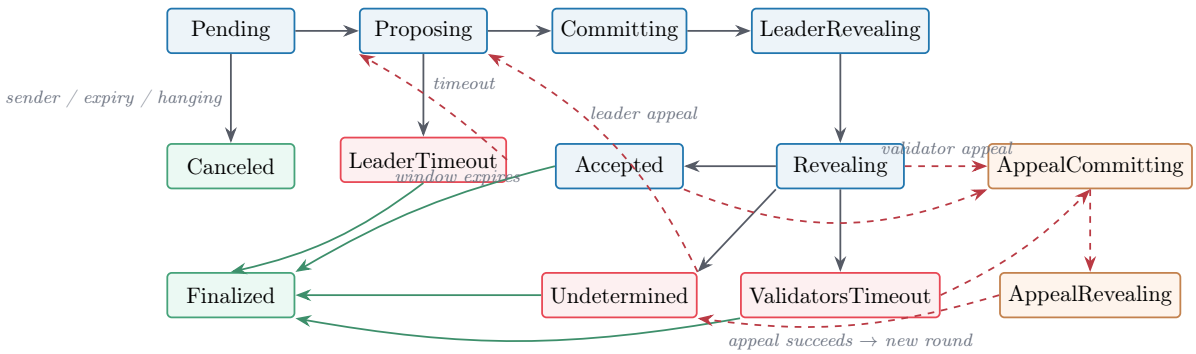

Figure 4: Simplified transaction lifecycle (the canonical machine has 14 stored states plus one virtual state and 42 effective transitions; this figure omits rotation self-loops, the terminal-capacity mode, the tribunal, and revert paths). Solid: normal flow. Green: finalization when the appeal window expires. Dashed red: the appeal surface. *ReadyToFinalize* is computed, never stored: finalization writes *Finalized* directly.

Pending termination. `validUntil` is the latest permissible activation time, not a whole-lifecycle time-to-live. Equality is still activatable; expiry is strict and applies only while the transaction is `Pending`. Once activation reaches `Proposing`, phase-specific timeouts govern progress. Expiry, sender cancellation, and hanging abort are distinct termination causes, checked in that order. Expiry terminates without first rotating or penalizing the activator or incrementing the recomputation count.

Committees and voting. The validator universe is frozen for each transaction. A stake-weighted VRF selects the activator, the leader, and the committee from that universe. The pinned VRF selection design prevents validator-side selection grinding under its stated beacon assumptions; deployment security additionally depends on seed integrity and the history-dependent inclusion law (Section 7). Validators vote by **commit–reveal** with a leader-reveals-first rule: votes are sealed until the relevant reveal phase, limiting within-round vote copying. This supplies the simultaneity premise of Part 4's auxiliary coordination game; it does not itself ensure statistical independence or eliminate pre-coordinated herding conventions (Section 6).

Outcomes. The protocol resolves a strict-majority outcome category: agree, timeout, disagree, or deterministic violation; if none has a strict majority, the outcome is no majority. Agree accepts; timeout produces `ValidatorsTimeout`; disagree or no majority replaces the leader when an exact replacement is available and otherwise produces `Undetermined`. On an ordinary reveal timeout, missing ballots count as disagree and the non-revealers are sanctioned. On an appeal reveal timeout, missing ballots instead support the challenged status quo.

A *deterministic violation* — a state-delta mismatch, which is provable wrongdoing unlike judgment disagreement — convenes the **tribunal** in parallel with the transaction's continued progress. The original round's ballots are pre-counted and its committee members do not vote again. The external electorate is the frozen validator universe minus that committee; it commits and reveals against the frozen tribunal quorum.

The appeal ladder. Any party may appeal a decision during its appeal window by posting a bond (Figure 5). The nominal FeeManager entries are

$$5, 7, 11, 13, \ldots, 767, 769, 1535, 1537.$$

The realized path is not this fixed interleaving: appeal feasibility depends on the round and the available unconsumed validators, and the companion specification applies a minus-two sizing brake after consecutive unsuccessful appeals. Every successor has completely reset commit/reveal state: ballots are never carried across rounds. A normal successor re-enters proposing with a new leader and execution result; an appeal jury re-evaluates the challenged proposal. Validator identities need not all be new. Appeal juries draw from previously unconsumed validators, while normal successors reuse eligible survivors and fill their remaining places from the unconsumed set. Leader replacement after majority disagreement stays within the current normal round; recovery after `LeaderTimeout` or `Undetermined` likewise constructs a normal successor from eligible survivors.[1]

An appeal succeeds when the appeal jury's outcome category differs from the challenged category. No majority and deterministic violation are categories for this comparison, so either can constitute a successful challenge.

