Incentives compatibility constraints for vote-based consensus protocols

Angelo Murano[1,*], Bruna Bruno[2] and Vincenzo Vespri[3]

[1]Università degli Studi di Salerno, Via Giovanni Paolo II, 132, 84084 Fisciano SA
[2]Università degli Studi di Salerno, Via Giovanni Paolo II, 132, 84084 Fisciano SA
[3]Università degli Studi di Firenze, Viale Morgagni, 67/a I-50134 Firenze

Abstract

For non-cryptocurrency permissioned blockchains to be widely adopted in various business and organizational applications, it is crucial to view nodes as economic agents rather than abstract entities. These agents exhibit behavior driven by their individual interests. Therefore, the consensus protocol must be designed to consider effective incentive compatibility constraints. The aim of this study is to explore blockchain technologies in depth, focusing on the integration of fundamental blockchain features in economic theory and applications. Economics provides a deep understanding of the behaviour, dynamics, and interactions among agents, while computer science focuses on system design, algorithm optimization, and information management. Although economics and computer science have different approaches to problem-solving and decision-making, they can be integrated to create blockchain solutions that are both technically robust and economically sustainable. The paper is structured as follows. After the introduction, the next section provides an overview of economic literature on blockchain technologies. Then, the model description is presented, explaining the conceptual framework. Finally, the concluding section summarizes the key findings and implications of exploring blockchain technologies in an economic context.

Keywords

Non-cryptocurrency, Incentives compatibility, PBFT

1. Introduction

Blockchain technologies have received significant attention from economists, especially in the study of non-fiat currencies and their impact on traditional monetary systems. Non-fiat currencies are those that are not issued or regulated by a central government and lack physical backing like traditional currency. Bitcoin [28] is an example of a decentralised cryptocurrency that operates on the blockchain. Unlike traditional currencies, it is not issued or regulated by a central bank or government. The acceptance and validity of the blockchain system depend on the network of users participating in it and their trust in the security and integrity of the underlying technology.

However, it is important to acknowledge that blockchain is a valuable technology with different applications, exerting significant influence across various sectors including environment, education, health, financial services, trading, utilities, manufacturing, as well as both private and public administrations. Beyond its origins in cryptocurrency, blockchain technology offers innovative solutions in numerous fields, such as supply chain management, healthcare, and digital identity management. For instance, in supply chain logistics, blockchain provides a transparent and tamper-proof record system that enhances traceability and accountability, thereby reducing fraud and ensuring product authenticity [1]. In the healthcare sector, blockchain can secure and streamline the sharing of medical records, improving patient outcomes and safeguarding data privacy [12]. Similarly, digital identity systems

DLT 2024 - 6th Distributed Ledger Technology Workshop
*Corresponding author.
amurano@unisa.it (A. Murano); brbruna@unisa.it (B. Bruno); vincenzo.vespri@unifi.it (V. Vespri)
https://rubrica.unisa.it/persone?matricola=052051 (A. Murano); https://docenti.unisa.it/003707/home (B. Bruno); https://people.dimai.unifi.it/vespri/ (V. Vespri)
0009-0006-2554-1921 (A. Murano); 0000-0001-6767-1069 (B. Bruno); 0000-0002-2684-8646 (V. Vespri)

based on blockchain technology can offer a decentralized and secure platform for identity verification, reducing identity theft and enhancing user control over personal information [15].

Central to the success of these non-cryptocurrency applications are the incentive mechanisms that encourage participation and ensure the integrity of the system. These mechanisms often involve rewarding participants with tokens or digital assets for their contributions to the network, such as validating transactions or providing storage space, which helps maintain the high-security standards necessary for these applications. For example, in blockchain-based supply chains, participants including suppliers and transporters may receive tokens as incentives for maintaining accurate and timely data entries, which are crucial for the effectiveness and reliability of the supply chain [17]. This form of incentivization not only promotes system integrity but also fosters a cooperative environment where all stakeholders are motivated to contribute positively.

This study aims to explore blockchain technologies in-depth, including the integration of fundamental blockchain features within economic theory and applications. It is important to note that applying an economic perspective to the functioning and attributes of blockchain may provide valuable insights to the extensive body of literature that has predominantly evolved within the realm of computer science. It is important to recognise the unique approach of economics in addressing problems and in choices making, which differs from the hypotheses and objectives typically used in computer science investigations.

Contribution of economic theory to computer science practice on blockchains is especially relevant in the realm of some potential applications of blockchain technologies where free-riding or other opportunistic behaviours may emerge. In the economic perspective, blockchain nodes are individuals, whose behaviours and objectives must be considered.

In contexts where blockchain is used outside of cryptocurrency - such as in digital supply chains, public record management, or consensus models for collective decision-making - these issues become particularly pronounced. Free-riding and opportunistic behaviors are crucial concerns in non-cryptocurrency blockchain applications due to the decentralized nature and lack of a central authoritative entity which in traditional systems often acts as a monitor or enforcer. For example, in a blockchain-based digital supply chain, a participant may under-report sales to evade fees or engage in selective misreporting to gain competitive advantages. Without the direct oversight typical in centralized systems, such actions can undermine the integrity and trust upon which the blockchain operates. Real-world data and theoretical analysis show that free riding can lead to significant inefficiencies and increase the cost of maintaining blockchain systems, potentially making them economically unviable without proper incentive mechanisms in place [30]. Furthermore, other studies [7] demonstrate through simulations how differing incentives can lead to varied degrees of participation and honesty among nodes, highlighting the fragile balance required to sustain cooperative behaviour in blockchain networks.

