SoK

SoK: A Stratified Approach to Blockchain Decentralization

Christina Ovezik[1] *, Dimitris Karakostas[1], and Aggelos Kiayias[1,2]

[1] University of Edinburgh
[2] IOG
{c.ovezik,dkarakos,akiayias}@ed.ac.uk

Abstract. Decentralization has been touted as the principal security advantage which propelled blockchain systems at the forefront of developments in the financial technology space. Its exact semantics nevertheless remain highly contested and ambiguous, with proponents and critics disagreeing widely on the level of decentralization offered by existing systems. To address this, we put forth a systematization of the current landscape with respect to decentralization and we derive a methodology that can help direct future research towards defining and measuring decentralization. Our approach dissects blockchain systems into multiple layers, or strata, each possibly encapsulating multiple categories, and it enables a unified method for measuring decentralization in each one. Our layers are (1) hardware, (2) software, (3) network, (4) consensus, (5) economics ("tokenomics"), (6) client API, (7) governance, and (8) geography. Armed with this stratification, we examine for each layer which pertinent properties of distributed ledgers (safety, liveness, privacy, stability) can be at risk due to centralization and in what way. We also introduce a practical test, the "Minimum Decentralization Test" which can provide quick insights about the decentralization state of a blockchain system. To demonstrate how our stratified methodology can be used in practice, we apply it fully (layer by layer) to Bitcoin, and we provide examples of systems which comprise one or more "problematic" layers that cause them to fail the MDT. Our work highlights the challenges in measuring and achieving decentralization, and suggests various potential directions where future research is needed.

1 Introduction

Bitcoin [130], the first blockchain-based distributed ledger,[3] put forth a new paradigm, that inspired numerous systems to enhance and expand its model and

* The order of the authors follows the Blockchain Technology Laboratory's Author Ordering Policy (https://www.ed.ac.uk/informatics/blockchain/btl-papers/aop).

[3] For the rest of this work we use the terms "blockchain" and "distributed ledger" interchangeably, even though strictly speaking, the latter describes an objective while the former is a means to it.

thousands of applications to be built on them. Alongside, a research discipline emerged across cryptography, distributed systems, game theory and economics, to analyze the properties and capabilities of this paradigm-shifting protocol.

Bitcoin's arguably most important contribution was offering a solution to the consensus problem [113,139] in an open setting. Contrary to classic protocols, cf. [74], Bitcoin participants are not known a priori; instead, the system only assumes a peer-to-peer (P2P) synchronous network and a public setup.[4] Bitcoin's core security argument is that, if a majority of computational power acts honestly, the protocol solves the consensus problem and implements a distributed ledger, as shown formally in [75,138,76]. This, in conjunction with the premise that computational power is widely distributed over the network participants, gives rise to the "security via decentralization" proposition: the system has no single point of failure, as any network participant is individually too weak to influence the properties of the protocol, no matter how they behave. Intuitively, a high degree of decentralization suggests that the trust for safe system operation is spread across the largest possible set of parties.

The appeal of this narrative, and the emergence of ledgers like Ethereum with APIs of higher functionality, gave rise to various "Decentralized Finance" (DeFi) [179] applications. Such systems have drawn the attention of industry, governments, regulators, and banks worldwide. Nonetheless, there is no agreement as to whether blockchain systems are decentralized, or even what "decentralization" entails, despite it being a topic of interest for centuries and across different disciplines [18,172,92]. Proponents often tout the existence of diverse communities, wide geographical distribution, or a theoretical ability of open participation as evidence of decentralization [6]. Antagonists point to power concentration around a few entities when it comes to system maintenance, protocol upgrades, or wealth ownership [155]. Interestingly, both sides might be correct at the same time — to some extent. Blockchains may exhibit high levels of decentralization w.r.t. some aspects, but not others. Thus, the pertinent question is more nuanced than the simple binary one "is the system decentralized or not?" — we are interested to know to what degree and in which aspects the system is (de)centralized.