Capacity behavior. Contract appeal feasibility is checked against the current round and available unconsumed validators. Successor construction can combine the prior normal committee (minus its leader) with the appeal jury, or snapshot the active set when the configured

[1] v1 described appellate votes as added to the initial votes. The protocol instead resets the ballots in every successor round, including when validator identities are reused.

condition is met. Voting always tests strict majority of the realized committee length; there is no separate frozen-universe quorum in the pinned voting contract.

For the documented $N = 1543$ scenario, eight appeal juries through size 769 can lead to a 1535-member normal successor. Its ballot state is reset and its strict-majority threshold is 768. This is an illustrative capacity path, not a protocol-wide appeal-count invariant.

Five useful distinctions from the adaptive statistical analysis [P2]:

- Part 2's G-AE–ADS represents the deployed distinction explicitly: an appeal jury moves down the evaluator rows of one frozen resolution, while a return to proposing can move to a newly generated column. Its fixed-column formulas remain benchmarks, and the deployed generator transition kernel is not estimated;
- for a declared workload, likelihood, evaluator cost, latency, and decision loss, a finite-horizon Bellman recursion identifies the social planner's action. One four-round mixture instance is solved exactly; it is illustrative rather than fitted to deployed transaction values;
- on the non-repeated nominal size grid through 1535, an oracle that stops at the first adequate rung spends less than six times the evaluator seats of one adequate panel. Neither this schedule result nor the planner result proves that private appellants implement the rule;
- an oracle that observes the population decision has an $O(1/\gamma^2)$ cost benchmark on a geometric statistical ladder, but that side information prevents combining the result with a vote-only lower bound to claim protocol optimality;
- uniform sampling without replacement admits Serfling concentration, and a separate theorem covers one successive probability-proportional-to- stake draw around its inclusion-probability design mean. Multi-round identity reuse and replacement, the distinguished leader seat, seed and shared-state assumptions, the minus-two path, and the capacity construction still require history-uniform transfer results.

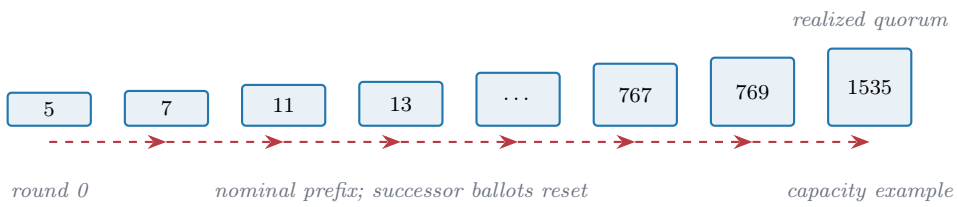

Figure 5: The appeal ladder. Bars show nominal committee sizes; validator identities follow the reuse and fresh-allocation rules in the text, while every successor round resets its ballots. Each escalation requires a bond equal to the worst-case payroll of the round it triggers; bonds grow geometrically. Section 6 gives covered-path financial accounting; neither fact makes the challenged outcome immutable.

Benchmark error correction. Figure 6 reports Part 2's homogeneous independent-round simulation, not a deployed network estimate. Votes are i.i.d. within and across model rounds and a single *watcher* has a fixed, calibrated probability $p_a = 0.9$ of agreeing with the Part 1 population decision. At $\gamma = 0.2$ the model's 5-vote round disagrees with that decision 16.3% of the time and the watcher policy produces about 2.0% end-to-end disagreement. The result does not cover the deployed multi-round weighted selection process, validator reuse or replacement, distinguished leader behavior, path-dependent sizing/capacity, changing normal-round resolutions, or an empirically uncalibrated watcher.