Agents often face complex and dynamic situations that involve rational, emotional, and social elements. Economics takes a unique approach to problem-solving and decision-making, based on a deep understanding of the behaviour, dynamics, and interactions among agents. In contrast, computer science is primarily focused on system design, algorithm optimization, and information management. The issues tackled in computer science are usually well-defined and structured, governed by clear logic and rules. The primary goal is to ensure the efficiency and accuracy of technical solutions. The integration of these perspectives is crucial in the development of blockchain solutions that are not only technically robust but also economically sustainable.

The traditional paradigm in distributed computation categorises agents as either 'good' or 'bad' [2]. In this scenario, individuals who follow the prescribed protocol are considered 'good', while those who use any means necessary to disrupt it are considered 'bad'. Research has shown that the protocol is successful when a given percentage of individuals take on the 'bad' role, highlighting the binary nature of traditional computational perspectives.

To understand the ethical and security challenges presented by blockchain technology, it is useful to analyse a case study that clearly demonstrates the complex relationship between technological vulnerabilities, ethical concerns, and the involvement of malicious actors in shaping the development

of blockchain technology. In 2016, a hacker exploited a vulnerability in the code of the Decentralised Autonomous Organisation (DAO) and transferred 3.6 million Ether, valued at approximately 50 million dollars, into a personal account. The DAO was a pioneering initiative in blockchain-based governance, aiming to democratise investment in Ethereum-based start-ups by enabling investors to allocate funds to proposed projects using Ethers. The breach not only highlighted vulnerabilities within such platforms but also sparked a controversial debate over the blockchain's immutable nature and the ethical principle of 'code is law'. The Ethereum community's decision to rectify the theft through a network fork resulted in significant fallout, including the emergence of Ethereum Classic and a revaluation of hard and soft fork applications in the blockchain domain [25]. Further complicating the situation, the hacker communicated directly with the Ethereum and DAO communities after the attack through an open letter. The hacker defended the extraction of funds as legal under the jurisdiction's laws, exploiting a loophole in the DAO's smart contract code. The hacker claimed that exploiting a systemic flaw was not a crime and threatened legal action against any attempts to recover the stolen Ether. This increased tensions within the community. In addition, the hacker's efforts to sabotage soft fork proposals by offering significant bribes to Ethereum miners highlight the difficulties in implementing technical solutions to ethical violations. This controversial exchange resulted in the abandonment of a soft fork due to identified vulnerabilities, ultimately requiring a hard fork. The division of the Ethereum blockchain into Ethereum (ETH) and Ethereum Classic (ETC) resolved the immediate crisis and sparked a significant discussion on censorship resistance and the principle of blockchain immutability [37].

Other contributions draw attention to the contrasting methodologies employed in computer science and economics within the context of blockchain technology [8]. In computer science and modern cryptography, assumptions typically concern the actions of agents, which are categorized as either honest or faulty nodes. Honest nodes consistently exhibit honest behaviour, while faulty nodes engage in misconduct. In contrast, economists tend to make assumptions about primitives, such as agents' utility functions, and subsequently analyse their strategic behaviours within equilibrium frameworks [8]. The expansive domain for economic exploration within blockchain technology becomes particularly evident when addressing conditions of asymmetric information. In other studies it is emphasized the substantial latitude available for manoeuvring within this nascent field of research [5]. Furthermore, some works accentuate the disparity between computer science hypotheses and economic traditions regarding individual behaviour [4]. The authors emphasise the contrast between the game-theoretic approach, which assumes rationality among all players or processes, and the conventional computer science approach. In the latter, some processes are deemed correct and expected to adhere to the specified protocol, while others are designated as Byzantine, introducing a nuanced perspective on individual behaviour [4].

A practical example of this dynamic is Dogecoin, a cryptocurrency conceived purely ironically in 2013. Contrary to the rational expectations of game theory, Dogecoin gained popularity and market value despite its unconventional origin [34]. From an economic perspective, the existence of cryptocurrencies such as Dogecoin underlines how psychological and social factors can influence economic valuation. This highlights that, in blockchain, economic valuation cannot completely disregard human behavioural dynamics, adding a complex dimension to the relationship between information technology and economics in this emerging field. The term 'Byzantine' originates from the Byzantine generals' problem [20], a theoretical framework that illustrates the complexity of coordinating distributed processes when certain entities may disseminate erroneous information or exhibit malicious behaviour.

In the context of blockchain and consensus protocols, it is crucial for nodes to adhere to the established protocol to ensure network consistency and security. The use of the 'Byzantine' concept recognises the possibility of non-compliant actions by certain nodes. Therefore, it is necessary to develop consensus algorithms that can handle such deviations in behaviour, ensuring the security and integrity of the system, even in the presence of defective or malicious nodes.

A specific investigation delves into numerous economic implications and challenges within the realm of economic science [8]. However, this discussion focuses mainly on the domain of cryptocurrency blockchain. Therefore, it is necessary to conduct an extensive economic analysis of non-cryptocurrency

blockchain, which is highly relevant for economic applications. It is important to note that non-cryptocurrency blockchains differ significantly from their cryptocurrency counterparts in terms of explicit monetary incentives. Given the relevance of non-cryptocurrency blockchain for economic application, economic perspective of non-cryptocurrency blockchain must be investigated.

Non-cryptocurrency applications of blockchain are often private blockchains, due to limitations observed in public blockchains when applied to industrial use cases. It is important to note that public blockchains exhibit constraints that prove detrimental to corporate and administrative applications [9].

