Committee-based Blockchains as Games Between Opportunistic players and Adversaries *

Yackolley Amoussou-Guenou
Université Paris-Saclay, CEA, List, F-91120, Palaiseau, France

Bruno Biais
HEC Paris, 1 Rue de la Libération, 78350, Jouy-en-Josas, France

Maria Potop-Butucaru
Sorbonne Université, CNRS, LIP6, F-75005, Paris, France

Sara Tucci-Piergiovanni
Université Paris-Saclay, CEA, List, F-91120, Palaiseau, France

We offer a game-theoretic study of consensus in a committee-based blockchain. Proposers send blocks to other nodes, who can check block validity and vote for blocks. Blocks with a majority of votes are accepted. Opportunistic players, who maximize expected rewards, interact with adversaries, who seek to disrupt consensus and propose invalid blocks. Free riding implies the prescribed protocol is not an equilibrium and there exist equilibria without consensus, even when adversaries are few. There also exists an equilibrium achieving consensus, in which valid blocks are committed. We propose a modification of the prescribed protocol implementing this equilibrium.

Key words: blockchains, committee, consensus, coordination, free-riding, voting, perfect bayesian equilibrium

1. Introduction

Distributed ledgers share information within a network of participants. Blockchains are distributed ledgers recording information in an evolving list of ordered blocks, each consisting of one or more transactions. This technology could be very useful in many areas, such as, e.g., cadastres, traceability or payments.

For the blockchain technology to fulfill its promise, however, consensus on the distributed ledger must be reached. This can be difficult, especially in public blockchains, where there is no central authority and a large number of anonymous participants. A major challenge, therefore, is to design a blockchain protocol achieving consensus.

* Thanks for insightful comments to participants at The Economics of Distributed Ledger Technology Conference at Cambridge University, The Crypto and Blockchain Economic Research Forum webinar, The Central Bank Research Association conference, The Carnegie Mellon University Distinguished Blockchain Seminar, The International Conference on Autonomous Agents and Multi-agents Systems, and also to Denis Gromb, Cyril Monnet, Tristan Tomala, and Luigi Zingales.

Protocols are sequences of instructions which the participants are to follow. The protocols are designed so that, if participants follow them, consensus is achieved. There is a risk, however, that participants would not follow the protocol. First, to the extent that participants intervene via machines, or processes, there is the risk that the latter could break down. Second, the participants behind the machines have their own agenda. Computer scientists have thus underscored the possibility that participants could be "Byzantines" or "adversaries", aiming at disrupting the blockchain. This could reflect geostrategic motivations, e.g., one country trying to disrupt payment systems in another country. Adversaries could also have economic motivations. For example, the sponsors of one blockchain could try to disrupt a competing blockchain. Even if participants are not adversaries and are not interested in disrupting the blockchain, they might choose to deviate from the prescribed protocol if it is profitable for them to do so. Thus, in addition to adversaries, we consider self interested participants maximizing their expected net profits, whom we refer to as "opportunistic".

Consensus is achieved when the outcome of the interaction between the participants satisfies the three following properties: *Termination:* every non-adversary participant commits on a value B (i.e., decides which block to append to the chain); *Agreement:* if there is a non-adversary participant that commits a value B, then all the non-adversary participants commit B; *Validity:* a committed value by any non-adversary participant is valid, it satisfies an application-defined predicate locally verifiable by any participant.

Different protocols are employed in various blockchains. In the proof-of-work protocol (see Nakamoto (2008)), participants are miners, and one randomly drawn miner gets to append a new block to the chain. The major problem with proof-of-work is that mining consumes a lot of energy (see Stoll et al. (2019), de Vries (2020)). Committee-based protocols, in which several participants exchange messages on blocks, e.g., Algorand, Casper FFG the potential next protocol for Ethereum, Dfinity, DiemBFT for the Facebook's cryptocurrency project, HoneyBadger, HotStuff, Tenderbake or Tendermint,[1] consume much less energy and are therefore more sustainable (see Decker et al. (2016), Eyal et al. (2016), Kokoris-Kogias et al. (2016), Miller et al. (2016), Gilad et al. (2017), Amoussou-Guenou et al. (2018), Hanke et al. (2018), Yin et al. (2019), Baudet et al. (2020), Buterin et al. (2020), Aştefanoaei et al. (2021)). Another potential advantage of committee-based blockchains is that, instead of being made by a single participant, the decision to commit a new block results from the interaction between several participants. This could reduce the exposure of the blockchain to the risk of being disrupted by a single adversary. But can we expect committee-based blockchains to achieve consensus when participants can be opportunistic or adversaries? This is the issue we examine in the present paper.

[1] On which the Cosmos blockchain is based.

We model interaction between participants in a committee-based blockchain as a game, and study whether consensus can emerge in equilibrium. Because we analyze their interaction within a game, we hereafter refer to the participants as players. All players in our analysis are rational and play best responses, but there are two types of players, with different preferences: *Opportunistic players* maximize expected net rewards, while *adversaries* seek to disrupt blockchain consensus at any cost.

To clarify the analysis, we only consider a simplified version of committee based blockchain protocols. But the simplified protocol we consider includes the key elements on which real world committee-based blockchains hinge. We thus consider a setting in which, at each height of the blockchain, a committee is selected. Committee members interact in a sequence of rounds. At the beginning of each round, one of the committee members is selected to be the proposer and sends a block to the others. When receiving a block, committee members decide whether to check its validity or not, and whether to send a vote for the block or not.[2] When a block receives a qualified majority of votes in a round, it is committed and appended to the blockchain. In that case, the network participants move to the next height in the chain. Otherwise, if the number of votes for the block is below the majority threshold, the block is not appended to the chain. In that case, the protocol moves to the next round, in which a new proposer sends a block to the other committee members. For simplicity, we assume the system is synchronous, messages cannot be lost, and proposers can only propose one block. These assumptions imply that agreement is not a problem in our setup and, in our analysis of consensus, we focus on termination and validity.[3]

Checking validity is costly (in terms of electricity and computer space) and so is sending votes. When a block is committed, the committee members who voted for the block receive a reward. When an invalid block is committed, opportunistic players incur a cost. While the reward for blocks is specified in the protocol, the cost incurred by opportunistic players reflects the reputational loss of the blockchain in which they participate, and of which they hold the native currency (in which they are rewarded). This sequence of moves and gains and losses defines a dynamic game.

Our first contribution is to write down the pseudo-codes of the prescribed protocol and of the strategies followed by opportunistic players and adversaries, as well as the extensive form of the corresponding dynamic game, which clarifies the link between the two. Our second contribution is to analyze the Perfect Bayesian Equilibria of that game. We find that there exist multiple equilibria, in which consensus may fail, and that the prescribed protocol is not an equilibrium (when

[2] Sending a vote message amounts to voting in favour of the block, while not sending any message means that the player is not in favour of the block.

[3] Relaxing these assumptions would raise issues regarding agreement. While interesting, these issues are beyond the scope of the present paper and we leave their analysis for further work.

anticipating that the others will follow the prescribed protocol, a non adversary player prefers to deviate from the prescribed protocol). Our third contribution is to propose a modification of the prescribed protocol, such that following the modified prescribed protocol is an equilibrium.

The equilibria without consensus arise because of coordination and free-riding problems. Coordination problems reflect that opportunistic players prefer not to vote, when they anticipate the others won't vote. Because of coordination problems, there exists an equilibrium without termination. Free riding problems reflect that opportunistic players prefer not to check validity, when they expect all the others to check validity (since in that case no single player is pivotal).[4] Because of free riding, there exists an equilibrium without validity. It is also because of free riding that the prescribed protocol (in which all players are instructed to check validity and vote for valid blocks only) is not an equilibrium. There also exists an equilibrium with consensus, however. In that equilibrium some opportunistic players (with index below a threshold) check block validity (thus following the prescribed protocol), while the others (with index above the threshold) don't check validity (thus deviating from the prescribed protocol). The players who are to check validity are pivotal, which prevents free riding. We suggest a modification of the prescribed protocol that implements this equilibrium.[5]

The next section discusses the literature to which our paper is related. Section 3 presents the system model and protocol. Section 4 presents the game and Section 5 its equilibria. Section 6 briefly concludes.

2. Literature

2.1. Computer science

Our analysis is related to the literature on Byzantine consensus-based blockchains. Castro and Liskov (1999) analyze consensus in a Byzantine Fault Tolerant (BFT) protocol. In order to use a BFT protocol in an open setting such as blockchains, recent research endeavoured to find secure mechanisms to select committees of fixed size over time (e.g. Gilad et al. (2017), David et al. (2018)) and to propose incentives to promote participation (e.g. Abraham et al. (2016)). While, this literature allows messages to be delayed and Byzantine processes to equivocate (see Lamport et al. (1982)), we assume away these difficulties to simplify the problem. On the other hand, while the literature on BFT protocols often focuses on correct and Byzantine processes, we in contrast consider opportunistic players and adversaries, which both are rational, i.e., maximize well specified objectives given rational expectations.

[4] There is also a coordination problem with regard to validity checks: If I anticipate the other opportunistic players won't check validity, then I prefer to also not check validity. My own validity check could only alter my own vote, which is not pivotal.

[5] This modified prescribed protocol is "incentive compatible", in the sense that players prefer to follow it if they expect the others to also follow it.

Our analysis, therefore, is in line with the seminal contributions of Abraham et al. (2006), Afek et al. (2014), Aiyer et al. (2005), Groce et al. (2012), Halpern and Vilaça (2016), which took a game theoretic approach to the study of distributed computing and consensus. Groce et al. (2012) consider an environment with rational and honest participants and provide protocols that tolerate rational adversaries, but they take a cooperative game approach, while we take a non-cooperative game approach. Also in a non-cooperative game framework, Halpern and Vilaça (2016), prove that in a fully rational setting, if participants can fail by crashing, there is no *ex post* Nash equilibrium solving the *fair* consensus problem (where fair means that the input of every agent is committed with equal probability), even with only one crash. Afek et al. (2014) proposes protocols solving consensus and renaming under the assumption that participants may be rational and offers a game theoretic analysis. While we consider committee-based protocols, Kiayias et al. (2016), Kroll et al. (2013), and Biais et al. (2019), study Nash equilibria in proof-of-work blockchain protocols.

An important ingredient in models of rational choices is the specification of players' utilities. Lysyanskaya and Triandopoulos (2006) propose an incentive compatible protocol robust to a coalition of up to f faulty players, when the players' utilities reflect whether a decision is reached or not, as well as the value of the decision. Abraham et al. (2006) propose an incentive-compatible protocol for secret sharing with rational participants when some utilities can be unknown. Halpern and Vilaça (2020) show that, when participants can fail by crashing, there is a Nash equilibrium achieving fair consensus when an agent's utility depends only on the consensus value achieved and not on the number of messages the agent sends. In contrast with these analyses, in our model, utilities are known and reflect not only whether and which blocks are appended to the chain, but also the costs of the actions (checking, sending) taken by the players. These costs make the analysis more complex, but also give rise to new economic effects, such as free riding.