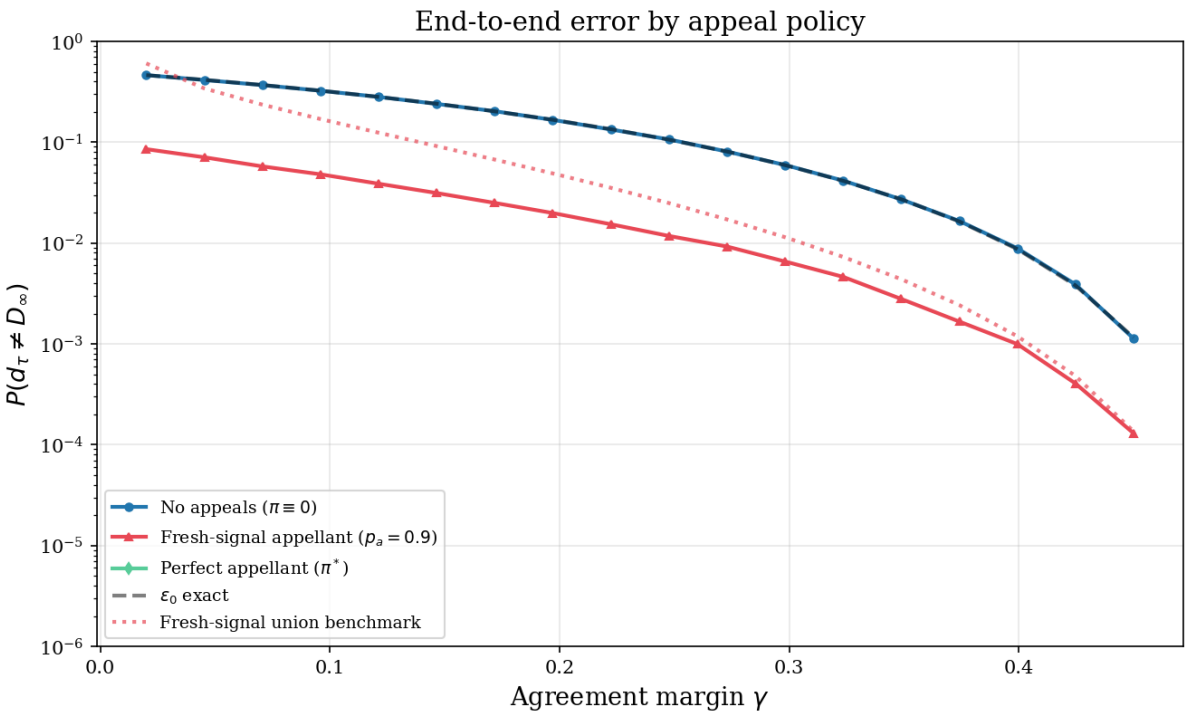

Figure 6: End-to-end population-disagreement probability in Part 2's homogeneous independent-round benchmark. The fresh-signal watcher can substantially reduce model error under its calibrated assumptions. The plug-in-score gate at 2/3 coincides with no appeals and is therefore omitted from the plot; this is a property of an uncalibrated score, not proof that every public-information appeal is negative expected value. The oracle curve is unimplementable.

5 Design margins and a possible contentious outcome

Near-threshold tasks require larger panels in the homogeneous model. At $n = 1535$ and target disagreement probability 10^{-6}, Hoeffding supplies the *sufficient* margin $\gamma \approx 0.0671$; exact homogeneous binomial calculation reaches the target at about 0.06043 (Figure 7) [P2]. Neither number is a necessary clarity floor. Failure to meet either criterion is inconclusive about another decision rule, sampling law, or appeal policy.

A distinct **contentious** outcome remains a promising protocol-design option, not an implemented consequence of Part 1. The operational theorem can calibrate a fixed K using a separate, predeclared generator–evaluator matrix and then certify future independent panels over a declared workload. It does not authorize estimating a live task's clarity from the same votes that decide that task and treating non-certification as proof of ambiguity. Deploying a live flag therefore needs its own anytime- or optional-stopping-valid calibration, explicit false-contentious/false-decision losses, and integration with the current outcome categories. Until then, `Undetermined` is the protocol outcome and "contentious" is future work.

6 Economic design: what is machine-checked

The fee layer has an **executable specification**: a Python mirror of the fee contracts, exercised over every generated simulator path, checking 24 invariants on each; parity with the contracts is enforced by generated test vectors. Three invariants are the mechanism's economic axioms — no profit from griefing, cost of contention, bounded amplification — and the results below are proven against them [P4].