Another research has highlighted that incentive-compatible mechanisms widely employed in existing blockchain systems are not seamlessly applicable to non-cryptocurrency scenarios [24]. This statement highlights the need for consensus protocols in business applications that use non-cryptocurrency blockchains. Though non-cryptocurrency blockchains are already employed in some business applications, wider organizational structures with multiple agents interacting on the blockchain (e.g. in public administration) may exhibit more occurrences of opportunistic behaviours to be prevented. Therefore, it is necessary to investigate incentive-compatible mechanisms enforcing non crypto-currency consensus protocols for suitable economic applications.

Contribution of the paper is the analysis of different rewarding schemes ensuring the compatibility of incentives for self-interest agents maximizing their utility functions. As also emphasised recently [24], to our best knowledge the work carried out to develop this approach is still limited. Paper is organized as follows. The next Section presents a brief overview of economic literature investigating blockchain technologies. The model description in following section, while the last Section summarizes concluding remarks.

2. Overview on previous works

In blockchain, incentivization mechanisms play a crucial role in motivating miners to participate in validation processes, contributing significantly to the system's overall maintenance. While effective in cryptocurrency applications, extending existing blockchain systems to non-cryptocurrency domains with distinct business models poses a significant challenge. The lack of effective incentives can be a significant obstacle to the wider adoption of blockchain technology.

An incentive compatibility mechanism characterizes a process where participants would not find it advantageous to violate the rules of the process [11]. In the blockchain domain, it ensures that the consensus protocol will induce validators to exhibit honest and cooperative behaviour, adhering to the prescribed protocol rules and actively contributing to the overall security and stability of the network. Incentives may be in economic form, such as cryptocurrency rewards or penalties for misconduct, or in reputational form, encompassing trustworthiness scores or penalties for any deviations from the prescribed behaviour.

The main challenge is to create an incentive mechanism that can be easily adapted to non-cryptocurrency applications. The need for a compatible incentive mechanism becomes evident in the context of non-cryptocurrency blockchain applications that employ 'Voting-based' consensus algorithms. Research in this domain is currently limited and mostly focused on incentive mechanism design for cryptocurrency applications that use 'Proof-based' consensus algorithms [24].

It is essential to point out that the 'Voting-based' consensus algorithms are used also in cryptocurrency permissionless blockchains, where probabilistic consensus can be achieved due to the large numbers of not identified nodes. In this framework, fast probabilistic consensus protocols [32] [27] may reduce the high communication costs associated to voting mechanisms at a large scale, without incurring in PoW energy costs [33].

It is important to note that studies on incentive compatibility for Proof-of-Work (POW) are not applicable to PBFT-based blockchains due to the inherent disparity between 'Proof-based' and 'Voting-based' consensus algorithms. Bridging this gap is essential for the successful integration of blockchain technology into non-cryptocurrency applications.

To tackle the issue of creating digital identities on a blockchain and implementing an anonymous

voting system to prevent the creation of fictitious digital identities, an approach is needed where the blockchain is used as an immutable ledger to securely store digital identities [19]. In this context, smart contracts should define rules and conditions related to digital identities. For example, they could establish the requirement for verification by a central authority to create or update an identity [14].

To ensure anonymity in the voting system, cryptographic techniques such as zero-knowledge proofs and homomorphic encryption voting systems can be implemented. These techniques allow participants to prove they possess a certain piece of information without revealing it. The proposed zero-knowledge proofs and homomorphic encryption voting system ensures integrity, confidentiality, and authenticity of votes while preserving voter anonymity [26]. Smart contracts manage the voting process and record each vote on the blockchain. Cryptography ensures the anonymity of votes, while smart contracts verify their validity.

To prevent the creation of fictitious digital identities, identity validation is performed before voting through offline verification procedures or interactions with certification authorities. Anti-Sybil mechanisms can be implemented that require approval by existing members before assigning an identity [31]. In addition, smart contracts define approval rules for adding new identities to the registry, requiring the consent of the majority of authorised actors [35].

Despite the shift in consensus algorithms from 'Proof-based' to 'Voting-based', concerns about incentive compatibility remain, particularly in PBFT-based consensus algorithms. Free riders within peer-to-peer networks exploit benefits to the detriment of honest actors [16]. Therefore, it is necessary to have robust mechanisms that can identify and penalise instances of free-riding behaviour. Our proposed approach, based on an economic framework, examines the complex study of participant interactions and strategic behaviours. The establishment of robust and efficient consensus mechanisms, critical for the optimal functioning of blockchain systems, relies heavily on the behaviour of participating entities [23]. However, in distributed systems, especially in open and permissionless settings, it is impractical to assume consistent proactive and honest behaviour from every participant. Therefore, the integration of incentives plays a crucial role in attracting participants and motivating them to adhere to desired behaviours. However, incompatible incentive mechanisms have been identified as a source of compromise to honesty, leading to significant security vulnerabilities.

Acknowledging the rational and self-interested nature of participants, the design of incentive mechanisms rooted in game theory proves indispensable in addressing these challenges [36]. The efficacy of such mechanisms in blockchain systems hinges on the ability to attract a substantial number of participants and motivate them to act proactively. Some studies underscore the positive impact of game theory-based incentives in enhancing the overall effectiveness of blockchain systems [38]. Although game theory has made significant contributions to incentive mechanism design in blockchain systems, limited efforts have been directed towards applying it to PBFT consensus algorithms.