Another common fallacy is perceiving decentralization as a goal, instead of a means to an end, and equating it with security, stability, or even efficiency. In reality, decentralization guarantees none of these properties. It can be synergistic to them, but in practice centralized systems can be more secure and fail-safe than decentralized ones and vice versa, depending on the relevant threat model. Still, it can be argued that decentralization's major advantage from a security perspective is related to the system's *resilience to single points of failure.*

With this as a starting point, our work sets on exploring decentralization across different layers, or strata, of blockchain systems. In particular, we select layers that influence a distributed system's security properties, e.g., privacy or

[4] Bitcoin uses the following newspaper headline as the common setup string: "The Times 03/Jan/2009 Chancellor on brink of second bailout for banks."

fault tolerance. Thus, centralization in one of our layers points to the existence of a single point of failure for the system as a whole w.r.t. one of those properties.

Our systematization effort aims to inform users, practitioners, and researchers, and to support policymaking and law enforcement processes. Decentralization — or the lack of it — plays a major role in policy discussions and the debate over the regulation of blockchain systems. For example, to determine if a digital asset constitutes a security, and particularly an investment contract, the US Securities and Exchange Commission (SEC) focuses on whether asset owners expect to profit via the efforts of "active participants" (APs), e.g., promoters or sponsors [162] (see also Appendix A). If a system is deemed decentralized across all layers, in effect there is no AP that the system's stakeholders rely on for profiting, so the underlying token would not be classified as a security under this criterion.

We note that many blockchain systems can be argued to have a *potential* for decentralization, due to their permissionless nature. Specifically, by allowing any party to join, they may find themselves in a decentralized state. Nonetheless, our work focuses on characterizing the decentralization of systems as manifested in specific points in time based on the engagement they attract, thus exploring to what degree these systems realize their decentralization potential in the real world, irrespective of whether they can be decentralized in theory.

Related Work. Various research works have addressed the decentralization — or lack thereof — of blockchain systems, from some particular perspective. The research of Zhou [188] and Cho [43] highlights the risk of centralization that arises in the context of *hardware*, when specialized equipment is used by system maintainers to create blocks. This tendency is also acknowledged in the work of Ekblaw *et al.* [62]. Choi *et al.* [44] and Reibel *et al.* [146] reveal high levels of similarity in the codebases of different blockchain projects, alluding to centralization around the *software* used in distributed ledgers. An empirical study by Azouvi *et al.* [12] also looks at software centralization within a single project, i.e., when few individuals undertake the majority of the development process. Neudecker *et al.* [133] identify several ways in which the underlying *network* of a distributed ledger can impact its overall degree of decentralization, while Apostolaki *et al.* [5] examine centralization on the level of Autonomous Systems (ASes) as an enabler of routing attacks on blockchains. A plethora of studies, such as those by Gencer *et al.* [79], Gervais *et al.* [80], Valdivia *et al.* [173] or Lin *et al.* [117], have focused on the decentralization of the *consensus* layer, by measuring the "mining power" ratio of a system's block producers. Another blockchain dimension whose decentralization has been thoroughly studied is the one pertaining to the economics of cryptocurrencies — often termed *tokenomics*. Sai *et al.* [157], Cheng *et al.* [42] and Ron and Shamir [152] analyze the distribution of transactions and tokens across parties, while Moore and Christin [128] touch on the subject of secondary markets and the risk carried by their potential centralization. Chatzigiannis *et al.* [39] point out that most blockchain light *client* schemes are vulnerable to centralization because of their reliance on centralized servers or full nodes, a concern also shared by Moxie Marlinspike [122].

Gervais *et al.* [80] present examples of centralization from the space of blockchain *governance*, and particularly conflict resolution, while Azouvi *et al.* [12] complement this work with a more systematic exploration of the contributors behind improvement proposals and discussions. Various works, such as those of Mariem *et al.* [121] and Sun *et al.* [165], turn their attention to the *geographic* dispersion of participants and infrastructure within a blockchain ecosystem.