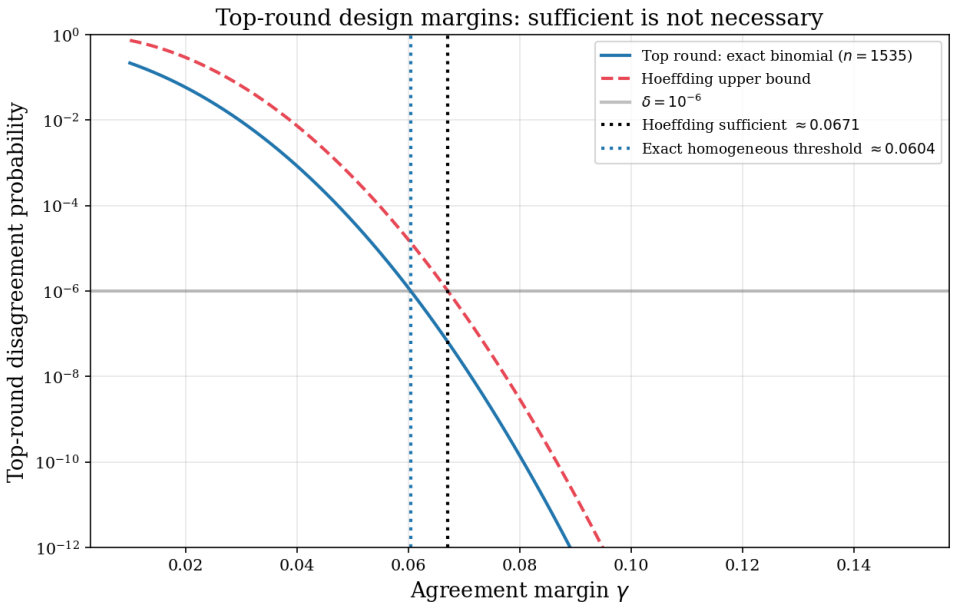

Figure 7: Top-round design margins in Part 2's homogeneous model. The exact binomial curve reaches 10^{-6} at $\gamma \approx 0.06043$; Hoeffding gives the conservative sufficient value 0.0671.

6.1 Who pays is different from whether to continue

Part 2's general planner may stop, re-evaluate the current resolution, or regenerate a new one. Part 4 studies its binary stop–re-evaluate interface: the next panel is purchased only when some appellant expects a positive private return. Its central negative result is that, when social loss and private transaction value may vary independently, even a bond which exactly funds the triggered work cannot align these decisions uniformly — no fixed bond and reward work across all workloads. Part 4 also prices the resulting welfare loss exactly, as an expected sum of per-round Bellman gaps along the equilibrium path, so premature stopping and excessive escalation are measured in the planner's own welfare units; estimating those terms requires a declared workload and loss model together with a calibrated multi-appellant equilibrium [P4].

6.2 Validators are paid for round-outcome alignment

Fix the two fee units: L (leader work) and W (one validator's work). Per determined round, a validator earns $+W$ when voting with the round's outcome and forfeits W otherwise (the symmetric coefficient $\kappa = 1$); undetermined rounds pay all participants. The direct accounting identity is $W(2q_i - 1)$ in a determined round, where q_i is the probability of matching that round's outcome. Under the additional homogeneous signal model used in [P4], this becomes

$$\mathbb{E}[\pi_i] \;=\; W\,(2p_i - 1)(1 - 2\varepsilon)$$

where p_i is agreement with the declared population decision and ε is the round's population-disagreement probability. This factorization requires a stated conditional-independence model and neglects the validator's pivotal effect; it is not a semantic-accuracy guarantee or a theorem for the deployed stake-weighted sampler (Figure 8). Two deliberate choices shape the accounting: the symmetric *fee* penalty — equivalence-vote minorities are **not** stake-slashed[2] — and pay-all

[2] v1 stated that appealed responses cost validators "rewards and staked tokens." The implemented design deliberately does not equate ordinary disagreement with provable wrongdoing. Raising minority penalties changes participation risk on low-coordination workloads and should be stress-tested rather than justified by calling the minority dishonest. Stake slashing is reserved for idleness (1%) and deterministic violations (1% validator / 5% leader, capped per epoch).

on undetermined rounds.