Several studies have analysed the compatibility problems that arise in blockchain systems based on Practical Byzantine Fault Tolerance (PBFT), particularly regarding incentive mechanisms. One notable endeavour, the SmartCast system [18], has systematically scrutinized compatibility challenges through the introduction of an off-chain incentive mechanism. However, this solution's efficacy becomes impractical when applied to applications functioning within a public peer-to-peer network. This is particularly exemplified by certain Internet of Things (IoT) applications [21]. In a distinct proposal, Solidus [3], attempts to address incentive compatibility by presenting a consensus algorithm meticulously designed for Proof-of-Work (POW)-based PBFT. However, it still relies on Proof of Work (POW) for selecting proposers, limiting its use to cryptocurrency contexts rather than a wider range of non-cryptocurrency applications. An alternative scholarly work (Wei et al, 2020) undertakes a game theory-based analysis with a principal focus on the incentive structure of the business model. Unfortunately, this study does not provide a detailed analysis of the complex dynamics involved in maintaining the blockchain network. It focuses mainly on incentives related to business considerations. Concurrently, efforts have been made to enhance Practical Byzantine Fault Tolerance (PBFT) consensus algorithms by incorporating reputation-based methods [8], [22]. Although these algorithms significantly contribute to enhancing the PBFT consensus mechanism, their primary objective is to formulate efficient

selection methods rather than directly addressing the challenges associated with incentive compatibility intrinsic to PBFT-based blockchain architectures. The need for comprehensive solutions that address both consensus efficiency and incentive compatibility remains a significant concern in the developing field of blockchain research.

3. Model

A fundamental process in blockchain functioning is the broadcast process, implying a problem concerning asymmetric information, because nodes have private information on the content of the message they receive. The message content may be a dichotomous variable, as in the retrieve/attack message of the Byzantine problem [20]. In a non-crypto blockchain used for certification purposes (e.g. competencies, property rights, characteristics of individuals or things), the dichotomous choice may be a true/false alternative. Consequently, a faulty or malicious node may have incentives in wrong certifications, namely in certificating features that are not true by broadcasting a message different from the one received. A self-interested validator may adopt opportunist behaviours, intentionally fraudulent. An honest node broadcasts the same message received. Utility functions for honest and malicious nodes may be different, and the incentive compatibility constraints for each type of node should be verified.

To comprehend and enhance the incentive mechanisms in such a system, an understanding of game theory is indispensable. The concept of Nash equilibrium, where each player, recognizing the strategies of others, has no incentive to unilaterally change their strategy, can be crucial in designing systems where honest behaviour is maintained [29]. Furthermore, incorporating dominant strategies, which are advantageous regardless of others' actions, can ensure that maintaining honesty becomes the most beneficial course of action for all nodes. By aligning these utility functions with game theory principles such as Nash equilibrium and dominant strategies, blockchain systems can be designed to be robust against dishonest behaviours and ensure a more secure and trustworthy environment for all participants.

Consider a non-crypto currency blockchain with *n=3k+1* validators, including faulty/malicious validators who are interested in blockchain malfunctioning and/or wrong certifications. In a PBFT protocol the block is accepted if at least *2k+1* validators agree on the block acceptance. The agreement of each validator i is expressed through a *vi* vote *(i=1,..,n)*, which can take two values:

$$v_i = \begin{cases} 1 & \text{if the validator agree,} \\ 0 & \text{if the validator does not agree.} \end{cases}$$

Therefore, the final decision on the block can be summarized as follows:

$$\text{If } \sum_{i=1}^{n} v_i > \frac{2}{3}n, \text{ the block is accepted} \tag{1}$$

$$\text{If } \sum_{i=1}^{n} v_i \leq \frac{2}{3}n, \text{ the block is rejected} \tag{2}$$

This simplified representation of a PBFT protocol can be used to represent validators' behaviours when receiving a block proposal message *M=(x,y)*, where x is the content and y is the result of the voting process on the message. In details:

$$x = \begin{cases} 1 & \text{if the content is true} \\ 0 & \text{if the content is false} \end{cases}$$

$$y = \begin{cases} 1 & \text{if the message reaches the consensus} \\ 0 & \text{if the message does not reach the consensus} \end{cases}$$

Furthermore, we assume that a reputation grade $w_i > 0$, or another kind of non-monetary reward, is associated with the voting (consensus) process.

Finally, we assume that the validators have the capability to access the data encapsulated within the blocks, although they cannot comprehend or interpret the contents. From the validator's perspective, the data housed within the block is simply construed as uncomplicated sequences of bits, even singular bits. The validator's role is emphasised as impartial through this abstraction, with a focus on the binary nature of the data rather than its semantic interpretation.

