CONJURA · OPEN PROBLEM

Consensus in the Partial PKI Setting Beyond One-Quarter Resilience

up to the classical one-half bound

Status. Statement: AI-written from Damiano Abram, Marshall Ball, Juan Garay and Aggelos Kiayias, "Permissionless Consensus from a Common Random String", IACR Cryptology ePrint Archive, 2026; abridged version to appear in Proc. CRYPTO 2026, not yet checked by a human. Settled positively for $\gamma < (1/4)(1-3\epsilon)$ with $\epsilon \in (0,1/3)$ by the paper's Partial-PhaseKing protocol (Theorem 5), and negatively at $\gamma = 1/4$, $\epsilon = 0$ for the class of "simple" protocols (Theorem 6, proved there for deterministic protocols); open for $1/4 \le \gamma < 1/2$, with no statement made about non-simple protocols at any $\gamma \ge 1/4$.

Category. *Consensus & Distributed Cryptography*

Conjecture (informal). *Suppose parties with no pre-existing identities bootstrap a public-key directory by spending computational effort: each party registers a key, and each party ends up with its own list of keys it considers valid. The lists are not identical — a key one honest party accepts, another may never have seen — but they overlap heavily, they all contain essentially every honest key, and honest keys make up all but a bounded fraction of the union of the lists. Given such a directory, this paper builds a Byzantine agreement protocol tolerating any adversarial share of the registered keys strictly below one quarter, and shows that one quarter is exactly the ceiling for the natural class of protocols in which honest parties sign everything they send and discard anything mentioning a signer they do not recognize. The classical limit for agreement of this kind is one half, and nothing is known to rule it out here; the authors expect that reaching it needs proofs of work that one party can forward to another, so that a proof accepted by one honest party is accepted by all. The conjecture is that agreement is achievable in this setting for every adversarial share below one half.*

1 The setting

Permissionless consensus without a PKI is often reached by first bootstrapping a *partial* PKI: parties self-certify public keys by investing computational resources, as sketched by Aspnes, Jackson and Krishnamurthy and formalized by Andrychowicz and Dziembowski [1] in an idealized model. The directory that results is inconsistent: honest parties do not share the same view of which keys are valid.

This paper obtains such a directory in the plain CRS model, from a new primitive it calls a multi-verifier signature of work. Abstracting away from how it is built, the resulting object is a (γ, ϵ)-partial PKI: each honest party P_i owns a run-time identity consisting of a verification key and a numerical identifier, and holds a set G_i of run-time identities it recognizes; writing H for the set of honest run-time identities, every G_i contains all but an ϵ fraction of H, and H makes up all but a γ fraction of $\bigcup_i G_i$.

On top of this abstraction the paper builds consensus by the graded-consensus-to-full-consensus route of Feldman and Micali [5], using a variant of the phase king protocol of Berman, Garay and Perry [3] adapted so that the "king" of a phase is whichever recognized identity carries the identifier equal to the phase number. This tolerates $\gamma < \frac{1}{4}(1 - 3\epsilon)$ for $\epsilon \in (0, 1/3)$ (Theorem 5). The paper then proves a matching barrier: no *simple* protocol — one in which honest parties sign every message they send and ignore any message they cannot fully verify, whether because the signer is unrecognized or because the message references a signed message by an unrecognized signer — achieves both Agreement and Validity at $\gamma = 1/4$, $\epsilon = 0$ (Theorem 6, proved in the appendix for deterministic protocols).

The classical resilience limit for consensus of this type is one half [6], and the paper leaves the gap between $1/4$ and $1/2$ open, naming the obstruction: multi-verifier signatures of work are not transferable, so one honest party may accept a proof that another rejects. With transferable proofs one would approach a *ranked* partial PKI, from which $1/2$-resilient consensus follows via an adaptation of Dolev–Strong [4], as in [1]. Note that in the end-to-end permissionless protocol the effective resilience is $\beta\gamma$, where β is the slowdown factor coming from the fine-grained hardness assumption [2]; the conjecture below concerns γ alone, which is where the paper's

positive result and its impossibility both live.