Despite the breadth and depth of the research around blockchain (de)centralization and its manifestations, there have been few efforts so far to generalize or systematize this knowledge. Sai *et al.* [156] offer a blockchain centralization taxonomy, based on an algorithmic literature review and expert interviews. They treat ledgers as multi-layer systems, capturing 13 aspects of centralization over 6 architectural layers: Application, Operational, Incentive, Consensus, Network, and Governance. However, their work neglects some components, such as software centralization (as identified in [44,146,12]), or geographic decentralization pertaining to layers other than the network (for example, the decentralization of consensus participants, as studied by Sun *et al.* [165]). More recently, Zhang *et al.* [185] propose a taxonomy around five facets of decentralization: Consensus, Network, Wealth, Governance, and Transactions. They focus primarily on transaction centralization (w.r.t. the distribution of transactions to users), which is mainly a measure of adoption and usage, rather than a dimension with security implications. Their systematization also does not account for several factors identified in previous research, including hardware [188,43], software [44,146,12], or geographic [121,165] decentralization. Last, there exist some studies that approach the topic of blockchain decentralization from different perspectives, e.g., economic or social [26,25]. Notably, while all these works offer ample information on blockchain decentralization, none of them propose a consistent methodology for determining the decentralization level across all relevant layers.

2 Methodology

Decentralization in the context of blockchains is often reduced to particular aspects of the system e.g., consensus participation. Nonetheless, distributed ledgers comprise multiple essential, interacting components. Drawing from all sources of prior research, our work discerns the layers that form a ledger in a "bottom-up" manner.[5] Starting from the physical layer, i.e., *Hardware*, we systematize blockchain decentralization in multiple strata, all the way up to *Governance*. We also include *Geography* as a dimension that touches upon all other layers (Figure 1). We note that this layering is applied only on the ledger's stack; exploring decentralization in exogenous infrastructure (e.g., physical links, Internet routing, operating systems *etc.*) is an interesting question, but outside the scope of this paper.

A first step to understand the importance of decentralization for such systems is to identify the properties of interest that distributed ledgers should satisfy and

[5] Our stratification is inspired by the OSI conceptual network model, cf. [166].

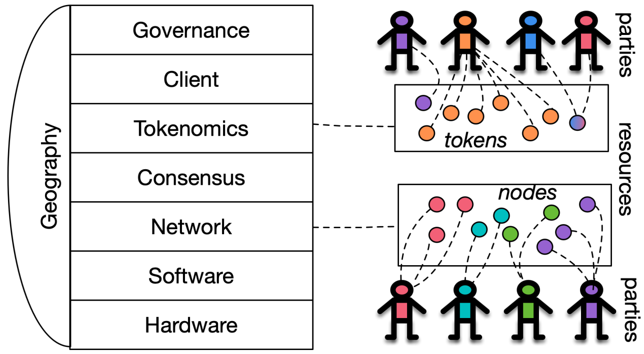

Fig. 1. Illustration of our methodology: the layers of a blockchain system, identification of some type of resources in two of the layers (tokens, peers), their assignment to relevant parties exhibiting a higher (network) and lower (tokenomics) degree of decentralization and an example of equal joint ownership of one token.

which can be affected by the ledger's degree of decentralization. The two core security properties that each ledger should guarantee are *safety* and *liveness*. Safety ensures that all honest users hold the same, "settled" view of the ledger. Liveness reflects the ability to update the ledger's settled view regularly, as new transactions are submitted. We note that safety and liveness typically incorporate and express other useful properties for real-world applications, such as transaction finality or censorship resistance. A third property, *privacy*, guarantees that users' actions enjoy a certain degree of dissociation from their real-world identities and individually are unlinkable. Finally, specifically in the context of blockchain systems, *price stability* captures the property that the ledger's core digital asset's supply and market price are predictable (to a reasonable extent). In particular, price stability is violated if the asset's market price demonstrates high volatility in the short term (e.g., monthly).[6] For ease of reading, we will refer to this property only as "stability" for the rest of the paper.

We view these properties through the lens of (cyber-)security, i.e., in the context of an adversary who wishes to subvert them. This is strictly stronger compared to settings where failures are assumed to be benign (e.g., crash faults due to power outages). A single point of failure exists when a single party, if controlled by the adversary, can violate one or more of the ledger's properties.