The following behavioral hypotheses are adopted:
H1: Validators can read the message content as either true or false.
H2: Honest validators receive positive satisfaction from the acceptance of true blocks and the rejection of false blocks (with the same satisfaction gain). Faulty validators receive positive satisfaction from the acceptance of false blocks and the rejection of true blocks (with the same satisfaction gain).
H3: Honest and faulty validators receive positive satisfaction for positive w_i.
With these basic elements of the blockchain, we can describe the utility functions of the faulty (F) and honest (H) validators as $U_{H,F} = f(M, w_i)$ as follows:

$$U'_{H_M} = \frac{\partial U_H}{\partial M} \begin{cases} > 0 & \text{if } M = \begin{cases} (1,1) & \text{a true message is accepted} \\ (0,0) & \text{a false message is rejected} \end{cases} \\ = 0 & \text{otherwise} \end{cases} \tag{3}$$

$$U'_{F_M} = \frac{\partial U_F}{\partial M} \begin{cases} > 0 & \text{if } M = \begin{cases} (1,0) & \text{a true message is accepted} \\ (0,1) & \text{a false message is rejected} \end{cases} \\ = 0 & \text{otherwise} \end{cases} \tag{4}$$

In other words, these utility functions imply that the honest validator receive positive satisfaction from the acceptance of a true block or the rejection of a false block, whereas the faulty validator pursue the objective of false blocks accepted and true blocks rejected (H2). Note that the derivatives are independent from the vote expressed by the validators. As to the reputation grade, both types of validators have increasing utility from the reward (H3).

$$U'_{H_w} = \frac{\partial U_H}{\partial w} > 0 \tag{5}$$

$$U'_{F_w} = \frac{\partial U_F}{\partial w} > 0 \tag{6}$$

The remuneration scheme is fundamental in a mechanism design. Two different remuneration schemes will be analysed, considering alternatively the payoffs for all types of validators at the end of the consensus process (final payoffs), or in the while the consensus process is still running (expected payoffs), that is when the y value is still unknown. The final payoffs are described in a strategic form, the matrix representation of a simultaneous (two players) non cooperative game, with the validator as row player and the State of Nature as column player, who selects the message proposed to validator. Payoffs are quantified only for the row player (the validator). Validator has two available actions/strategies

when receiving the message, corresponding to the values (1,0) chosen for the vote expressed with vi. For each strategy, resulting payoffs are reported according to the different States of Nature.

Scheme A: the validator receives a positive w_i if the block is accepted, and the vote is $w_i = 1$. No reward is provided for the 'no agreement' vote.

Final payoffs – Scheme A

Table 1
Honest Nodes

Actions/Messages	M=(1,1)	M=(1,0)	M=(0,0)	M=(0,1)
$V_i = 1$	U'$_{HM} + U'_{H_w}$		U'$_{H_M}$	U'$_{Hw}$
$V_i = 0$	U'$_{HM}$		U'$_{HM}$	

Table 2
Faulty Nodes

Actions/Messages	M=(1,1)	M=(1,0)	M=(0,0)	M=(0,1)
$V_i = 1$	U'$_{F_w}$	U'$_{FM}$		U'$_{F_M} + U'_{F_w}$
$V_i = 0$		U'$_{F_M}$		U'$_{F_M}$

Table 1 shows the payoffs for honest nodes in Scheme A. It is easy to verify that the agreement vote ($v_i = 1$) is a dominant strategy, even for false messages (when $M = (0, 0)$ or $(0, 1)$). Similarly, for faulty nodes (Table 2), the $v_i = 1$ vote is a dominant strategy. This is because the reward scheme only awards validators who agree on the block. Furthermore, one can observe that for honest nodes, rejecting a false block is preferred to acceptance only if $U'_{HM} > U'_{Hw}$, whereas for faulty nodes, the condition to prefer rejecting a true block to its acceptance is that $U'_{FM} > U'_{Fw}$.

Nevertheless, when validators give their vote, they know if the message content is true or false, but they do not know if the message will be added to the block or not. They may have an expected result y^e, based on the PBFT hypothesis on faulty/honest nodes. If all faulty nodes vote 0 to true blocks and 1 to false block, whereas honest nodes do the opposite, the probability of each type of block to be accepted/rejected are:

$$\pi(1,1) = \pi(0,0) = \frac{2}{3}n + 1 \tag{7}$$

$$\pi(1,0) = \pi(0,1) = \frac{1}{3}n - 1 \tag{8}$$

If they can recognize the true/false feature of the block, they can evaluate the expected payoffs as follows. Let's consider first the expected values for honest nodes (EV_H) when facing true blocks.

$$EV_H(1, y^e) = \begin{cases} (\frac{2}{3}n + 1)(U'_{HM} + U'_{Hw}) & \text{if } v_i = 1 \\ (\frac{2}{3}n + 1 - 1)U'_{HM} & \text{if } v_i = 0 \end{cases} \tag{9}$$

From the above descriptions it easy to verify that $v_i = 1$ is a dominant strategy for an honest node facing true blocks. When coming at false blocks, the expected values are as follows.

$$EV_H(0, y^e) = \begin{cases} (\frac{2}{3}n + 1 - 1)U'_{HM} + (\frac{1}{3}n - 1 + 1)U'_{Hw} & \text{if } v_i = 1 \\ (\frac{2}{3}n + 1)(U'_{HM}) & \text{if } v_i = 0 \end{cases} \tag{10}$$

Here, one would expect that $v_i = 0$ is a dominant strategy, but this is true only if

$$\frac{U'_{HM}}{U'_{Hw}} > \frac{1}{3}n \tag{11}$$

.

Consequently, the false block rejection will depend on the specific form of the utility function (the validator preferences for honest behaviours and reward) and on the blockchain dimension n.

When coming at the expected values for faulty nodes (EV_F), for true blocks it would be:

$$EV_F(1, y^e) = \begin{cases} (\frac{2}{3}n + 1 + 1)U'_{Fw} + (\frac{1}{3}n - 1 - 1)U'_{FM} & \text{if } v_i = 1 \\ (\frac{1}{3}n - 1)U'_{FM} & \text{if } v_i = 0 \end{cases} \tag{12}$$

The faulty should vote 0 to the true block, but this is true only if

$$\frac{U'_{FM}}{U'_{Fw}} > \frac{2}{3}n + 2 \tag{13}$$

.