Manshaei et al. (2018) also offers a non-cooperative game-theoretic analysis of free-riding in committees. In the protocol they consider, multiple committees run in parallel to validate non-intersecting sets of transactions (shards). They show that rational agents can free-ride when rewards are equally shared. Differences between our analysis and Manshaei et al. (2018) are that i) we do not consider shards, but ii) we consider interaction between adversaries and opportunistic players in a dynamic game, and iii) due to the presence of adversaries some blocks can be invalid.

A key assumption in our analysis is that following the prescribed protocol is costly for players. This is similar to the assumption in Kiayias and Stouka (2020). There are two major differences between their model and ours: First, Kiayias and Stouka (2020) analyze the interaction between an adversary, who might deviate from the prescribed protocol, and honest processes, who are supposed to follow the protocol. In contrast, we analyze the interaction between an adversary and opportunistic players, who follow the prescribed protocol only if it is individually optimal for them.

In this context, we show that following the prescribed protocol is not an equilibrium strategy for opportunistic players. Second, Kiayias and Stouka (2020) apply their methodology to study Bitcoin and Fruitchain. This differs from our focus on committee-based blockchains.

We also share the assumption that following the protocol is costly with Fooladgar et al. (2020). Taking a game theoretical approach, they study the Algorand protocol as a static game, and under the assumption that all players are opportunistic agents maximizing expected rewards net of costs. They note that, in the Algorand protocol, all players equally share the reward obtained when a block is appended to the chain. In this context, they show that following the prescribed protocol is not an equilibrium: if all the other players follow the prescribed protocol, I am better off defecting, since I will receive the same reward, without incurring the cost. They show that a modified version of the protocol, in which players receive the reward only if they follow the protocol, is an equilibrium. Our analysis differs from theirs for the following two main reasons: First, while their model only features opportunistic agents, ours features opportunistic agents and adversaries. It is the presence of adversaries which creates the risk that proposed blocks are invalid. Second, in their model it is observable whether players follow the prescribed protocol or defect. Exploiting this feature of their setting, Fooladgar et al. (2020) put forward a modification of the protocol in which players are rewarded only if they follow the protocol. In contrast, in our model, it is not observable whether players check block validity or not. Therefore, it is impossible to directly make rewards contingent on following the prescribed protocol. This is why, in our setting, pivotality is key to incentives. In contrast, in the equilibrium of the modified protocol of Fooladgar et al. (2020), all players follow the prescribed protocol, while none is pivotal.

Apart from the consensus protocol and the financial aspects of blockchains, in the computer science literature, economics tools have been used to study different technical aspect of blockchains such as cross-blockchains mechanisms (see Belotti et al. (2020)), transactions fees and the design of block rewards (see Basu et al. (2019), Chen et al. (2019)), robustness properties of blockchains (see Zappalà et al. (2020)), and the choice of pools of users (see Eyal (2015), Belotti et al. (2018)). See Liu et al. (2019) for a survey. These analyses differ to ours since we study committee-based blockchains on the consensus level, while they study other aspects of blockchains.

2.2. Economics

Our analysis is also related to the political economy literature on voting. Starting with the seminal contributions of Downs (1957) and Riker and Ordeshook (1968) a stream of papers surveyed in Feddersen (2004) note that, when their probability of being pivotal is infinitesimal, citizens should not vote, as long as the cost of doing so is bounded away from zero. The probability of being pivotal is low when the population of voters is large, and in that case rational choice models predict

citizens should not vote. For example, Palfrey and Rosenthal (1985) show, in a model in which the cost of voting varies across agents, that, as the size of the electorate gets larger, the fraction of agents who vote decreases, and in the limit goes to 0. Our analysis differs from that literature in several respects:

First, we consider a relatively small group of voters (the committee members), so the force that made voters non pivotal in the political economy literature is not at play in our model.[6] In a sense, our point of view is the opposite of that literature:

- The political economy literature considers a large number of voters and underscores that observed high voter participation is surprising: with costly voting, basic rational choice implies citizens should not vote.
- We consider a relatively small number of voters and warn that, in spite of that relatively small number, committee members can be non pivotal, which induces free-riding.

Second, the majority rule we consider is not in terms of fraction of the voters (as in the political economy settings discussed above), but in terms of the number of voters. This restores the possibility that the impact of one vote is zero, inspite of the number of potential voters being relatively small.[7]

Third, while for plausibility the political economy literature assumes citizens cannot be paid for voting, in our setting, when a block is committed, committee members are rewarded conditional on having voted.[8]

Fourth, the classical political economy analyses consider voters that are initially endowed with information, see, e.g., Austen-Smith and Banks (1996), and Feddersen and Pesendorfer (1996, 1997, 1998, 1999). In contrast, Persico (2004), Gershkov and Szentes (2009), and Smorodinsky and Tennenholtz (2006) consider agents who can acquire information, at a cost, before voting. In their analyses, as in ours, key issues are free riding and pivotality. Our analysis differs from theirs, however, because we assume some participants are adversary, while they consider all voters have the same preferences. The presence of adversaries rules out requiring unanimity, which differs from Persico (2004). Also, our focus on distributed consensus rules out central authorities. This differs

[6] Riboni and Ruge-Murcia (2010) also analyze voting within a relative small size committee (the monetary policy committee of a central bank.) In their analysis, however, there is no cost of voting, and committee members are assumed to vote as if they were pivotal.

[7] For example, in Ledyard (1984), when nobody else is voting, a single vote is pivotal. In contrast, in our approach, when nobody else is voting, a single vote has no impact, because it is not enough to raise the number of votes to the majority threshold.

[8] Morton (1987, 1991) study "group based models" in which political group leaders can spend resources to incentivize the members of their group to vote. As noted by Feddersen (2004), "Turnout occurs in group-based models with costly voting for the same basic reason that it occurs in costly voting games with a small number of voters." This contrasts with the risk of lack of participation we undescore.

from Gershkov and Szentes (2009) and Smorodinsky and Tennenholtz (2006), where a central planner designs an optimal mechanism and makes decisions based on the reports elicited from participants.

Our analysis is also related to analyses of free riding in takeovers. Bradley (1980) and Grossman and Hart (1980) consider take over bids when there is a continuum of atomistic shareholders. They show that, if the value of a share after a successful takeover exceeds the offer price, then stockholders (who are non pivotal since they are atomistic) will free ride and fail to tender. Bagnoli and Lipman (1988), in contrast, study the case of a finite number of stockholders. In that case, some stockholders can be pivotal, which eliminates their incentive to free ride. While we also underscore free riding and pivotality, the game we consider is very different from theirs. In their analysis, stockholders have only one decision to make: whether to tender their shares or not, and this decision is observable. In contrast, in our analysis, players have three decisions to make: i) whether to propose a block and which, ii) whether to vote or not, and iii) whether to check validity or not. Since checking validity is unobservable, there is a form of moral hazard, making free riding even more of a concern.

3. System model and protocol

Blockchain: A blockchain is a growing sequence of blocks, to which new blocks can be appended. Once a block is in the blockchain, it cannot be modified nor removed. The block at position $h \geq 0$ in the blockchain is said to be at height h, and the first block at height 0 is the initialization block.

System model: For each given height h, we consider a system composed of a finite and ordered set Π, called *committee*, of n synchronous sequential processes or players. Hereafter, the words "player" and "process" are taken to have the same meaning. $\Pi = \{p_1, \ldots, p_n\}$ and process p_i is said to have index i. We assume each player is aware of its index. In the following, we refer to process/player p_i by its index, say process/player i.

Processes behaviour: The prescribed protocol defines the sequence of actions to be taken by processes. A correct process is a process that follows the prescribed protocol. In this paper, we consider a variant of the BAR model Aiyer et al. (2005) where all processes are rational, i.e., take actions that are optimal for them, given their objective and their rational expectations. These processes follow the prescribed protocol if and only if doing so is optimal for them.

The rational processes in our analysis are either *opportunistic* or *adversary*. *Opportunistic processes* are self-interested and seek to maximize their expected utility, which is equal to the expectation of the rewards they obtain net of the costs they incur. In line with Aiyer et al. (2005), the objective of *adversary processes* is to prevent the protocol from achieving its goal, no matter the cost. We denote by f the number of adversary processes in the network.

Rounds, phases and steps: The n processes interact during at most n rounds. Each round is divided in two sequential phases: the PROPOSE phase and the VOTE phase. These two phases encapsulate the main ideas of consensus protocol for committee-based blockchains.[9]

Each phase is divided into three sequential steps: the send step, the delivery step and the compute step. We assume that the send step is atomically executed at the beginning of the phase and the compute step is atomically executed at the end of the phase.

The phase has a fixed duration that allows collecting all the messages sent by the processes at the beginning of the phase during the delivery step. At the end of a phase, a process exits from the current phase and starts the next one.

The processes communicate by sending and receiving messages through a reliable broadcast primitive.[10] Messages are created with a digital signature, and we assume digital signatures cannot be forged. When process i delivers a message, it knows the process j that created the message. We assume messages cannot be lost.

Algorithm 1 Prescribed Protocol for a given height h at any process i

```
1:  Initialization:
2:    vote := nil
3:    t := 0                                                    /* Current round number */
4:    committedValue := nil

5:  Phase PROPOSE(t):
6:    Send step:
7:      if i == isProposer(t, h) then
8:        proposal ← createValidValue(h)  /* The proposer of the round generates a block, i.e. the value to be proposed */
9:        broadcast ⟨PROPOSE, h, t, proposal⟩
10:   Delivery step:
11:     delivery ⟨PROPOSE, h, t, v⟩ from proposer(t)            /* The process collects the proposal */
12:   Compute step:
13:     if isValid(v) then
14:       vote ← v                          /* If the delivered proposal is valid, then the process sets a vote for it */

15: Phase VOTE(t):
16:   Send step:
17:     if vote ≠ nil then
18:       broadcast ⟨VOTE_i, h, t, vote⟩    /* If the proposal is valid, the process sends the vote for it to all the validators */
19:   Delivery step:
20:     delivery ⟨VOTE, h, t, v⟩            /* The process collects all the votes for the current height and round */
21:   Compute step:
22:     if |⟨VOTE, h, t, v⟩| ≥ ν ∧ committedValue = nil ∧ vote ≠ nil ∧ vote = v then
23:       committedValue ← v; exit          /* The valid value is committed if the threshold is reached */
24:     else
25:       vote ← nil
26:       t ← t + 1
```

Prescribed protocol: Algorithm 1 presents the pseudo-code for a correct process, i.e., a process who follows the prescribed protocol. For each round $t \in \{1, ...n\}$ the committee member with index

[9] Chan and Shi (2020) extended this two phases approach to multiple communication and failure models. They point out the importance and sufficiency of the PROPOSE and VOTE phases in consensus algorithms for blockchains.