We lay out the following methodology: for each system layer we identify (a) one or more *resources*, that can be thought of as the basic "unit" of the layer pertaining to the ledger's security properties; (b) the *relevant parties* that control, either directly or indirectly, said resources; (c) the ledger's *properties* that are at risk, if the resources' distribution across the relevant parties becomes

[6] The cryptocurrency market is notoriously volatile, so one could apply different thresholds for "reasonable" levels of volatility. An interesting line of future research would be to identify non-cryptocurrency assets, which could serve as a base of comparison for which levels of volatility are acceptable in this case.

centralized. For example, considering Bitcoin's consensus layer, the resource is hashing power and the relevant parties are the miners; the properties at risk are safety, liveness and, to a somewhat lesser degree, stability and privacy. Table 1 provides a summary of our systematization, with resources and relevant parties presented for each identified layer and sub-category.

Notably, a resource might be modular, with some parts considered more important than others. For instance, software products are typically not monolithic, with e.g., documentation being less crucial than a library or a configuration file. Therefore, the parties that maintain the former have (arguably) less influence over the resource (i.e., the software product) than the coders of the latter. To resolve this concern, one could compute an aggregate level of decentralization, after *weighing* each component based on its significance. Such aggregation methodology could also be applied to compute the decentralization level of the whole system, assuming weights for each layer (cf. Section 12).

Another issue is that one relevant party may encompass multiple real-world identities. For instance, consider two software products, one maintained by a single organization with many members, the other maintained by a handful of independent developers. Although the first may be more decentralized in terms of people, from a legal perspective the second may be deemed more decentralized, as the first is authored by a single legal entity.

By projecting the relevant parties of each category to *legal persons*, we articulate a test that can be useful in assessing systems w.r.t. their decentralization in a legal sense (Definition 1). Here, a legal person can be an organization, e.g., a company or non-profit foundation, or an individual.

Definition 1. *A blockchain system fails the* Minimum Decentralization Test *(MDT) if and only if there exists a layer (cf. Table 1) for which there is a single* legal person *that controls a sufficient number of* relevant parties *so that it is able to violate a property of interest.*

In the following sections, we provide detailed explanations as to why each identified layer is important and how it fits into our framework (Sections 3-10). Then, we apply our methodology on case studies (Section 11), and finally, we suggest various directions for future research that our work points to (Section 12).

3 Hardware

The role of hardware in the potential decentralization of blockchain systems has been reported in various research works [188,43]. To be specific, by "hardware", we refer to the machines that host and/or support the consensus software, which can be anything from personal computers to purpose-built devices. In many cases, the hardware is also provided as a service from cloud providers to consensus participants. To account for all possibilities, we segment this layer into two categories, namely "physical" and "virtual" hardware.

Physical hardware. This category covers all hardware that is used directly by consensus participants. Bitcoin mining, and that of other Proof-of-Work (PoW)

Layer	Subcategory	Resources	Relevant Parties
Hardware [188,43]	Physical hardware	Participating power	Hardware manufacturers
	Virtual hardware	Participating power	Cloud providers
Software [44,146,12]	Protocol participation	1) Participating power 2) Full nodes	Developers of full node software
	Asset management	Tokens	Developers of wallet software
Network [133,5,79,73]	Topology	Component bridges	Owners of bridges
	Peer discovery	Bootstrapping nodes	Node operators
Consensus [79,80,173,117]		1) Participating power 2) Block content	Owners of participation nodes
Tokenomics [157,42,152,128]	Initial distribution	Bootstrapping tokens	Token holders
	Token ownership	Tokens	1) Addresses 2) Key managers 3) Legal stakeholders
	Secondary markets	Market liquidity	1) Exchanges 2) Trading pairs
Client API [39,122]		Tokens	Full node operators
Governance [80,12,156]	Conflict resolution	Decision-making power	All system entities
	R&D funding	Capital	Active developers
Geography [121,165,156]	Physical safety	All resources above	Regions
	Legal compliance	All resources above	Jurisdictions

Table 1. Overview of blockchain decentralization layers, including for each layer the literature that motivated it and the way it fits into our framework.

ledgers, started from regular CPUs, but quickly migrated to GPUs and, eventually, to dedicated devices (ASICs), which produce more hashes at a lower cost [167]. In the case of Proof-of-Stake (PoS) blockchains, participation is typically performed through generic hardware, although there are systems with particular requirements, which may restrict the compatible hardware options (an example such blockchain is The Internet Computer by Dfinity [56]). Some alternative schemes, such as certain forms of Proof-of-Useful-Work (PoUW) or Proof-of-Elapsed-Time (PoET) also make use of Trusted Execution Environments (TEEs) to guarantee higher security or efficiency levels [16].