2 Notation and parameters

- $\lambda \in \mathbb{N}$ is the security parameter; n is the number of parties $P_1, \ldots, P_n$ and $n_{\max}$ is a public upper bound on it, with $n \leq n_{\max}$. The value of n is not known to the parties.
- m is a public parameter, polynomial in $n_{\max}$, giving the range $[m] = \{1, \ldots, m\}$ of identifiers.
- $\mathsf{DSig} = (\mathsf{Gen}, \mathsf{Sign}, \mathsf{Verify})$ is an existentially unforgeable signature scheme. A *run-time identity* is a pair (pk, r) with $r \in [m]$; honest party P_i owns (pk_i, r_i) and knows sk_i.
- G_i is the set of run-time identities that P_i recognizes as valid, with $(\mathsf{pk}_i, r_i) \in G_i$. H is the set of run-time identities of honest parties.
- $\gamma \in [0,1)$ bounds the adversary's share of the recognized identities and $\epsilon \in [0,1)$ bounds the disagreement between honest views. The paper's protocol (Theorem 5, with Lemma 3) requires $\epsilon \in (0, 1/3)$ and $\gamma < \frac{1}{4}(1 - 3\epsilon)$; its setup lemma (Lemma 1) takes $\epsilon \in (0,1)$.
- t bounds the number of parties the adversary corrupts, statically; c bounds the computational steps any single party takes per round; θ bounds the messages the adversary sends per round; and β is the slowdown factor of the fine-grained hardness assumption.
- $\mathcal{F}_{\mathsf{diff}}$ is the diffusion channel of the execution model below.
- $b_i \in \{0,1\}$ is P_i's input bit and $b'_i \in \{0,1\}$ its output bit.
- R denotes a round bound and $|S|$ the cardinality of a set S.

Definition 1 (Execution model). Execution proceeds in synchronous rounds. Communication is over an unauthenticated diffusion channel $\mathcal{F}_{\mathsf{diff}}$: a message sent by an honest party in round ρ is delivered to every party at the beginning of round

$\rho + 1$; the adversary chooses the order in which messages arrive at different parties and may send its own messages to an adaptively chosen subset of the parties, so parties cannot tell who sent a message. The adversary is active/Byzantine and statically corrupts up to t parties. Parties and adversary are Interactive Word RAM machines with $O(\log \lambda)$-bit words; every party, honest or corrupted, executes at most c computational steps per round, the adversary sends at most θ messages per round, and a moderately hard function output requiring s steps for an honest party requires at least $\beta \cdot s$ adversarial steps. The adversary controls every run-time identity in $\bigcup_i G_i \setminus H$, including the corresponding secret keys. Each honest party P_i starts with input $b_i \in \{0,1\}$, the public parameters $(1^\lambda, n_{\max}, m)$, its key pair $(\mathsf{pk}_i, \mathsf{sk}_i)$, its identifier r_i, and the set G_i; the protocol logic may not depend on n or on network identities.

Definition 2 ((γ,ϵ)-partial PKI). An assignment of run-time identities and recognized sets $(G_i)_i$ to the parties of an execution is a (γ,ϵ)-*partial PKI* if

1. $|H \cap G_i| \geq (1-\epsilon)\,|H|$ for every honest P_i, and
2. $|H| \geq (1-\gamma)\left|\bigcup_i G_i\right|$.

Definition 3 (Consensus). A protocol in which every honest P_i has input $b_i \in \{0,1\}$ and outputs $b_i' \in \{0,1\}$ achieves *consensus* if

- **Agreement**: all honest parties output the same bit;
- **Validity**: if every honest party has the same input b, then every honest party outputs b;
- **Termination**: every honest party halts with an output.