[10] A broadcast is reliable if the following conditions hold: i) safety: every message delivered by a process has been previously sent by a source, and ii) liveliness: every process eventually delivers every message sent by a source.

t is designated to be the proposer for the round in a round robin fashion. The `isProposer`(t, h) function returns the index of the proposer for the current round and height (line 7). The function, by taking as parameter the current height, deterministically selects the proposer on the basis of the information contained in the blockchain up to h.

During the PROPOSE phase, the proposer of the round generates a block. In the prescribed protocol, the proposer uses the function `createValidValue`(h), which creates a valid block. Because a valid block must include the identifier of the previous block in the blockchain as well as the index h where the block should be, the height h is passed as parameter (line 8). Once the block is created, a message broadcasting the proposal is sent (line 9). For simplicity, we assume the proposer cannot broadcast more than one block. At line 11 the proposal is received through a delivery function. In the prescribed protocol, each process is correct and checks if the proposal is a valid value (line 13). If so, the process sets its vote to the value (line 14).

During the VOTE phase, any process that set its vote to the current valid proposal sends a message (i.e., a vote) to the other members of the committee (line 18).[11] During the delivery step, sent messages are collected by every process. During the compute step, each process verifies if a quorum of $\nu > 2$ votes for the current proposal has been reached. If the quorum is reached, the process voted for the value and did not already commit for the current height, then the current proposal it committed (lines 23) and the protocol ends. If the quorum is not reached, then a new round starts (line 26).

Consensus properties: Consensus is defined as follows:

Definition 1 *We say that consensus is achieved when the following properties hold:*

- *Termination: every non-adversary process commits on a value (a block);*
- *Agreement: if two non-adversary processes commit respectively on values B and B', then $B = B'$;*
- *Validity: a committed value by any non-adversary process is valid, it satisfies the predefined predicate.*

We use the concept of *external validity* introduced by Cachin et al. (2001). The validity predicate must be known and can be verified by all processes and is defined by the given application. External validity was later adapted by Crain et al. (2017) as a well-suited validity concept for blockchains. Because we assume the system is synchronous, messages cannot be lost, and proposers can only propose one block, agreement is not a problem in our setup. Thus, in our analysis of consensus, we focus on termination and validity. When there are only correct processes, following the prescribed protocol, consensus is reached.

Algorithm 2 Pseudo-code for a given height h modeling the rational process i's behaviour

```
1:  Initialization:
2:    vote := nil
3:    t := 0                                                        /* Current round number */
4:    committedValue := nil
5:    action^propose := nil
6:    action^check := nil
7:    action^send := nil
8:    validValue[] := {⊥, ⊥, ..., ⊥}                                /* validValue[r] ∈ {⊥, 0, 1} */

9:  Phase PROPOSE(t):
10:   Send step:
11:     if i == isProposer(h, t) then
12:       action^propose ← σ_i^propose()   /* σ_i^propose ∈ {−1, 0, 1} sets the action of proposing an invalid block, not proposing a
          block, or proposing a valid block respectively */
13:       if action^propose ≠ 0 then
14:         if action^propose == 1 then
15:           proposal ← createValidValue(h)
16:         else if action^propose == −1 then
17:           proposal ← createInvalidValue()
18:         broadcast ⟨PROPOSE, h, t, proposal⟩
19:   Delivery step:
20:     delivery ⟨PROPOSE, h, t, v⟩ from proposer(h, t)             /* If there is no block proposed, v is set to ⊥ */
21:   Compute step:
22:     action^check ← σ_i^check()   /* σ_i^check ∈ {0, 1} sets the action of checking or not the validity of the proposal. When v = ⊥,
        σ_i^check = 0 */
23:     if action^check == 1 then
24:       validValue[t] ← isValid(v)                                /* The execution of isValid(v) has a cost c_check */
25:     action^send ← σ_i^send(validValue[t]) /* σ_i^send ∈ {0, 1} sets the action of sending the vote or not. When v = ⊥, σ_i^send = 0 */
26:     if action^send == 1 then
27:       vote ← v                          /* The process decides to send the vote, the proposal might be invalid */

28: Phase VOTE(t):
29:   Send step:
30:     if vote ≠ nil then
31:       broadcast ⟨VOTE_i, h, t, vote⟩                            /* The execution of the broadcast has a cost c_send */
32:   Delivery step:
33:     delivery ⟨VOTE, h, t, v⟩                                    /* The process collects all the votes for the current height and round */
34:   Compute step:
35:     if |⟨VOTE, h, t, v⟩| ≥ ν ∧ committedValue = nil ∧ vote ≠ nil ∧ vote = v then
36:       committedValue = v; exit
37:     else
38:       vote ← nil
39:       t ← t + 1
```

Pseudo-code for opportunistic processes and adversaries: While correct processes just follow the prescribed protocol, opportunistic and adversary processes choose actions to maximize their objective. Processes choices are represented in the pseudo-code (Algorithm 2) by dedicated variables, namely, $action^{propose}$, $action^{check}$, and $action^{send}$, defined at lines $5-7$. Each action, initialized to *nil*, can take values from the set $\{0,1\}$, at the exception of $action^{propose}$ which takes its values in $\{-1,0,1\}$. Those values are set by calling the functions $\sigma_i^{\text{propose}}$, σ_i^{check}, and σ_i^{send}, respectively, returning the strategy for process i.

Strategy $\sigma_i^{\text{propose}}$ determines if the proposer i chooses to produce a valid proposal, in which case $action^{propose}$ takes the value 1, or an invalid one, in which case $action^{propose}$ takes the value

[11] To simplify, we assume the proposer does not need to send a vote for the block it broadcasted. Broadcasting a block includes voting for it.

-1, or not proposing a block, in which case $action^{propose}$ takes the value 0 (lines 12-17). When $|action^{propose}| = 1$, the proposal is sent in broadcast (line 18) and when $action^{propose} = 0$, there is no proposal, and no block is sent.[12] Notice that the proposer can send at most one block.

Strategy σ_i^{check} determines if the receiving process chooses to check the validity of the proposal, in which case $action^{check}$ takes the value 1, or not, in which case $action^{check}$ takes the value 0. If the process chooses to check the validity (line 23), it then updates its knowledge about the validity of the proposal. Otherwise, the process does not observe if the proposal is valid or not ($validValue[t]$ remains set to $\perp$).

Strategy σ_i^{send} determines if the receiving process chooses to send a vote, in which case $action^{send}$ takes the value 1, or not, in which case $action^{send}$ takes the value 0 (line 25-31). Note that the strategy σ_i^{send} depends on the knowledge the process has about the validity of the proposal. The strategy determines if the process chooses to send its vote for the proposal or not.

4. Game

In this section, we offer a game theoretic formulation of the protocol presented in the previous section. We consider a game between n processes, indexed by i in $\{1, ...n\}$, hereafter referred to as the players.

Number of adversaries: The number of adversaries is denoted by f. $\nu > 2$ denotes the majority threshold needed for a block to be committed. While ν is a parameter of the protocol, commonly known by all participants, f is the realization of the random variable denoted by $\tilde{f}$. $\tilde{f}$ takes its values in $\{1, ...\bar{f}\}$, with $\bar{f} < \min[\nu, n - \nu]$. $\bar{f} < \nu$ means that adversaries never have the majority, while $\nu < n - \bar{f}$ means opportunistic players always have the majority. Together these conditions imply that $\bar{f} < \frac{n}{2}$. One could expect these assumptions to imply that consensus should be achieved. This is not the case, however. As shown below, termination or validity may fail to hold in equilibrium.

Players' types: At the beginning of the game, nature draws the types of all the players, and correspondingly the number of adversaries. Each player i observes its own type, θ_i, which can be adversary ($\theta_i = \theta^a$) or opportunistic ($\theta_i = \theta^s$). The probability that a player is an adversary is independent from its index, and thus, a priori, the same for all i between 1 and n.

Action space: At round t the player with index t is selected to generate a block and send it to the other players. We refer to this player as "the proposer." The cost of generating a block is denoted by C, while the cost of sending the block is denoted by c. The proposer can choose whether to generate a valid block or an invalid one, or no block at all. If the proposer does not generate nor send any block, it does not incur costs C and c.

[12] When no block is proposed, there is no block to check, and no proposal to vote for. In that case, participants go to the next round.

Players who receive a proposed block must decide whether to check its validity. Since checking validity involves similar operations as generating a block, we assume the cost of checking block validity is also equal to C. Then the player decides whether to send a vote for the block or not. Again, for simplicity, we assume the cost of sending a message is equal to the cost of sending a block, c. If no block is proposed, players don't have the opportunity to check validity or vote. In that case, they simply terminate the current round and go to the next one.

Our framework is in line with Kiayias and Stouka (2020) and Fooladgar et al. (2020) who also assume that following the prescribed protocol involves costs.

Information sets: We assume opportunistic players only observe their own type.[13] Adversaries, in contrast are assumed to know the types of all players. This assumption is in line with that, often made in computer science, that adversaries (or Byzantine processes) are very powerful and can collude. Also consistent with that assumption, we assume adversaries can perfectly coordinate their actions. The goal is to test whether the blockchain is robust to adversary attacks. If a blockchain resists the attack of very powerful adversaries, it means it is very robust.

The information set of player i, at the beginning each round t, which we denote by η_i^t, includes

- the observation of the round number t,
- the player's own type θ_i, and also when i is an adversary the types of all the other players,
- when $t > 1$, the observation of what happened in previous rounds, namely (i) whether i decided to check validity, and in that case the knowledge of whether the block was valid or not, (ii) how many messages were sent at each round, and (iii) whether a block was proposed or not.

Then, in the Compute step of the PROPOSE phase, when receiving a proposed block, each player decides whether to check its validity or not. Denote by b_t the block proposed at round t. When the player decides not to check validity, $\texttt{isValid}(b_t)$ is the null information set, while if the player decides to check, $\texttt{isValid}(b_t)$ is equal to 1 if the block is valid and 0 otherwise. So, the player information set becomes $H_i^t = \eta_i^t \cup \texttt{isValid}(b_t)$.