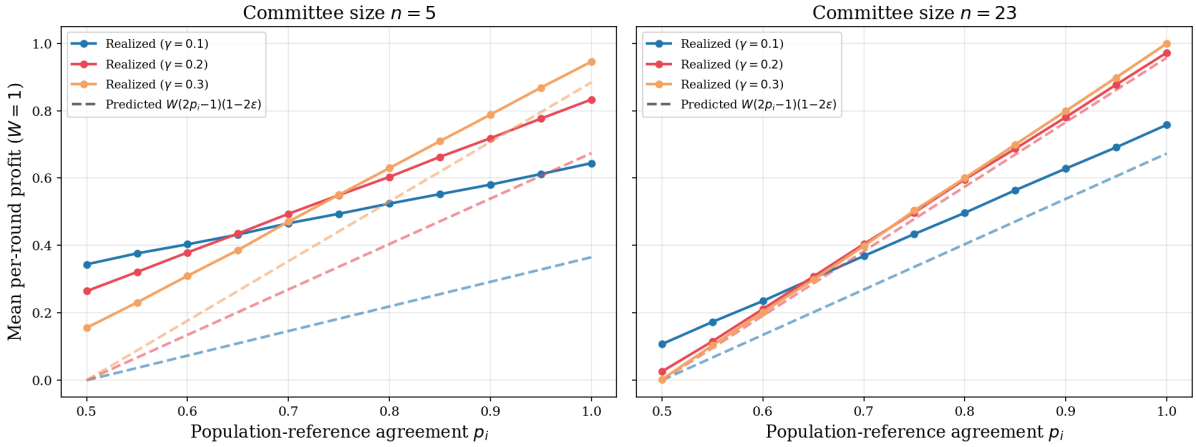

Figure 8: Outcome-alignment fees in the homogeneous independent-ballot model: realized mean per-round profit versus population-reference agreement, against the conditional factorization, at committee sizes 5 and 23. The deviation at $n = 5$ is the validator's pivotality effect. The plot is not semantic-accuracy calibration.

6.3 Escalation is self-financing

Each appeal bond equals the worst-case payroll of exactly what it triggers, per appeal type (verified against the contracts' bond computation and the simulator):

$$B_{\text{val}} = n'W, \qquad B_{\text{lead}} = (\rho + 1)\,(L + mW), \qquad B_{\text{lt}} = (\rho + 1)\,(L + n_{\text{cur}}W),$$

for validator appeals (n' new validators re-evaluate the existing proposal; appeal rounds are leaderless, so no L component), leader appeals from Undetermined (a full new normal round of size m with up to ρ budgeted rotations), and leader-timeout appeals (the rotated existing committee). A successful appeal returns 1.5× the bond; a failed one is forfeit, and the bond-split rules route it to the validators who did the re-evaluation. Escalation never draws down the sender's budget beyond what was pre-funded. This funding identity is not a participation theorem: the nominal validator appeal bond grows from $7W$ to $769W$ over the displayed eight-jury sequence, so the set of actors able to post it can shrink by tier.

6.4 Covered-path financial griefing bounds

On non-voided simulator paths covered by the economic invariants, minority griefing yields no positive coalition profit, a failed appeal costs at least the bond, and financial damage to others does not exceed the attacker's cost. These are accounting guarantees, not outcome immunity: an additional round can replace a population-consistent current decision with a population-inconsistent successor. The current contracts separately impose round-index and available-validator feasibility checks.

6.5 Honest, open item: the appeal-pricing upgrade

The flat bond with 1.5× return sets a 2/3 financial break-even success probability for a risk-neutral appellant who values only the bond payoff. That threshold is a fact about the pricing, not about the workload: it matches the social planner's escalation threshold only when the workload-specific threshold happens to equal 2/3. Part 4's positive result is that, under a calibrated single-crossing model, there exists a success reward that implements the planner's threshold exactly;