3 The conjecture

Conjecture 1 (Consensus below one-half resilience). *For every constant γ with $0 \le \gamma < 1/2$ there exist a protocol Π and a polynomial R such that Π achieves consensus in the sense of Definition 3 with resilience γ in the paper's permissionless common-random-string setting: for every $n_{\max} \in \mathbb{N}$, every $n \le n_{\max}$, every input vector $(b_1, \ldots, b_n) \in \{0,1\}^n$ and every adversary subject to the conventions of Definition 1 whose share of the registered run-time identities is at most γ, all honest parties halt within $R(n_{\max}, \lambda)$ rounds and, except with probability negligible in λ, their outputs satisfy Agreement, Validity and Termination.*

Nothing in Conjecture 1 constrains how Π obtains its identities. In particular Π may bootstrap a (γ, ϵ)-partial PKI in the sense of Definition 2, or a stronger setup abstraction — for instance the ranked partial PKI that transferable proofs of work would yield, from which $1/2$-resilient consensus follows via an adaptation of Dolev–Strong [4], and which is the route the paper itself points to.

Conjecture 2 (Fixed-setup strengthening; chosen by the harvester, not stated by the source paper). *As Conjecture 1, but with the setup fixed in advance: for every constant γ with $0 \le \gamma < 1/2$ there exist a protocol Π and a polynomial R such that the conclusion of Conjecture 1 holds in every execution of Π in the model of Definition 1 whose run-time identities and recognized sets form a $(\gamma, 0)$-partial PKI.*

Remark 1. By Theorem 6 of the source paper, any Π witnessing Conjecture 2 with $\gamma \ge 1/4$ must not be "simple": it cannot both sign every message honest parties send and discard every message it cannot fully verify. This rider attaches to the fixed-setup strengthening only. It says nothing about Conjecture 1, whose protocol is free to leave the (γ, ϵ)-partial PKI abstraction behind, in which case Theorem 6 does not apply.

Remark 2. At $\epsilon = 0$ the positive result is a consequence of Theorem 5 rather than an instance of it: a $(\gamma, 0)$-partial PKI satisfies the conditions of Definition 2 for every $\epsilon > 0$, so choosing ϵ small enough that $\gamma < \frac{1}{4}(1 - 3\epsilon)$ yields consensus for every $\gamma < 1/4$ at $\epsilon = 0$.

Note also that $\epsilon = 0$ is not a setting the paper's PartialPKISetup can realize — Theorem 4 with Lemma 1 gives $\epsilon > 0$ — so Conjecture 2 at $\epsilon = 0$ is narrower than the paper's question about its own setting, and a proof at $\epsilon = 0$ does not automatically give one at $\epsilon > 0$.

Remark 3. Both conjectures are stated in terms of γ, the adversary's share of the run-time identities. In the end-to-end permissionless protocol the effective resilience is $\beta\gamma$, with β the slowdown factor of the fine-grained hardness assumption; a solver should say which of the two quantities a claimed $1/2$ bound refers to.

4 Bibliography

[1] Marcin Andrychowicz and Stefan Dziembowski. Pow-based distributed cryptography with no trusted setup. Annual Cryptology Conference, pages 379–399, Springer, 2015.

[2] Marshall Ball, Juan A. Garay, Peter Hall, Aggelos Kiayias, and Giorgos Panagiotakos. Towards permissionless consensus in the standard model via fine-grained complexity. CRYPTO 2024, Part II, volume 14921 of LNCS, pages 113–146, Springer, Cham, 2024.

[3] Piotr Berman, Juan A. Garay, and Kenneth J. Perry. Towards optimal distributed consensus (extended abstract). 30th Annual Symposium on Foundations of Computer Science, pages 410–415, IEEE Computer Society, 1989.

[4] Danny Dolev and H. Raymond Strong. Authenticated algorithms for byzantine agreement. SIAM J. Comput., 12(4):656–666, 1983.

[5] Pesech Feldman and Silvio Micali. An optimal probabilistic protocol for synchronous byzantine agreement. SIAM Journal on Computing, 26(4):873–933, 1997.

[6] Matthias Fitzi. Generalized communication and security models in Byzantine agreement. PhD thesis, ETH Zurich, 2002.