Whereas for false blocks, the faulty node has a dominant strategy in $v_i = 1$

$$EV_H(0, y^e) = \begin{cases} (\frac{1}{3}n - 1)(U'_{FM} + U'_{Fw}) & \text{if } v_i = 1 \\ (\frac{1}{3}n - 1 - 1)U'_{FM} & \text{if } v_i = 0 \end{cases} \tag{14}$$

The results for the scheme put in evidence that for honest nodes facing true blocks the agreement strategy is always dominant, but the rejection of false blocks is conditioned. Vice versa, faulty nodes will always agree on false blocks, but the rejection of true blocks is conditioned. In this scheme more likely could happen that both true and false blocks will be accepted, compared with the previous scheme.

Scheme B: the validator receives a positive wi when the block is accepted and the vote is vi=1; a positive wi is also received when the block is rejected, and the vote is vi=0.

Final payoffs – Scheme B

Table 3
Honest Nodes

Actions/Messages	M=(1,1)	M=(1,0)	M=(0,0)	M=(0,1)
$V_i = 1$	$U'_{HM} + U'_{H_w}$		U'_{H_M}	U'_{Hw}
$V_i = 0$	U'_{HM}	U'_{H_w}	$U'_{HM} + U'_{H_w}$	

Table 4
Faulty Nodes

Actions/Messages	M=(1,1)	M=(1,0)	M=(0,0)	M=(0,1)
$V_i = 1$	U'_{F_w}	U'_{FM}		$U'_{F_M} + U'_{F_w}$
$V_i = 0$		$U'_{FM} + U'_{F_w}$	U'_{F_w}	U'_{F_M}

There is no dominant strategy both for honest and faulty nodes. The expected payoffs for honest nodes are as follows.

$$EV_H(1, y^e) = \begin{cases} (\frac{2}{3}n + 1)(U'_{HM} + U'_{Hw}) & \text{if } v_i = 1 \\ (\frac{2}{3}n + 1 - 1)U'_{HM} + (\frac{1}{3}n - 1 + 1)U'_{Hw} & \text{if } v_i = 0 \end{cases} \tag{15}$$

It easy to verify that $vi = 1$ is a dominant strategy for an honest node facing true blocks. When coming at false blocks, the expected values are as follows.

$$EV_H(0, y^e) = \begin{cases} (\frac{2}{3}n + 1 - 1)U'_{HM} + (\frac{1}{3}n - 1 + 1)U'_{Hw} & \text{if } v_i = 1 \\ (\frac{2}{3}n + 1)(U'_{HM} + U'_{Hw}) & \text{if } v_i = 0 \end{cases} \tag{16}$$

It easy to verify that $vi = 0$ is a dominant strategy for an honest node facing false blocks. When coming at the expected values for faulty nodes (EV_F), for true blocks it would be:

$$EV_H(1, y^e) = \begin{cases} (\frac{2}{3}n + 1 - 1)U'_{Fw} + (\frac{1}{3}n - 1 - 1)U'_{FM} & \text{if } v_i = 1 \\ (\frac{1}{3}n - 1)(U'_{FM} + U'_{Fw}) & \text{if } v_i = 0 \end{cases} \tag{17}$$

The faulty should vote 0 to the true block, but this is true only if

$$\frac{U'_{FM}}{U'_{Fw}} > \frac{1}{3}n + 3 \tag{18}$$

.

Whereas for false blocks, the faulty node has a dominant strategy in vi=1, if the same conditions as above holds.

$$EV_F(0, y^e) = \begin{cases} (\frac{1}{3}n - 1)(U'_{FM} + U'_{Fw}) & \text{if } v_i = 1 \\ (\frac{1}{3}n - 1 - 1)U'_{FM} + (\frac{2}{3}n + 1 + 1)U'_{Fw} & \text{if } v_i = 0 \end{cases} \tag{19}$$

If the remuneration scheme award both positive votes for accepted blocks and negative votes for rejected blocks, the honest validators are always incentivized to behave correctly, by agreeing on true blocks and rejecting the false ones. On the other hand, the faulty nodes will be incentivized to behave maliciously only if their motivation is strong enough, that is if

$$\frac{U'_{FM}}{U'_{Fw}} > \frac{1}{3}n + 3 \tag{20}$$

.

The above discussion shows that the remuneration scheme A is not incentive compatible for honest nodes, who will prefer to accept also false blocks, also when considering expected payoffs. The false block rejection will depend on the validator's preferences and on the blockchain dimension. Conversely, the second scheme (B) is incentive compatible for honest nodes, who have a dominant strategy in accepting true blocks and rejecting false blocks. Faulty nodes will always agree on false block acceptance, whereas the rejection of true blocks is conditioned.

4. Concluding remarks

This paper starts from the idea that, to widely employ non-cryptocurrency permissioned blockchain in business and organizational applications, it is important to consider that nodes are not abstract entities, but economic agents who behave pursuing their own interests.

With this perspective, the consensus protocol must be designed considering effective incentive compatibility constraints, at least for those currently defined as honest nodes. The proposed analysis of individual behaviour is based on two different types of agents, with different objective function. It compares two different remuneration schemes, showing that the design of mechanisms may be relevant for addressing individual choices toward protocol integrity and secure network environment.