Strategies: At each round $t \geq 1$, the strategy of player i is a mapping from its information set into its actions. If $i = t$, the player is selected to propose the block and its choice is given by $\sigma_i^{\text{propose}}(\eta_i^t)$. Then, in the Compute step of the PROPOSE phase, if $i \neq t$, the player's strategy is given by $\sigma_i^{\text{check}}(\eta_i^t)$. Finally, in the Send step of the VOTE phase, player $i \neq t$ must decide whether to send a vote or not, and that decision is given by $\sigma_i^{\text{send}}(H_i^t)$.

Rewards and Costs for Opportunistic Players: We assume that, when a block is produced, only the players who sent a message receive a reward, denoted by R. This is in line with practice, e.g., in the Tendermint protocol, see Amoussou-Guenou et al. (2018). In addition, we assume that when

[13] If adversaries were detectable, they could be excluded, and therefore could not harm the system.

an invalid block is produced, all opportunistic players incur cost κ. Including an invalid block in the chain reduces the reputation and credibility of the chain, and therefore its business prospects and valuation. κ is borne by all the non adversary players, whether they voted for the block or not. It is reasonable to assume that the reputational cost is κ large and greater than the reward R. On the other hand, the reward R must be set above the cost C of checking validity or generating a block. Otherwise, players would not be willing to participate in the blockchain. C, however, is larger than the cost c of sending a message. Checking validity requires significant internal computations and inspecting the whole chain of blocks, while sending a message is less costly for the player, since it involves less and simpler computations. Thus, overall we assume

$$\kappa > R > C > c > 0.$$

Rewards and Costs for Adversaries. The adversaries have lexicographic preferences over the outcome of the game, in order, they prefer:

1. Outcomes that do satisfy Termination, but not Validity;
2. Outcomes that do satisfy Validity, but not Termination;
3. Outcomes that do satisfy Termination and Validity;

Adversaries are assumed to care only about the outcome of the protocol and neglect the costs of their own actions.

Objective of opportunistic players: Let T be the endogenous round at which the game stops. If a block is produced at round $t \leq n$, then $T = t$. Otherwise, if no block is produced, $T = n + 1$. In the latter case, the *termination* property is not satisfied.

At the beginning of round $t \geq 1$, the continuation payoff of the opportunistic player with information set η_i^t is[14]

$$W_{i,t}(\eta_i^t) = E\left[\begin{array}{c}(R * \mathbf{1}_{(\sigma_i^{\text{send}}(H_i^T)=1)} * \mathbf{1}_{(\text{block produced at } T)} - \kappa \mathbf{1}_{(\text{invalid block produced})}) \\ -\sum_{s=t}^{T}\left(\mathbf{1}_{(i=s)}(C+c)\mathbf{1}_{(|\sigma_i^{\text{propose}}(h_i^t)|=1)} + \mathbf{1}_{(i\neq s)}\left(C\mathbf{1}_{(\sigma_i^{\text{check}}(h_i^s)=1)} + c\mathbf{1}_{(\sigma_i^{\text{send}}(H_i^s)=1)}\right)\right)\end{array} | \eta_i^t\right].$$

$\mathbf{1}_{(.)}$ denotes the indicator function. The first line reflects the reward received by committee members when a new block is committed, minus the penalty incurred when that block is invalid. The second line reflects the flow of costs incurred by player i at rounds $s = t, ...T$, when this player is a proposer $(i = s)$ or a voter $(i \neq s)$.

[14] This is comparable with the specification of players' utility in Kiayias and Stouka (2020) as absolute rewards minus absolute costs.

Equilibrium concept: Since we consider a dynamic game with asymmetric information, the relevant equilibrium concept is Perfect Bayesian Equilibrium (Fudenberg and Tirole (1991)), intuitively defined as follows:

Definition 2 *A Perfect Bayesian equilibrium is an n-tuple of strategies, such that all players 1) choose strategies that are optimal given their preferences and beliefs, 2) rationally anticipate the strategies of the others, and 3) draw rational inferences from what they observe, using their expectations about the strategies of the others and Bayes law whenever it applies.*

A Perfect Bayesian Equilibrium (PBE) is a Nash equilibrium (Nash (1951)), so players best-respond to one another. It imposes additional restrictions, to take into account the fact that the game is dynamic and that players can have private information and must draw rational inferences. Rationality implies that each player's beliefs are consistent with Bayes law, when computing probabilities conditional on events that have strictly positive probability on the equilibrium path. Perfection implies that at each node starting a subgame the players' strategies form a Nash equilibrium of that subgame. In this context, to show that a strategy is optimal it is sufficient to show that it dominates any one-shot deviation Blackwell (1965).

5. Equilibria

5.1. Equilibrium without termination

Consider an opportunistic player who receives a block and anticipates the others won't vote. This player anticipates that, even if it voted, no block would be committed, since the number of votes would still be below the majority threshold. So, when an opportunistic player anticipates the others won't vote, he/she prefers to also abstain from voting: his/her message would not change the outcome of the vote, and would only result in the player incurring cost c. A fortiori, it is suboptimal for this player to check validity, since it would only result in incurring cost C without otherwise changing the outcome of the game.

Now turn to adversaries, who, in contrast with opportunistic players, can coordinate their moves. Should they choose to all vote for invalid blocks? Such a move would result in $f < \nu$ votes in favour of the invalid block. So the block would still not be committed. Consequently, adversaries (weakly) prefer not to send votes, and, a fortiori, not to check validity.

The above remarks imply that, when players expect the others not to vote, their best response is also not to vote. Anticipating this, proposers find it optimal not to send any block. These remarks are summarized in our first proposition:

Proposition 1 *There exists an equilibrium in which no block is ever proposed.*

In the equilibrium of Proposition 1, there is no termination. This failure of consensus reflects a coordination failure among players, who coordinate on a bad equilibrium in which no one votes. Votes, however, are observable. So one could modify the protocol to reimburse c to the players. As shown below, this would not be enough to eliminate coordination problems.

5.2. Equilibrium without validity

Our next proposition shows there can also be issues with block validity checks.

Proposition 2 *If $C \leq \frac{\kappa}{n-1}$, there exists a perfect Bayesian equilibrium in which: i) An opportunistic proposer sends a valid block while an adversary sends an invalid block. ii) When receiving blocks, opportunistic players do not check validity but send a vote. iii) Adversaries vote in favour of invalid blocks only. Thus, when the proposer is opportunistic a valid block is committed, while when the proposer is adversary an invalid block is committed.*

In the equilibrium of Proposition 2 there is termination, at the first round, but validity does not always hold. Adversaries, propose invalid blocks and also vote for invalid blocks, which end up committed. Opportunistic proposers send valid blocks, which end up committed. In that case, they earn $R-C-c$. This is larger than what they would have obtained if they had sent an invalid block, $R-C-c-\kappa$.

Consensus can fail in the equilibrium of Proposition 2 due the behaviour of opportunistic players, who send votes without checking validity, which can result in committing invalid blocks. Why do opportunistic players find it optimal to do so? Why don't they prefer to follow the prescribed behaviour, which is to check validity and vote only if the block is valid? The problem is that, when the other opportunistic players follow the equilibrium strategies of Proposition 2, an opportunistic player deviating to the prescribed behaviour would be unable to alter the outcome of the protocol. Valid blocks would still get $n-f$ votes and be committed, while invalid blocks would get $n-1$ votes and also be committed. Thus, following the prescribed behaviour would give opportunistic players the following expected payoff

$$R-C-\Pr(\text{valid})*c-\Pr(\text{invalid})*\kappa,$$

which is lower than their equilibrium payoff

$$R-c-\Pr(\text{invalid})*\kappa$$

because $C>c$. To complete the proof of Proposition 2, in the appendix we show that, at round t, the opportunistic player with index $i=t$ is better off proposing blocks than not proposing, while the opportunistic players with index $i\neq t$ are better off voting than not voting.

In Proposition 2 consensus fails because each opportunistic player finds it suboptimal to check block quality. Checking validity (and voting for valid blocks only) can be interpreted as producing a public good. In the equilibrium of Proposition 2, opportunistic players free ride on the supply of the public good. The problem is more acute than in Proposition 1: In the case of Proposition 1, one way to solve the coordination problem was to reimburse the cost of sending messages. This does not work for Proposition 2 because validity checks, while costly, are unobservable. In that sense, Proposition 2 reflects moral hazard among the opportunistic players.

Our game theoretic approach, in which all players (or processes) are rational, contrasts with the standard approach in computer science, in which some processes are correct and assumed to follow the prescribed protocol, while other processes are Byzantines. In that standard computer-science approach, a typical result is that, if the number of Byzantines is lower than a threshold, then consensus obtains. In contrast, Proposition 2 states that even if f is very low, as long as it is not zero, consensus may fail. This is because, in addition to attacks by adversaries, we consider a new source of fragility: coordination problems and free riding among rational, but opportunistic, players.

5.3. The prescribed protocol is not an equilibrium

While the above analysis shows that deviations from the prescribed protocol can form an equilibrium, it does not rule out the possibility that following the prescribed protocol would be an equilibrium strategy. We now examine that point. Recall that the prescribed protocol entails i) proposing valid blocks, ii) checking the validity of blocks received and iii) voting for valid blocks only. If all the opportunistic players follow that strategy, then consensus obtains: As long as the proposer is an adversary the block is invalid and rejected, and the first time the proposer is opportunistic, the proposed block is valid and a majority $n-f>n-\bar{f}>\nu$ of committee members vote for that block, which is therefore committed. Unfortunately, if $\nu>\bar{f}+1$, following the prescribed strategy is not an equilibrium because, when a participant anticipates the others to follow that strategy, that participant prefers to shirk on validity checks, to avoid bearing cost C.

To see this, consider opportunistic player i receiving a block and anticipating the others follow the prescribed strategy. If i also follows the prescribed strategy, then if the block is valid it gets at least $n-\bar{f}$ votes and gets committed (since $n-\nu>\bar{f}$), while if it is invalid the block gets at most $\bar{f}$ votes and does not get committed (since $\nu>\bar{f}$). However, the outcome is unchanged if, instead, the opportunistic player deviates from the prescribed strategy and sends a vote without checking block validity: If the block is valid, it still gets at least $n-f>\nu$ votes and gets committed, while if the block is invalid, it gets at most $\bar{f}+1$ votes. In that case, the block is still not committed

if $\nu > \bar{f} + 1$. Therefore, under that condition, as stated in the next proposition, the prescribed protocol is not an equilibrium.[15]

Proposition 3 *If $\nu > \bar{f} + 1$, then it is not an equilibrium that all opportunistic players follow the prescribed protocol at all rounds.*

Note that Proposition 3 obtains whatever the majority threshold ν, as long as it is larger than $\bar{f} + 1$. In particular the result holds even if $\nu = n - \bar{f}$ so that (when the number of adversaries is $\bar{f}$) the votes of all opportunistic players are needed for the valid block to be committed. While $\nu = n - \bar{f}$ can make opportunistic players pivotal in the vote (if one of them fails to vote the valid block is not committed when the number of adversaries is $\bar{f}$), it does not make them pivotal in validity checks. Thus, opportunistic players free ride on validity checks, which prevents the prescribed strategy from being an equilibrium. In contrast, as explained in the next subsection, when $\nu = \bar{f} + 1$ we are in a special, limiting, case of Proposition 5, in which opportunistic players are pivotal and check validity.