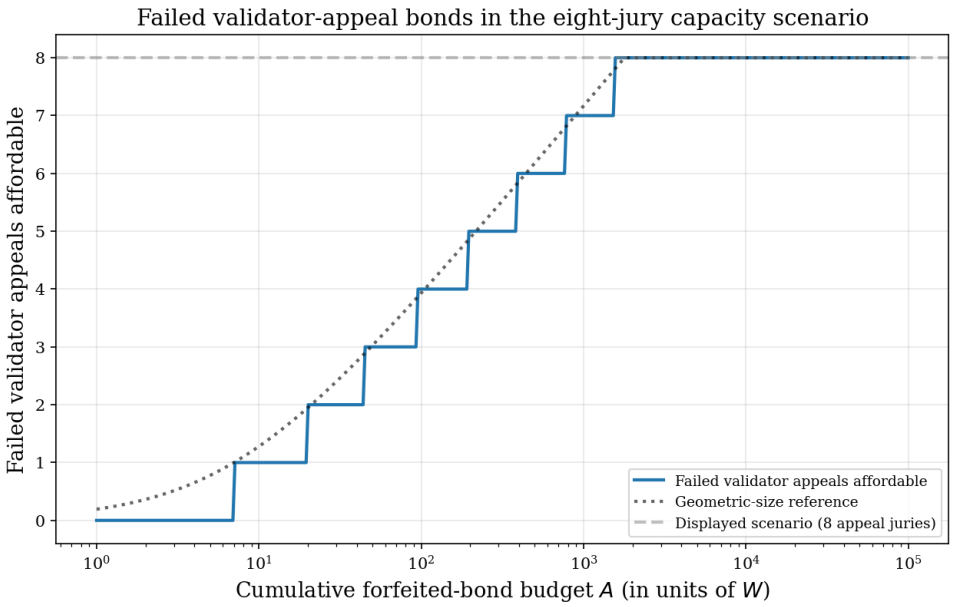

Figure 9: Nominal validator-appeal bonds along the documented eight-jury capacity scenario, assuming every displayed appeal fails and its bond is forfeited. The staircase counts how many such forfeitures fit in a budget; it does not model elapsed delay, successful-return recycling, outcome transitions, the minus-two brake, capacity checks, reuse, or voided paths.

the closed form and its assumptions are given there [P4]. With several potential appellants, the first-valid-transaction game has no-entry, full-entry, and interior regimes — the interior combines a positive probability of missed appeals with redundant submissions — race-style priority competition can dissipate the appellant's rent, and growing bonds can exclude capital-constrained actors. None of these declared stress models estimates the deployed sequencer or appellant wealth. The zero appeal rate in Figure 6 comes from applying 2/3 to Part 2's *uncalibrated* plug-in score; it is not a deployed estimate. Any state-dependent bond schedule first requires calibrated transition probabilities, ordering and capital data, and an equilibrium model.

6.6 Staking

Delegated proof-of-stake: 42,000 GEN validator minimum, 42 GEN delegation minimum, 7-epoch unbonding, stake-weighted VRF selection. Idleness is slashed (1%) and repeat offenders are banned; deterministic violations are slashed via the tribunal. Epoch-level distribution — inflation schedule and the staker/validator/developer/DAO splits — is specified in the contracts and detailed in the Tokenomics companion document.

7 Security: characterized failure modes and countermeasures

We state these plainly because the analysis exists; a security story that hides its assumptions is worth less than one that exposes them. The protocol uses stake-weighted VRF selection, leader-first commit–reveal, and a tribunal for provable violations. Part 1 does not certify that sampler by substituting a weighted mean into a uniform-panel bound. Part 2 covers one fixed-eligible-set successive draw around its panel-size-dependent inclusion mean, but deployment security still requires the history-dependent joint inclusion law, reuse and replacement history, active-set snapshot, seed integrity, ballot dependence, task clarity, and attack objective. Accordingly, adversarial tolerance is a curve or service-level profile rather than one universal percentage.

Majority aggregation amplifies the population, in both directions [P1]. Under the declared i.i.d. model, larger panels reproduce the declared evaluator-population decision increasingly reliably whenever its mean is separated from the threshold. That target need not match external truth: vote-only data do not identify semantic correctness without independent labels or assumptions. *Countermeasure:* treat population design, independently adjudicated workloads, audits, and ex-post channels as separate controls rather than relabeling coordination as correctness.