The model we consider for incentive compatibility constraints is inherently static, as the decision-making of each validator is contingent solely upon its conduct within the current round, without regard to its past performance or prevailing network conditions. Such a static approach may entail potential inefficiencies and vulnerabilities within the system. Transitioning from a static to a dynamic model

necessitates several adjustments to accommodate the evolving dynamics of validator participation, potentially fortifying the consensus system within the framework of voting-based blockchains. The dynamic model permits the system to adapt to changing conditions or alterations in validator engagement, thereby enhancing the system's resilience and adaptability.

Within the dynamic model, the reputation of validators becomes subject to fluctuations over time based on their actions and performance. Active participation, consistent voting patterns, and constructive contributions to consensus can lead to an augmentation of a validator's reputation. Conversely, unethical behaviour, erratic voting, or inadequate participation may result in a deterioration of reputation. Dynamic reputation is contingent upon metrics such as voting consistency, adherence to consensus commitments, and overall participation. This framework creates incentives for validators to cultivate a positive, enduring reputation, fostering trust within the network.

Furthermore, the dynamic model takes into consideration the interactions among validators during various consensus phases. For instance, a validator altering its vote during the preparation phase without a valid justification might incur a reputation penalty. Interactions among validators can shape the interpretation of votes and decisions during the consensus process.

Should the incentive compatibility constraints model evolve into a dynamic framework, mechanisms could be implemented to assess the trustworthiness of each validator based on historical behaviour, network participation, quality of service, reputation, and other relevant factors. This augmentation could bolster the security and efficiency of the protocol, rendering it more resistant to potential attacks and scalable. The dynamic model, by incorporating the evolution of validators' actions and decisions over time, introduces a temporal and adaptive component into the consensus system. Time assumes a pivotal role in this model, as past and present actions of validators influence their reputation and, consequently, their future rewards, thereby encouraging consistent and reliable conduct over time.

Situations characterized by a framework akin to the one recently delineated find representation in game theory through repeated games [6]. However, despite existing literature applying game theory to blockchain, none of these works delve into the interaction among validators using an infinitely repeated game theoretical approach to address the issue. Infinitely repeated games provide insights into the structure of individual behaviour when interacting repeatedly, as the outcomes of prior interactions influence their future behaviours. Such games contribute to reinforcing cooperative behaviour [10]. In this form of interaction, players strive to optimize rewards in both the short and long term.

Within an infinitely repeated prisoner's dilemma, the "grim trigger strategy" has been identified as a mechanism to encourage cooperation among players [13][6]. In a grim strategy, a player initiates cooperation (playing honestly) and continues to cooperate unless the other player deviates (acting maliciously) at a certain point. This could potentially represent a future evolution within the context of our work, as a player must commence with cooperation and persist in cooperation to build their reputation.

References

[1] S. A. Abeyratne and R. P. Monfared, "Blockchain ready manufacturing supply chain using distributed ledger," *International journal of research in engineering and technology*, vol. 5, no. 9, pp. 1–10, 2016.

[2] I. Abraham, D. Dolev, R. Gonen, and J. Halpern, "Distributed computing meets game theory: Robust mechanisms for rational secret sharing and multiparty computation," in *Proceedings of the twenty-fifth annual ACM symposium on Principles of distributed computing*, 2006, pp. 53–62.

[3] I. Abraham, D. Malkhi, K. Nayak, L. Ren, and A. Spiegelman, "Solida: A blockchain protocol based on reconfigurable byzantine consensus," *arXiv preprint arXiv:1612.02916*, 2016.

[4] Y. Amoussou-Guenou, B. Biais, M. Potop-Butucaru, and S. Tucci-Piergiovanni, "Committee-based blockchains as games between opportunistic players and adversaries," *The Review of Financial Studies*, vol. 37, no. 2, pp. 409–443, 2024.

[5] L. Ante, "A place next to satoshi: Foundations of blockchain and cryptocurrency research in business and economics," *Scientometrics*, vol. 124, no. 2, pp. 1305–1333, 2020.

[6] R. Axelrod and W. D. Hamilton, "The evolution of cooperation," *science*, vol. 211, no. 4489, pp. 1390–1396, 1981.

[7] G. Bissias, A. P. Ozisik, B. N. Levine, and M. Liberatore, "Sybil-resistant mixing for bitcoin," in *Proceedings of the 13th Workshop on Privacy in the Electronic Society*, 2014, pp. 149–158.

[8] L. Chen, L. W. Cong, and Y. Xiao, "A brief introduction to blockchain economics," in *Information for Efficient Decision Making: Big Data, Blockchain and Relevance*, World Scientific, 2021, pp. 1–40.

[9] O. Dib, K.-L. Brousmiche, A. Durand, E. Thea, and E. B. Hamida, "Consortium blockchains: Overview, applications and challenges," *Int. J. Adv. Telecommun*, vol. 11, no. 1, pp. 51–64, 2018.

[10] A. K. Dixit, S. Skeath, and D. McAdams, *Games of Strategy: Fifth International Student Edition*. WW Norton & Company, 2020.

[11] J. Eatwell, M. Milgate, and P. Newman, *Allocation, information and markets*. Springer, 1989.

[12] M. A. Engelhardt, "Hitching healthcare to the chain: An introduction to blockchain technology in the healthcare sector," *Technology Innovation Management Review*, vol. 7, no. 10, 2017.

[13] C.-P. Fan, "Teaching children cooperation—an application of experimental game theory," *Journal of Economic Behavior & Organization*, vol. 41, no. 3, pp. 191–209, 2000.