5.4. Equilibrium consensus

Our results, so far, are negative in the sense that they present equilibrium situations in which consensus is not achieved and the prescribed strategy is not followed. Is the committee-based protocol we examine doomed to fail? Fortunately, the answer to that question is negative. We hereafter show that there exists an equilibrium in which consensus is achieved.

To have termination and validity, it must be that sufficiently many opportunistic players prefer to check block validity and vote for valid blocks. The problem in Propositions 2 and 3 was that opportunistic players were tempted to free-ride, and let the others bear the cost of checking. To avoid this situation, it must be that (at least some) opportunistic players are pivotal, i.e., if they vote without checking block validity, this can lead to committing invalid blocks. In this section, we present an equilibrium in which the strategies of the opportunistic players imply they are endogenously pivotal.

In this equilibrium, a key ingredient is that some opportunistic players are expected to check validity, while others are not. It is because some opportunistic players don't check validity that the others are pivotal with some probability. As a first step in the equilibrium construction, the next proposition (whose proof is in the appendix) characterizes the expected continuation payoff of opportunistic players in this context.

[15] When $\nu = \bar{f} + 1$ we are in a special, limiting, case of Proposition 5, presented below. $\nu = \bar{f} + 1$ is consistent with $\nu < n - \bar{f}$ iff $\bar{f} < (n-1)/2$.

Proposition 4 *Consider a candidate equilibrium in which i) opportunistic proposers propose valid blocks, ii) some opportunistic voters check block validity and then send a vote if and only if the block is valid, iii) the other opportunistic voters send votes without checking validity, iv) when proposed, valid blocks are committed while invalid blocks are rejected. In this candidate equilibrium, the time-t continuation payoff of the opportunistic voters who are to check block validity is*

$$\pi_{check}(t) = (R-c) - \phi(t)C, \ \forall t \leq \bar{f}+1, \ \textit{with } \phi(\bar{f}+1) = 0 \ \textit{and } \phi(\bar{f}) = 1, \tag{1}$$

while that of the opportunistic voters who are not to check block validity is

$$\pi_{send}(t) = R - \psi(t)c, \ \forall t \leq \bar{f}+1, \ \textit{with } \psi(\bar{f}+1) = 1. \tag{2}$$

The proof of Proposition 4 is in appendix. It gives the definition of ϕ and ψ and offers a detailed analysis of (1) and (2). The intuition for (1) is the following: When a valid block is committed, opportunistic voters, who all send a message, get payoff $R-c$. This is the first term in $\pi_{check}(t)$, reflecting that eventually a valid block will end up committed. The second part of $\pi_{check}(t)$, $\phi(t)C$, is the expected cost of checking the block validity, where $\phi(t)$ is the expected number of times (from round t on) the player expects to check validity before a block is committed. We have $\phi(\bar{f}) = 1$ because, when round $\bar{f}$ is reached, opportunistic players know this is the last time they have to check block validity. c shows up only once in (1) because the players who are to check validity vote only once, when they receive a valid block, and in that round the block is committed.

Similarly, in $\pi_{send}(t)$, $\psi(t)C$ is the expected cost of sending messages, where $\psi(t)$ is the expected number of times the player expects to send messages before a block is committed. We have $\psi(\bar{f}+1) = 1$ because, when that round is reached, opportunistic players who are not to check validity know this is the last time they have to vote.

Denote by $f(t)$ the conditional expectation of $\tilde{f}$, from the point of view of an opportunistic player knowing that all participants with indexes $i \leq t$ are adversaries. Relying on Proposition 4, our next proposition describes more precisely our candidate equilibrium and states the conditions under which it is indeed an equilibrium.

Proposition 5 *If the cost κ of committing an invalid block is large enough in that*

$$\kappa > \max(\alpha(t)C - \beta(t)c, C/\bar{p}), \forall t < \bar{f}, \tag{3}$$

and the reward R for committing a block is large enough in that

$$R \geq c + \frac{n-t+1}{n-f(t)}C, \forall t \leq \bar{f}+1, \tag{4}$$

then there exists a Perfect Bayesian Nash equilibrium achieving consensus, in which the strategy of opportunistic players is the following:

- *Adversaries propose invalid blocks and vote for invalid blocks.*
- *At any round $t \leq \bar{f}$, (i) opportunistic proposers propose valid blocks, (ii) opportunistic voters with index $i \in \{t, \ldots, n - \nu + \bar{f} + 1\}$ check validity and send votes for valid blocks, and (iii) opportunistic voters with index $i \in \{n - \nu + \bar{f} + 2, \ldots, n\}$ do not check validity but send votes.*
- *If round $t = \bar{f} + 1$ is reached, no opportunistic player proposes a block until round n. When round n is reached, the proposer proposes a valid block, and all opportunistic players send a message without checking validity. At this point, the block is valid and committed.*

The proof of Proposition 5 is in the appendix. It gives the definition of $\alpha(t)$, $\beta(t)$ and $\bar{p}$. The intuition of the proposition is the following. In the equilibrium of Proposition 5, invalid blocks (proposed by adversaries) are rejected, while valid blocks (proposed by opportunistic players) are committed. This implies that, if round $t = \bar{f} + 1$ is reached, the players know that the number of adversaries was $\bar{f}$ and that during the $\bar{f}$ previous rounds the proposers were adversaries (to draw this inference, the opportunistic players use their anticipation that all participants play equilibrium strategies). Consequently, at round $\bar{f} + 1$, in equilibrium, the proposer is opportunistic and the proposed block is valid. So, no opportunistic player needs to check the validity of the block but all send a message, which brings them $R - c$. This is larger than their gain from deviating (e.g., by not sending a message or by checking the block.)

Similarly, at round $t \leq \bar{f}$, in equilibrium all $t - 1$ previous proposers were adversaries and there remains $f(t-1) - (t-1)$ adversaries with index strictly larger than $t - 1$. In this context, do the equilibrium strategies of the opportunistic players preclude invalid blocks from being committed? To examine this point, consider the maximum possible number of messages that can be sent if the proposer is an adversary. In equilibrium, opportunistic players with indexes strictly larger than $n - \nu + \bar{f} + 1$ are to send a vote without checking block validity. The worst case scenario (maximizing the number of messages sent when the block is invalid) is that none of these players are adversaries and that $f = \bar{f}$. In that case, in equilibrium, the number of messages sent when the block is invalid is $\bar{f} + (\nu - \bar{f} - 1) = \nu - 1$, so that we narrowly escape validation of the invalid block. If, in that worst case scenario one of the opportunistic players with index below $n - \nu + \bar{f} + 1$ deviated from equilibrium and sent a message without checking the block, this would lead to committing an invalid block. Thus, in that sense, the opportunistic players with index lower than $n - \nu + \bar{f} + 1$ are pivotal. Hence, they check block validity, because, under condition (3), the cost of committing an invalid block is so large that opportunistic players do not want to run the risk of tipping the balance.

One could worry that the equilibrium in Proposition 5 is a bit complex, making it hard for players to discover equilibrium and coordinate on it. This difficulty can be circumvented by modifying the

prescribed protocol. In the prescribed protocol, all players are instructed to check block validity. In the modification of the prescribed protocol we suggest, players are instructed to check block validity if and only if their index is strictly lower than a given threshold. Proposition 5 implies that, when opportunistic processes expect the others to follow the modified prescribed protocol, their best response is to also follow the modified prescribed protocol. Thus, the modified prescribed protocol can be interpreted as a way to make equilibrium consensus a focal point, in the sense of Schelling (1960).

As mentioned above, $\nu = \bar{f} + 1$ is an interesting limiting case of Proposition 5. In that case, in equilibrium, all opportunistic players check block validity (and thus follow the prescribed protocol) up to round $\bar{f}$.[16] This is because, when the majority threshold is at $\bar{f}+1$, all opportunistic players are pivotal if the proposed block is invalid and the number of adversaries is $\bar{f}$. Relative to the general case of Proposition 5, this case is relatively simple, since opportunistic players' strategies don't depend on their index. This leads to another suggestion on how to obtain consensus in the committee-based protocol we consider: set the majority threshold to $\nu = \bar{f} + 1$. This ensures the majority threshold is high enough, so that adversaries don't have the majority when the opportunistic players follow the protocol, but at the same time not too high, so that adversaries would have the majority if opportunistic players deviated from the protocol.

6. Conclusion

In distributed ledgers, such as blockchains, there is no central authority. The advantage is that this eliminates the risk of a malevolent central authority harming network participants. The drawback is that this creates the risk of coordination failures and free riding. We study whether, in spite of this risk, consensus can emerge in committee-based blockchains.

We show that coordination failures can arise concerning vote messages. This is because a minimum number of votes is requested for a decision to be made, so that sending a vote is useless when the others don't vote. In that sense, votes are strategic complements. We also show there can be free riding concerning validity checks. In contrast with votes, validity checks are strategic substitutes: a player prefers not to check validity if he/she expects the others to check. We show that, because of coordination failures and free-riding, there exist equilibria in which Termination or Validity fail. On the other hand, we also show that there exists an equilibrium in which consensus (with Termination and Validity) is achieved. An important issue is how to avoid "bad equilibria," without consensus. We suggest a modification of the prescribed protocol which implements the "good equilibrium,"

[16] After that round the opportunistic players deviate from the prescribed protocol, as they no longer check validity. This is because all proposed blocks are valid and therefore don't need to be checked.

with consensus. To the extent that it facilitates coordination on that equilibrium, the modified protocol can be interpreted as a focal point, see Schelling (1960).

An important avenue of further research is to extend the analysis to settings in which agreement would be an issue because the system would be non synchronous or messages could be lost. Other promising avenues of further research are to rely on mechanism design and relate the blockchain protocol to currency valuation, as in Auer et al. (2021) or Persico (2004).