More validators cannot fix a biased population [P1]. Per-task convergence in the i.i.d. binary benchmark depends on the task-specific mean. Increasing committee size concentrates around that mean and therefore cannot repair a biased target. Under a chosen mixture population, changing one family's component can move the mean by at most that family's mixture weight; this is a sensitivity identity, not evidence that any named family is independent or accurate. *Countermeasure candidate:* measure and vary the evaluator population at the **routing layer**, then validate the effect on independently labelled workloads. Validator LLM dispatch is a pure, seeded, fingerprintable *policy algebra*[3]: per-validator policies with canonical fingerprints (auditable identity), seeded parameter jitter, and per-call policy selection. These are engineering controls whose benefit must be measured; seeds do not remove shared provider state, and mean-preserving variance calculations do not prove semantic robustness. Model-family concentration and fallback-graph diversity are useful candidate health metrics, preferably reported under several population weightings.

Adversarial content moves the target, not the mechanism [P1]. An adversarial input can change the task-specific verdict function and hence its population mean without corrupting validator identities. If the shift makes the declared population coordinate strongly on the attacker's answer, the votes alone cannot distinguish that shared error from another clear population convention. Candidate controls include independently labelled adversarial workloads, input normalization, policy randomization, provider diversity, and developer-side input hygiene. Their transfer rates across families and time must be measured; Part 1 does not prove that any one of them strips an entire attack class.

8 Verification: don't trust — run it

Three artifacts, public and runnable, connect model claims to pinned code baselines (Figure 10). Their revision scope matters: Part 3's reported TLC campaign predates the current clarification that successor rounds reset ballots while validator identities may be reused. That S6 property is pending re-verification rather than retroactively covered.

1. **Canonical state graph** (`opt-dem-spec`): the consensus state machine — 14 stored states, one virtual state, 42 effective transitions — extracted from and anchored to the Solidity implementation, with the anchor check in CI at a pinned revision.
2. **TLA$^+$ specification** (`od-consensus-tla`): model-checks 26+ safety, liveness, tribunal, and commit–reveal properties at the reported baseline; its transition topology is checked against the canonical graph. Current successor-round identity semantics require an updated run.
3. **Fee simulator**: the executable economic specification — 24 invariants over generated path enumeration, with generated vectors enforcing parity against the fee contracts.

[3]Superseding v1's "greyboxing." The shipped greybox was a hand-configured script; the policy algebra makes routing a canonical, hashable data object: same policy, catalog, context, and seed ⇒ bit-identical decision, on-chain and off. That determinism is also what makes watcher re-execution and ex-post audits cheap.

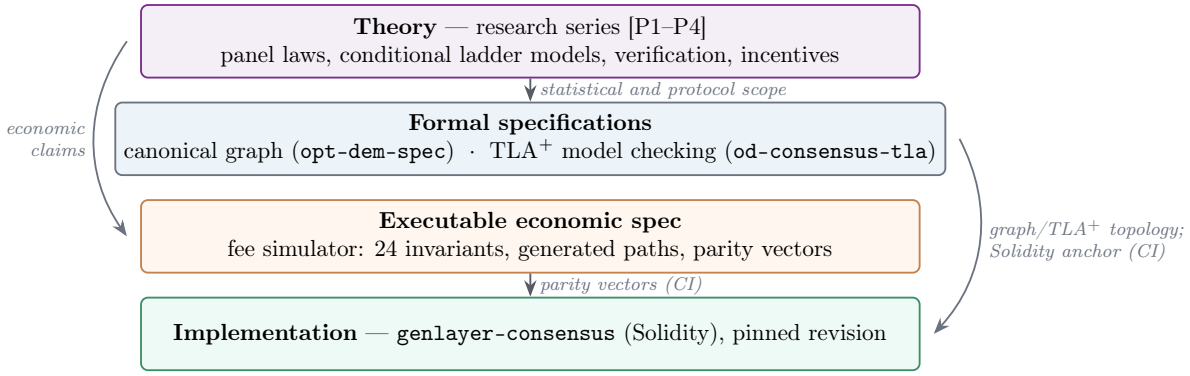

Figure 10: The verification stack. The state-machine and fee artifacts have separate checks against their pinned implementation targets; claims do not silently carry forward when the protocol or specification changes.

It works — because it finds things. The verification track has caught real defects ahead of mainnet: a transaction-ordering race in appeal resolution, three natural invariants falsified by voided appeal rounds, an over-broad cancellation guard in the specification, and dead expiry code in the contracts. All are published. A verification process with findings is the credible kind; the alternative is a process that has never looked hard enough.