[14] K. Gilani, F. Ghaffari, E. Bertin, and N. Crespi, "Self-sovereign identity management framework using smart contracts," in *NOMS 2022-2022 IEEE/IFIP Network Operations and Management Symposium*, IEEE, 2022, pp. 1–7.

[15] A. Jøsang, "Identity management and trusted interaction in internet and mobile computing," *IET Information Security*, vol. 8, no. 2, pp. 67–79, 2014.

[16] S. D. Kamvar, M. T. Schlosser, and H. Garcia-Molina, "Incentives for combatting freeriding on p2p networks," in *Euro-Par 2003 Parallel Processing: 9th International Euro-Par Conference Klagenfurt, Austria, August 26-29, 2003 Proceedings 9*, Springer, 2003, pp. 1273–1279.

[17] K. Korpela, J. Hallikas, and T. Dahlberg, "Digital supply chain transformation toward blockchain integration," 2017.

[18] A. Kothapalli, A. Miller, and N. Borisov, "Smartcast: An incentive compatible consensus protocol using smart contracts," in *Financial Cryptography and Data Security: FC 2017 International Workshops, WAHC, BITCOIN, VOTING, WTSC, and TA, Sliema, Malta, April 7, 2017, Revised Selected Papers 21*, Springer, 2017, pp. 536–552.

[19] S. R. Kumar and M. Goyal, "Administration of digital identities using blockchain," in *2022 5th International Conference on Contemporary Computing and Informatics (IC3I)*, IEEE, 2022, pp. 2179–2183.

[20] L. Lamport, R. Shostak, and M. Pease, "The byzantine generals problem," in *Concurrency: the works of leslie lamport*, 2019, pp. 203–226.

[21] L. Lao, X. Dai, B. Xiao, and S. Guo, "G-pbft: A location-based and scalable consensus protocol for iot-blockchain applications," in *2020 IEEE international parallel and distributed processing symposium (IPDPS)*, IEEE, 2020, pp. 664–673.

[22] K. Lei, Q. Zhang, L. Xu, and Z. Qi, "Reputation-based byzantine fault-tolerance for consortium blockchain," in *2018 IEEE 24th international conference on parallel and distributed systems (ICPADS)*, IEEE, 2018, pp. 604–611.

[23] T. Li, Y. Chen, Y. Wang, *et al.*, "Rational protocols and attacks in blockchain system," *Security and Communication Networks*, vol. 2020, pp. 1–11, 2020.

[24] X. Li, Q. Liu, S. Wu, Z. Cao, and Q. Bai, "Game theory based compatible incentive mechanism design for non-cryptocurrency blockchain systems," *Journal of Industrial Information Integration*, vol. 31, p. 100 426, 2023.

[25] M. I. Mehar, C. L. Shier, A. Giambattista, *et al.*, "Understanding a revolutionary and flawed grand experiment in blockchain: The dao attack," *Journal of Cases on Information Technology (JCIT)*, vol. 21, no. 1, pp. 19–32, 2019.

[26] Y. Miao, "Secure and privacy-preserving voting system using zero-knowledge proofs," *Applied and Computational Engineering*, vol. 8, no. 1, pp. 328–333, 2023. DOI: 10.54254/2755-2721/8/20230181.

[27] D. Mitra, A. Cortesi, and N. Chaki, "A two-hop neighborhood based berserk detection algorithm for probabilistic model of consensus in distributed ledger systems," in *International Conference on Computational Collective Intelligence*, Springer, 2023, pp. 379–391.

[28] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," *Decentralized business review*, 2008.

[29] J. F. Nash *et al.*, "Non-cooperative games," 1950.

[30] A. Pazaitis, P. De Filippi, and V. Kostakis, "Blockchain and value systems in the sharing economy: The illustrative case of backfeed," *Technological Forecasting and Social Change*, vol. 125, pp. 105–115, 2017.

[31] M. Platt and P. McBurney, "Sybil in the haystack: A comprehensive review of blockchain consensus mechanisms in search of strong sybil attack resistance," *Algorithms*, vol. 16, no. 1, p. 34, 2023.

[32] S. Popov and W. J. Buchanan, "Fpc-bi: Fast probabilistic consensus within byzantine infrastructures," *Journal of Parallel and Distributed Computing*, vol. 147, pp. 77–86, 2021.

[33] S. Popov and S. Müller, "Voting-based probabilistic consensuses and their applications in distributed ledgers," *Annals of Telecommunications*, pp. 1–23, 2022.

[34] I. A. Rahman, T. Indrakusuma, A. Widodo, and D. Nuryadin, "Dogecoin price volatility after economic recovery on covid-19 pandemic," *International Journal of Advanced Economics*, vol. 5, no. 6, pp. 129–137, 2023.

[35] V. Shermin, "Disrupting governance with blockchains and smart contracts," *Strategic change*, vol. 26, no. 5, pp. 499–509, 2017.

[36] Y. Wen, F. Lu, Y. Liu, and X. Huang, "Attacks and countermeasures on blockchains: A survey from layering perspective," *Computer Networks*, vol. 191, p. 107 978, 2021.

[37] X. Zhao, Z. Chen, X. Chen, Y. Wang, and C. Tang, "The dao attack paradoxes in propositional logic," in *2017 4th international conference on systems and informatics (ICSAI)*, IEEE, 2017, pp. 1743–1746.

[38] X. Zhu, Y. Li, L. Fang, and P. Chen, "An improved proof-of-trust consensus algorithm for credible crowdsourcing blockchain services," *IEEE Access*, vol. 8, pp. 102 177–102 187, 2020.