Appendix: Proofs

Proof of Proposition 2:

Suppose an opportunistic proposer decides not to propose a block. Since no block would be proposed, the protocol implies we would move to the second round. We assume that, after this out of the equilibrium path event, the players would then revert to their equilibrium strategies. Thus, at the second round a block would be proposed and end up committed. With probability $\frac{f}{n-1}$ the proposer would be an adversary. Since $f > 1$, this probability is larger than $\frac{1}{n-1}$. So the expected gain of the deviating opportunistic proposer would be lower than

$$R - c - \frac{\kappa}{n-1}.$$

This is lower than the equilibrium gain, $R - C - c$, under the condition stated in the proposition, which means that the cost of committing invalid blocks is high relative to the cost of generating valid blocks.

Another possible deviation for opportunistic players would be to abstain from voting (and from checking validity). This, however, would however give expected payoff equal to $-\Pr(\text{invalid}) * \kappa$, which is lower than the equilibrium payoff.

The other deviations: check but don't vote or check and vote irrespective of validity are obviously dominated by the equilibrium action.

Q.E.D.

Proof of Proposition 4:

The functions $\phi(t)$ and $\psi(t)$ are defined recursively as follows:

$$\phi(t) = 1 + \frac{f(t) - t + 1}{n - t + 1}\phi(t+1), \forall t < \bar{f}, \text{ and } \phi(\bar{f}+1) = 0, \tag{5}$$

$$\psi(t) = 1 + \frac{f(t) - t + 1}{n - t + 1}\psi(t+1), \forall t < \bar{f}, \text{ and } \psi(\bar{f}+1) = 1. \tag{6}$$

1) In the first part of the proof, we prove that the round t continuation payoff of opportunistic voters that are supposed to check validity is as in (1). To do so we proceed by backward induction.

First, we establish that (1) holds at round $t = \bar{f} + 1$. If round $\bar{f} + 1$ is reached, in equilibrium it means that no valid block has been proposed so far. That is, all $\bar{f}$ previous proposers were adversary and that now there are only opportunistic proposers. So they do not check block validity and send a message, which gives them payoff $R - c$, so that (1) holds at $t = \bar{f} + 1$.

Now turn to round $t < \bar{f}+1$. If round $t < \bar{f}+1$ is reached, the previous $t-1$ proposers were adversaries. There remains $n-(t-1)$ potential proposers. Denoting by $f(t)$ the conditional expectation of $\tilde{f}$ given that all participants with indexes $i \leq t$ were adversaries, the probability that the next proposer is adversary is

$$\frac{f(t)-(t-1)}{n-t+1},$$

and, with the complementary probability,

$$\frac{n-f(t)}{n-t+1}$$

the next proposer is opportunistic.

We now prove that if (1) holds at round $t+1 \leq \bar{f}+1$, i.e.,

$$\pi_{check}(t+1) = R - c - \phi(t+1)C,$$

then (1) holds at round t. Suppose the opportunistic voter follows the equilibrium strategy of checking and sending iff the block is valid. Its expected gain from round t on is

$$-C + \frac{n-f(t)}{n-t+1}(R-c) + \frac{f(t)-(t-1)}{n-t+1}\pi_{\text{check}}(t+1),$$

where the first term is the cost of checking the block at round t, the second term is the probability that the block is valid and committed multiplied by the payoff in that case, and the third term is the probability that the block is invalid and rejected multiplied by the payoff in that case. Substituting the value of $\pi_{check}(t+1)$ from (1) evaluated at round $t+1$, the expected gain writes as

$$-C + \frac{n-f(t)}{n-t+1}(R-c) + \frac{f(t)-(t-1)}{n-t+1}(R-c-\phi(t+1)C).$$

That is

$$R - c - \left(1 + \frac{f(t)-(t-1)}{n-t+1}\phi(t+1)\right)C,$$

which, by (5), is $R-c-\phi(t)C$. Consequently, (1) holds at round t.

2) In the second part of the proof, we prove that the round t continuation payoff of opportunistic players that are not supposed to check validity is as in (2).

Again, we proceed by backward induction, proving that if (2) holds at round $t+1 \leq \bar{f}+1$, i.e., $\pi_{send}(t+1) = R - \psi(t+1)c$, then it holds at round t. Suppose the opportunistic voter follows the equilibrium strategy of not checking blocks' validity and always sending a message. On the equilibrium path its expected gain from round t on is

$$-c + \frac{n-f(t)}{n-t+1}R + \frac{f(t)-t+1}{n-t+1}\pi_{send}(t+1),$$

where the first term is the cost of sending a message at round t, the second term is the probability that the block is valid and committed multiplied by the payoff in that case, and the third term is the probability that the block is invalid and rejected multiplied by the payoff in that case. Substituting the value of $\pi_{send}(t+1)$ from (2) evaluated at $t+1$, the expected gain writes as

$$-c+\frac{n-f(t)}{n-t+1}R+\frac{f(t)-t+1}{n-t+1}(R-\psi(t+1)c).$$

That is

$$R-\left(1+\frac{f(t)-t+1}{n-t+1}\psi(t+1)\right)c,$$

which, by (6), is $R-\psi(t)c$.

Q.E.D.

Proof of Proposition 5:

1) Our first step is to rule out possible deviations for adversaries. If an adversary deviated by sending a valid block, the block would be immediately committed, which would not make the adversary strictly better off. Moreover, since adversaries neglect the costs of checking validity and sending messages, they weakly prefer to check validity and send messages for invalid blocks, although, on the equilibrium path, these actions have no impact on the outcome.

2) The second step is to show that, when it is their turn to propose, opportunistic players prefer to propose a valid block.

First, we note that opportunistic proposers prefer to propose a valid block rather than an invalid one. If the opportunistic player proposes a valid block, it is committed and the proposer is rewarded, while if the opportunistic player proposes an invalid block, it is rejected, and we move to the next round, so the player gets less than its equilibrium gain.

Second, we study whether opportunistic proposers prefer to propose a valid block than to propose no block. To do so, we proceed by backward induction.

Suppose we reach round $\bar{f}+1$, which in equilibrium happens if $f=\bar{f}$ and all the players with the first $\bar{f}$ indexes are adversaries. A this point all the remaining proposers are opportunistic. They all anticipate that, if we reach round n, the proposer at that point will be an opportunistic agent and will propose a valid block, giving them gain: $R-c$. If instead an opportunistic agent proposed a valid block before round n, it would get $R-c-C$, which is lower. Thus, all agents with index stricly lower than n prefer not to propose any block, and the protocol moves to round n, at which a valid block is proposed and committed.

Now, suppose we reach round $\bar{f}$ (which implies $f\geq\bar{f}-1$ and all the previous proposers are adversaries). If the proposer is an adversary, it proposes an invalid block and we move to round $\bar{f}+1$. If instead the proposer is opportunistic, in equilibrium it proposes a valid block and gets

$$R-c-C.$$

What would happen if the player deviated and did not generate a block? This is off the equilibrium path. Assume the out of equilibrium belief of the other opportunistic players is that the deviator was adversary, so that they believe all proposers will be opportunistic. From that point on, there are two possible scenarios. Denote by $\bar{p}$ the probability that the number of adversaries is $\bar{f}$, if the $\bar{f}-1$ first proposers were adversary. With probability $1-\bar{p}$, all $n-\bar{f}-1$ remaining proposers are indeed opportunistic, in which case the deviator earns $R-c$ (at round n). Alternatively, with probability $\bar{p}$, one of the remaining potential proposers is an adversary. When it is selected to propose a block, the adversary proposes an invalid block. Suppose the other players interpret this as stemming from an opportunistic agent. Then they vote for the block without checking its validity, and the deviator gets $R-c-\kappa$. Hence, the expected gain of the agent who deviated at round $\bar{f}$ is

$$R-c-\bar{p}\kappa.$$

Consequently, the round $\bar{f}$ opportunistic proposer is better off proposing a block than not proposing if

$$R-C-c \geq R-c-\bar{p}\kappa.$$

That is

$$\kappa \geq C/\bar{p},$$

as implied by condition (3).

Finally, consider the case in which we reach round $t<\bar{f}$, because all proposers before t were adversaries. If the round t proposer is adversary, it proposes an invalid block, which is not committed and we move to the next round. If instead the proposer is opportunistic, in equilibrium it proposes a valid block which is committed. This gives the proposer an equilibrium gain equal to

$$R-C-c.$$

If this agent deviated and did not propose any block, this would be an out of the equilibrium action, interpreted by the others as meaning that the proposer was adversary. The players would afterwards follow the same behaviour as on the equilibrium path in the case in which the round t proposer had been adversary. What would be the expected gain of the deviator in that case? Note that the index of the deviating agent is $t \leq \bar{f}$ which is lower than $n-\nu+1+\bar{f}$. Hence, the expected gain of the deviating agent would be $\pi_{check}(t+1)$. Hence, the opportunistic proposer with index t prefers to propose a valid block at t than to propose no block if

$$\pi_{check}(t+1) = R-c-\phi(t+1)C \leq R-C-c.$$

That is $C \leq \phi(t+1)C$. That condition holds, since $\phi(t+1)$ is always larger than or equal to 1.

3) The third step is to analyze the actions of opportunistic players with index $i \leq n-\nu+\bar{f}+1$ when they receive a block at round $t < \bar{f}+1$. On the equilibrium path, these opportunistic players check block validity. To prove equilibrium we must show they prefer to do so rather than deviating once, by voting without checking, and then returning to their equilibrium strategy. When round t is reached, players know that the $t-1$ previous proposers where adversaries. So the average fraction of adversaries, among the $n-t+1$ players who have not been proposers yet is $\frac{f(t)-(t-1)}{n-t+1}$. Denoting the highest index of all adversary players by $i_A = \max\{i : \theta_i = \theta^a\}$, if the next proposer is an adversary, $i_A \leq n-\nu+\bar{f}+1$ and $f = \bar{f}$, then an opportunistic player who is supposed to check block validity is pivotal. Indeed, in that case the $\bar{f}$ adversaries vote for the block, as well as the $n-(n-\nu+\bar{f}+2)+1 = \nu-\bar{f}-1$ opportunistic players who are not supposed to check validity. So, if an opportunistic player who is supposed to check block validity deviates and votes without checking, the total number of votes is $\bar{f}+(\nu-\bar{f}-1)+1 = \nu$ so the block is committed. On the other hand, if $f < \bar{f}$ or $i_A > n-\nu+\bar{f}+1$, then an opportunistic player who is supposed to check block validity is not pivotal. Thus, the expected gain from the one-shot deviation "vote without checking" at round t is

$$\begin{aligned}
&\left(1-\frac{f(t)-(t-1)}{n-t+1}\right)(R-c)\\
&+\frac{f(t)-(t-1)}{n-t+1}\Pr(f=\bar{f}|f\geq t-1)\Pr(i_A\leq n-\nu+\bar{f}+1|\theta_t=\theta^a,\ f=\bar{f})\,(R-c-\kappa)\\
+&\frac{f(t)-(t-1)}{n-t+1}(1-\Pr(f=\bar{f}|f\geq t-1)\Pr(i_A\leq n-\nu+\bar{f}+1|\theta_t=\theta^a,\ f=\bar{f}))\,(\pi(t+1)-c)\,.
\end{aligned}$$

- The first term is the expected payoff of the deviating opportunistic player from the case in which the current block is valid and therefore immediately committed.
- The second term is the expected payoff of the deviating player in the "worst case scenario" in which he was pivotal and triggered to commit an invalid block.
- The third term corresponds to the case in which the deviating opportunistic player is not pivotal, and the invalid block is not committed so that we move to the next round.