9 Use cases

The mechanisms support the application patterns below. A calibrated *contentious* outcome could eventually give contracts a useful third branch, but Section 5 explains why it is not yet a theorem or an implemented protocol primitive.

- **Intelligent Oracles.** Contracts that answer arbitrary questions against the live web — election results for prediction markets, weather triggers for parametric insurance, protocol-attack detection for emergency shutdowns — with panel-population guarantees when their sampling assumptions are met, plus application-specific handling of `Undetermined` outcomes.
- **Trustless world data.** Aggregated, appealable, incrementally-verified databases of real-world facts, built by off-chain proposers and judged on-chain.
- **Autonomous organizations.** DAOs whose constitutions are written in natural language and enforced by the network: proposals evaluated against the constitution and decisions executed by messages; future calibrated abstention could surface low-resolvability calls.
- **Agentic commerce.** Agreements between AI agents (or agents and humans) with qualitative terms — deliverable quality, service levels — adjudicated by the same machinery, with the appeal ladder as the dispute-resolution court.

10 Token

GEN is the network's utility token: gas for execution, staking and delegation, appeal bonds, spam resistance. Supply, distribution, and epoch-level economics are specified in the Tokenomics companion document.

11 Conclusion

The 2024 vision paper described a design; this document separates what is implemented, what is proved under a declared model, what is measured in a diagnostic, and what remains open.

Each boundary is stated once, where its result lives: the decision layer certifies reproduction of a declared population decision (Section 3), the economic layer funds and bounds the work it triggers (Section 6), and the formal artifacts hold at pinned revisions and are re-run after semantic changes (Section 8). Publishing these boundaries is the standard a network holding real value owes its users.

A Protocol parameters (pinned revision)

Parameter	**Value**	**Where**
Nominal size entries	$5, 7, 11, 13, \ldots, 1535, 1537$	Section 4
Appeal feasibility	Round- and available-validator-dependent; minus-two companion rule	Section 4
Capacity example	$N = 1543$ gives a 1535-member successor after eight juries	Section 4
Example successor threshold	$\lfloor 1535/2 \rfloor + 1 = 768$	Section 4
Decision rule	strict-majority category; otherwise no majority	Section 4
Pending expiry	`validUntil` before activation; strict comparison	Section 4
Appeal return	$1.5\times$ bond on success	fee contracts
Bond, validator appeal	$n'W$	fee contracts
Bond, leader appeal	$(\rho + 1)(L + mW)$	fee contracts
Bond, leader-timeout appeal	$(\rho + 1)(L + n_{\text{cur}}W)$	fee contracts
Fee penalty (minority)	symmetric, $\kappa = 1$ (fee only)	fee contracts
Slashing: idleness	1% of stake	`Slash`
Slashing: det. violation	1% validator / 5% leader	`Slash`
Slashing cap per epoch	10%	`Slash`
Validator minimum stake	42,000 GEN	`Staking`
Delegation minimum	42 GEN	`Staking`
Unbonding period	7 epochs	`Staking`
Hoeffding-sufficient top-round margin (10^{-6})	$\gamma \approx 0.0671$; exact homogeneous benchmark ≈ 0.06043	[P2]

Protocol values are pinned to the documented contract baseline; the two statistical margins are model diagnostics rather than contract parameters.

References

[1] GenLayer Research. *Aggregate Disambiguation Systems.* arXiv, pending. [P1]

[2] GenLayer Research. *Adaptive Economic Aggregate Disambiguation.* arXiv, pending. [P2]

[3] GenLayer Research. *Layered Formal Verification of the Optimistic Democracy Protocol.* arXiv, pending. [P3]

[4] GenLayer Research. *Multi-Agent Appeal Games in Aggregate Disambiguation Systems.* arXiv, pending. [P4]

[5] A. Castellana, J. M. Lago, E. Nemše. *GenLayer: The Intelligent Contract Network* (vision paper, v1), 2024. `genlayer.com/whitepaper-v1`. [v1]

[6] Verification artifacts: `opt-dem-spec` (canonical state graph), `od-consensus-tla` (TLA$^+$ specification), fee distribution simulator. Public repositories; pinned revisions cited within.