Substituting the value of $\pi_{check}(t+1) = R-c-\phi(t+1)C$, the expected continuation value of the deviating player is

$$\begin{aligned}
&\left(1-\frac{f(t)-(t-1)}{n-t+1}\right)(R-c)\\
&+\frac{f(t)-(t-1)}{n-t+1}\Pr(f=\bar{f}|f\geq t-1)\Pr(i_A\leq n-\nu+\bar{f}+1|\theta_t=\theta^a,\ f=\bar{f})\,(R-c-\kappa)\\
+&\frac{f(t)-(t-1)}{n-t+1}(1-\Pr(f=\bar{f}|f\geq t-1)\Pr(i_A\leq n-\nu+\bar{f}+1|\theta_t=\theta^a,\ f=\bar{f}))\,(R-c-\phi(t+1)C-c)\,.
\end{aligned}$$

This simplifies to

$$R-c-\frac{f(t)-(t-1)}{n-t+1}\Pr(f=\bar{f}|f\geq t-1)\Pr(i_A\leq n-\nu+\bar{f}+1|\theta_t=\theta^a,\ f=\bar{f})\kappa$$
$$-\frac{f(t)-(t-1)}{n-t+1}(1-\Pr(f=\bar{f}|f\geq t-1)\Pr(i_A\leq n-\nu+\bar{f}+1|\theta_t=\theta^a,\ f=\bar{f}))\left(\phi(t+1)C+c\right).$$

The equilibrium condition is that this deviation payoff must be lower than the equilibrium continuation payoff of the player

$$R-c-\phi(t)C.$$

That is

$$R-c-\frac{f(t)-(t-1)}{n-t+1}\Pr(f=\bar{f}|f\geq t-1)\Pr(i_A\leq n-\nu+\bar{f}+1|\theta_t=\theta^a,\ f=\bar{f})\kappa$$
$$-\frac{f(t)-(t-1)}{n-t+1}(1-\Pr(f=\bar{f}|f\geq t-1)\Pr(i_A\leq n-\nu+\bar{f}+1|\theta_t=\theta^a,\ f=\bar{f}))\left(\phi(t+1)C+c\right)$$
$$\leq R-c-\phi(t)C.$$

That is

$$\frac{f(t)-(t-1)}{n-t+1}\Pr(f=\bar{f}|f\geq t-1)\Pr(i_A\leq n-\nu+\bar{f}+1|\theta_t=\theta^a,\ f=\bar{f})\kappa\geq\phi(t)C$$
$$-\frac{f(t)-(t-1)}{n-t+1}(1-\Pr(f=\bar{f}|f\geq t-1)\Pr(i_A\leq n-\nu+\bar{f}+1|\theta_t=\theta^a,\ f=\bar{f}))\left(\phi(t+1)C+c\right).$$

Or

$$(f(t)-(t-1))\Pr(f=\bar{f}|f\geq t-1)\Pr(i_A\leq n-\nu+\bar{f}+1|\theta_t=\theta^a,\ f=\bar{f})\kappa\geq(n-t+1)\phi(t)C$$
$$-(f(t)-(t-1))(1-\Pr(f=\bar{f}|f\geq t-1)\Pr(i_A\leq n-\nu+\bar{f}+1|\theta_t=\theta^a,\ f=\bar{f}))\left(\phi(t+1)C+c\right).$$

$$\kappa\geq$$
$$\frac{(n-t+1)\phi(t)-(f(t)-(t-1))(1-\Pr(f=\bar{f}|f\geq t-1)\Pr(i_A\leq n-\nu+\bar{f}+1|\theta_t=\theta^a,\ f=\bar{f}))\phi(t+1)}{(f(t)-(t-1))\Pr(f=\bar{f}|f\geq t-1)\Pr(i_A\leq n-\nu+\bar{f}+1|\theta_t=\theta^a,\ f=\bar{f})}C$$
$$-\frac{(1-\Pr(f=\bar{f}|f\geq t-1)\Pr(i_A\leq n-\nu+\bar{f}+1|\theta_t=\theta^a,\ f=\bar{f}))}{\Pr(f=\bar{f}|f\geq t-1)\Pr(i_A\leq n-\nu+\bar{f}+1|\theta_t=\theta^a,\ f=\bar{f}).}c.$$

Denoting

$$\alpha(t)=\frac{(n-t+1)\phi(t)-(f(t)-(t-1))(1-\Pr(f=\bar{f}|f\geq t-1)\Pr(i_A\leq n-\nu+\bar{f}+1|\theta_t=\theta^a,\ f=\bar{f}))\phi(t+1)}{(f(t)-(t-1))\Pr(f=\bar{f}|f\geq t-1)\Pr(i_A\leq n-\nu+\bar{f}+1|\theta_t=\theta^a,\ f=\bar{f})}$$
$$\beta(t)=\frac{(1-\Pr(f=\bar{f}|f\geq t-1)\Pr(i_A\leq n-\nu+\bar{f}+1|\theta_t=\theta^a,\ f=\bar{f}))}{\Pr(f=\bar{f}|f\geq t-1)\Pr(i_A\leq n-\nu+\bar{f}+1|\theta_t=\theta^a,\ f=\bar{f})},$$

the condition is

$$\kappa>\alpha(t)C-\beta(t)c,$$

as stated in the proposition.[17]

The other possible deviations for the opportunistic player supposed to check block's validity are easier to rule out:

First, the player could do nothing (neither check nor send). In this case, the player would expect to get

$$\frac{f(t)-(t-1)}{n-t+1}\pi_{check}(t+1),$$

while the equilibrium expected gain is

$$\frac{n-f(t)}{n-(t-1)}(R-C-c)+\frac{f(t)-(t-1)}{n-(t-1)}(\pi_{check}(t+1)-C).$$

The latter is larger than the former if

$$\frac{n-f(t)}{n-(t-1)}(R-c)\geq C,$$

which is condition (4).

Second, the player could check the block validity, and then send a message irrespective of whether the block is valid or not. This would generate a lower payoff than the main deviation, shown above to be dominated.

Third, the player could check validity but then send no message. When the current proposer is an adversary, this one-shot deviation yields the same payoff as the equilibrium strategy. When the current proposer is opportunistic, this deviation yields a payoff of $-C$, which is lower than the equilibrium payoff $R-c-C$.

Fourth, the player could check the block's validity and send a message only if the block is invalid, which is trivially dominated.

4) The last step is to analyze the actions of the opportunistic players with index $i\in\{n-\nu+\bar{f}+2,...n\}$ when they receive blocks at round $t<\bar{f}+1$. In equilibrium, these opportunistic players send messages without checking blocks' validity. To finalize the proof we need to show they prefer to follow this equilibrium strategy rather than deviating.

First, consider the possibility to abstain from sending a message. This economizes the costs c, but, in case the block is valid and committed, this implies the agent loses the reward R. So, the deviation is dominated if

$$\frac{n-f(t)}{n-(t-1)}R\geq c,$$

which holds under condition (4).

[17] $\alpha(t)$ and $\beta(t)$ are complicated functions, for which closed form solutions are not readily available, but they depend only on exogenous parameters and are themselves exogenous objects.

Second, consider the possibility of checking validity and sending a message only for valid blocks. This deviation would imply the agent would have to incur the cost of checking (C), but it would economize the cost of sending a message when the block is invalid. So the deviation is dominated if

$$C \geq \frac{f(t) - t + 1}{n - t + 1} c,$$

which holds, since by assumption $C \geq c$.

Other deviations, such as checking validity but never sending messages, or checking validity and always sending messages, or checking validity and sending only if the block is invalid, are trivially dominated.

Q.E.D.

References

Abraham I, Dolev D, Gonen R, Halpern JY (2006) Distributed computing meets game theory: robust mechanisms for rational secret sharing and multiparty computation. *Proceedings of the Twenty-Fifth Annual ACM Symposium on Principles of Distributed Computing, PODC 2006, Denver, CO, USA, July 23-26, 2006*, 53–62.

Abraham I, Malkhi D, Nayak K, Ren L, Spiegelman A (2016) Solidus: An incentive-compatible cryptocurrency based on permissionless byzantine consensus. *CoRR* abs/1612.02916v1.

Afek Y, Ginzberg Y, Feibish SL, Sulamy M (2014) Distributed computing building blocks for rational agents. *ACM Symposium on Principles of Distributed Computing, PODC '14, Paris, France, July 15-18, 2014*, 406–415.

Aiyer AS, Alvisi L, Clement A, Dahlin M, Martin J, Porth C (2005) BAR fault tolerance for cooperative services. *Proceedings of the 20th ACM Symposium on Operating Systems Principles 2005, SOSP 2005, Brighton, UK, October 23-26, 2005*, 45–58.

Amoussou-Guenou Y, Del Pozzo A, Potop-Butucaru M, Tucci-Piergiovanni S (2018) Correctness of tendermint-core blockchains. *22nd International Conference on Principles of Distributed Systems, OPODIS 2018, December 17-19, 2018, Hong Kong, China*, 16:1–16:16.

Aştefanoaei L, Chambart P, Del Pozzo A, Rieutord T, Tucci-Piergiovanni S, Zălinescu E (2021) Tenderbake – a solution to dynamic repeated consensus for blockchains. *International Symposium on Foundations and Applications of Blockchain (FAB)*.

Auer R, Monnet C, Shin HS (2021) Permissioned distributed ledgers and the governance of money. Technical Report 924, BIS working paper, Basel, Switzerland.

Austen-Smith D, Banks JS (1996) Information aggregation, rationality, and the condorcet jury theorem. *American Political Science Review* 34–45.

Bagnoli M, Lipman BL (1988) Successful Takeovers without Exclusion. *The Review of Financial Studies* 1(1):89–110.

Basu S, Easley D, O'Hara M, Sirer EG (2019) Towards a functional fee market for cryptocurrencies. *Available at SSRN 3318327* .

Baudet M, Ching A, Chursin A, Danezis G, Garillot F, Li Z, Malkhi D, Naor O, Perelman D, Sonnino A (2020) State machine replication in the diem blockchain. *The Libra Assn., Tech. Rep* URL `https://developers.diem.com/main/docs/state-machine-replication-paper`.

Belotti M, Kirati S, Secci S (2018) Bitcoin pool-hopping detection. *4th IEEE International Forum on Research and Technology for Society and Industry, RTSI 2018, Palermo, Italy, September 10-13, 2018*, 1–6 (IEEE).

Belotti M, Moretti S, Potop-Butucaru M, Secci S (2020) Game theoretical analysis of cross-chain swaps. *40th IEEE International Conference on Distributed Computing Systems, ICDCS 2020, Singapore, November 29 - December 1, 2020*, 485–495 (IEEE).

Biais B, Bisière C, Bouvard M, Casamatta C (2019) The blockchain folk theorem. *The Review of Financial Studies* .

Blackwell D (1965) Discounted dynamic programming. *The Annals of Mathematical Statistics* 36(1):226–235.

Bradley M (1980) Interfirm tender offers and the market for corporate control. *The Journal of Business* 53(4):345–376.

Buterin V, Reijsbergen D, Leonardos S, Piliouras G (2020) Incentives in ethereum's hybrid casper protocol. *International Journal of Network Management* 30(5).

Cachin C, Kursawe K, Petzold F, Shoup V (2001) Secure and efficient asynchronous broadcast protocols (extended abstract. *in Advances in Cryptology: CRYPTO 2001*, 524–541 (Springer).

Castro M, Liskov B (1999) Practical byzantine fault tolerance. *Proceedings of the Third USENIX Symposium on Operating Systems Design and Implementation (OSDI), New Orleans, Louisiana, USA, February 22-25, 1999*, 173–186.

Chan BY, Shi E (2020) Streamlet: Textbook streamlined blockchains. *IACR Cryptol. ePrint Arch.* 2020:88.

Chen X, Papadimitriou C, Roughgarden T (2019) An axiomatic approach to block rewards. *Proceedings of the 1st ACM Conference on Advances in Financial Technologies, AFT 2019, Zurich, Switzerland, October 21-23, 2019*, 124–131 (ACM).

Crain T, Gramoli V, Larrea M, Raynal M (2017) (Leader/Randomization/Signature)-free Byzantine Consensus for Consortium Blockchains. `http://csrg.redbellyblockchain.io/doc/ConsensusRedBellyBlockchain.pdf`.

David B, Gazi P, Kiayias A, Russell A (2018) Ouroboros praos: An adaptively-secure, semi-synchronous proof-of-stake blockchain. *Advances in Cryptology - EUROCRYPT 2018 - 37th Annual International*

Conference on the Theory and Applications of Cryptographic Techniques, Tel Aviv, Israel, April 29 - May 3, 2018 Proceedings, Part II, 66–98.

de Vries A (2020) Bitcoin's energy consumption is underestimated: A market dynamics approach. *Energy Research & Social Science* 70:101721, ISSN 2214-6296.

Decker C, Seidel J, Wattenhofer R (2016) Bitcoin Meets Strong Consistency. *Proceedings of the 17th International Conference on Distributed Computing and Networking Conference (ICDCN)*.

Downs A (1957) *An Economic Theory of Democracy* (Harper and Row).

Eyal I (2015) The miner's dilemma. *2015 IEEE Symposium on Security and Privacy, SP 2015, San Jose, CA, USA, May 17-21, 2015*, 89–103.

Eyal I, Gencer AE, Sirer EG, van Renesse R (2016) Bitcoin-NG: A Scalable Blockchain Protocol. *13th USENIX Symposium on Networked Systems Design and Implementation, (NSDI)*.

Feddersen TJ (2004) Rational choice theory and the paradox of not voting. *Journal of Economic perspectives* 18(1):99–112.

Feddersen TJ, Pesendorfer W (1996) The swing voter's curse. *The American Economic Review* 86(3):408–424, ISSN 00028282.

Feddersen TJ, Pesendorfer W (1997) Voting behavior and information aggregation in elections with private information. *Econometrica* 65(5):1029–1058, ISSN 00129682, 14680262.

Feddersen TJ, Pesendorfer W (1998) Convicting the innocent: The inferiority of unanimous jury verdicts under strategic voting. *American Political Science Review* 92(1):23–35, URL `http://dx.doi.org/10.2307/2585926`.

Feddersen TJ, Pesendorfer W (1999) Abstention in elections with asymmetric information and diverse preferences. *American Political Science Review* 93:381–398, ISSN 0003-0554, URL `http://dx.doi.org/10.2307/2585402`.

Fooladgar M, Manshaei MH, Jadliwala M, Rahman MA (2020) On incentive compatible role-based reward distribution in algorand. *50th Annual IEEE/IFIP International Conference on Dependable Systems and Networks, DSN 2020, Valencia, Spain, June 29 - July 2, 2020*, 452–463 (IEEE).

Fudenberg D, Tirole J (1991) Perfect bayesian equilibrium and sequential equilibrium. *Journal of Economic Theory* 53(2):236 – 260, ISSN 0022-0531.

Gershkov A, Szentes B (2009) Optimal voting schemes with costly information acquisition. *Journal of Economic Theory* 144(1):36–68, ISSN 0022-0531.

Gilad Y, Hemo R, Micali S, Vlachos G, Zeldovich N (2017) Algorand: Scaling byzantine agreements for cryptocurrencies. *Proceedings of the 26th Symposium on Operating Systems Principles, Shanghai, China, October 28-31, 2017*, 51–68.

Groce A, Katz J, Thiruvengadam A, Zikas V (2012) Byzantine agreement with a rational adversary. *Automata, Languages, and Programming - 39th International Colloquium, ICALP 2012, Warwick, UK, July 9-13, 2012, Proceedings, Part II*, 561–572.

Grossman SJ, Hart OD (1980) Takeover bids, the free-rider problem, and the theory of the corporation. *The Bell Journal of Economics* 11(1):42–64.

Halpern JY, Vilaça X (2016) Rational consensus: Extended abstract. *Proceedings of the 2016 ACM Symposium on Principles of Distributed Computing, PODC 2016, Chicago, IL, USA, July 25-28, 2016*, 137–146.

Halpern JY, Vilaça X (2020) Rational consensus. *CoRR* abs/2005.10141.

Hanke T, Movahedi M, Williams D (2018) DFINITY technology overview series, consensus system. *CoRR* abs/1805.04548.

Kiayias A, Koutsoupias E, Kyropoulou M, Tselekounis Y (2016) Blockchain mining games. *Proceedings of the 2016 ACM Conference on Economics and Computation, EC '16, Maastricht, The Netherlands, July 24-28, 2016*, 365–382.

Kiayias A, Stouka A (2020) Coalition-safe equilibria with virtual payoffs. *CoRR* abs/2001.00047.

Kokoris-Kogias E, Jovanovic P, Gailly N, Khoffi I, Gasser L, Ford B (2016) Enhancing Bitcoin Security and Performance with Strong Consistency via Collective Signing. *Proceedings of the 25th USENIX Security Symposium*.

Kroll JA, Davey IC, Felten EW (2013) The economics of bitcoin mining, or bitcoin in the presence of adversaries. *Proceedings of WEIS*, volume 2013, 11.

Lamport L, Shostak R, Pease M (1982) The byzantine generals problem. *ACM Transactions on Programming Languages and Systems* 4(3):382–401, ISSN 0164-0925.

Ledyard JO (1984) The pure theory of large two-candidate elections. *Public choice* 44(1):7–41.

Liu Z, Luong NC, Wang W, Niyato D, Wang P, Liang Y, Kim DI (2019) A survey on blockchain: A game theoretical perspective. *IEEE Access* 7:47615–47643.

Lysyanskaya A, Triandopoulos N (2006) Rationality and adversarial behavior in multi-party computation. *Advances in Cryptology - CRYPTO 2006, 26th Annual International Cryptology Conference, Santa Barbara, California, USA, August 20-24, 2006, Proceedings*, 180–197.

Manshaei MH, Jadliwala M, Maiti A, Fooladgar M (2018) A game-theoretic analysis of shard-based permissionless blockchains. *IEEE Access* 6:78100–78112.

Miller A, Xia Y, Croman K, Shi E, Song D (2016) The honey badger of BFT protocols. *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, Vienna, Austria, October 24-28, 2016*, 31–42.

Morton RB (1987) A group majority voting model of public good provision. *Social Choice and Welfare* 4(2):117–131, ISSN 01761714, 1432217X.

Morton RB (1991) Groups in rational turnout models. *American Journal of Political Science* 35(3):758–776, ISSN 00925853, 15405907.

Nakamoto S (2008) Bitcoin: A Peer-to-Peer Electronic Cash System. `https://bitcoin.org/bitcoin.pdf`.

Nash J (1951) Non-cooperative games. *Annals of Mathematics* 54(2):286–295, ISSN 0003486X.

Palfrey TR, Rosenthal H (1985) Voter participation and strategic uncertainty. *The American Political Science Review* 62–78.

Persico N (2004) Committee design with endogenous information. *The Review of Economic Studies* 71(1):165–191, ISSN 00346527, 1467937X.

Riboni A, Ruge-Murcia FJ (2010) Monetary Policy by Committee: Consensus, Chairman Dominance, or Simple Majority? *The Quarterly Journal of Economics* 125(1):363–416.

Riker WH, Ordeshook PC (1968) A theory of the calculus of voting. *The American Political Science Review* 62(1):25–42, ISSN 00030554, 15375943.

Schelling TC (1960) *The strategy of conflict* (Cambridge, Harvard University Press).

Smorodinsky R, Tennenholtz M (2006) Overcoming free riding in multi-party computations—the anonymous case. *Games and Economic Behavior* 55(2):385–406, ISSN 0899-8256.

Stoll C, Klaaßen L, Gallersdörfer U (2019) The carbon footprint of bitcoin. *Joule* 3(7):1647–1661, ISSN 2542-4351.

Yin M, Malkhi D, Reiter MK, Golan-Gueta G, Abraham I (2019) Hotstuff: BFT consensus with linearity and responsiveness. *Proceedings of the 2019 ACM Symposium on Principles of Distributed Computing, PODC 2019, Toronto, ON, Canada, July 29 - August 2, 2019.*, 347–356.

Zappalà P, Belotti M, Potop-Butucaru M, Secci S (2020) Brief announcement: Game theoretical framework for analyzing blockchains robustness. Attiya H, ed., *34th International Symposium on Distributed Computing, DISC 2020, October 12-16, 2020, Virtual Conference*, volume 179 of *LIPIcs*, 49:1–49:3 (Schloss Dagstuhl - Leibniz-Zentrum für Informatik).