When analyzing decentralization on this layer, the *resource* of interest is participating power (e.g., hashes per second or stake) and the *relevant parties* are the manufacturers of hardware products that are used for participating. That is, we consider a system to be decentralized in the hardware layer if the participating power is distributed across various pieces of equipment, manufactured by different entities.

As demonstrated by the evolution of Bitcoin mining, one pathway to hardware centralization is the development and adoption of specialized machines that outperform generic equipment and provide an advantage to their operators. Nowadays, an overwhelming majority of block production in PoW blockchains comes from specialized hardware, despite the fact that PoW does not require specialized hardware in theory [167]. Notably, this trend has motivated significant research in "ASIC-resistance" and the development of PoW algorithms that attempt to facilitate better hardware diversity [21,64,148]. In other cases, strict or "non-typical" protocol requirements (e.g., the requirement for trusted hardware) reduce the pool of possible manufacturers that can support those systems, potentially leading to increased centralization.

Concentration around few hardware manufacturers creates various hazards. Same-vendor products are more susceptible to collective faults, e.g., due to defective parts or hardware bugs. Such faults could result in sudden drops in the network's power, lowering the threshold for gaining a computational majority (*safety* and *liveness* hazard) and slowing down block production, at least until the PoW parameters are recalculated (*liveness* hazard). Manufacturers could also introduce backdoors, threatening the ledger's security and *stability*, albeit such hazards can possibly be mitigated via cryptographic techniques [8].

Virtual hardware. The emergence of mining data centers allowed for hashing power to be offered as a "cloud" service, effectively enabling miners to participate in block production without possessing their own hardware [119,169]. The advent of PoS protocols reinforced this trend towards "virtual" hardware, by decoupling Sybil resilience from physical requirements. In theory, this enables PoS nodes to run on generic hardware, e.g., even home equipment, but in practice, convenience often drives PoS users to employ cloud services, which offer uptime and connectivity guarantees that a DIY configuration cannot. This is exacerbated when PoS systems apply penalties to absent users and uptime guarantees become of utmost importance to guarantee profitability. Therefore, the *resource* of interest in this category is again participating power (either in the form of hashing power or stake) and the *relevant parties* are cloud providers.

When nodes that control significant participating power are hosted by the same provider, significant hazards arise for all properties. First, the provider may have access to private keys and hence is able to create conflicting blocks (*safety* hazard) or deanonymize users (compromising *privacy*). Second, the provider controls the node's network access, so it could prohibit communication (*liveness* hazard). Finally, it could also tamper with the system's *stability*, e.g., increasing price volatility via targeted interference or even stealing user rewards.

4 Software

Software development is another dimension of distributed ledgers that has been associated with potential centralization [44,146,12]. Diverse software development and usage is a core element of stability and safety of distributed ledgers, as it increases resilience to catastrophic bugs in a product's code. A vulnerability in one implementation may jeopardize a part of the system, but if the system is sufficiently decentralized, such vulnerabilities would not escalate to systemic threats. Following, we discuss the development of core blockchain software components, namely transaction validation and PoW mining (via full nodes) and management of keys and digital assets (via wallets).[7]

Protocol Participation. The principal type of software in blockchain systems is the *full node*. Full nodes implement the ledger protocol by: i) keeping a local chain; ii) validating new transactions; iii) extending the local chain with new blocks; iv) participating in the consensus mechanism to incorporate new blocks. To analyze decentralization, we identify two *resources* of interest: 1) (number of) full nodes; 2) participating power (e.g., computational or stake) that is hosted on full nodes. The *relevant parties* are full node software developers.

Relying on a handful of full node implementations introduces *safety*, *liveness*, and *stability* hazards. A bug that fails to validate correct transactions would hurt the system's liveness, whereas accepting incorrect transactions could hurt the system's safety and stability, e.g., via a network split or token forgery. Such bugs have been observed in Bitcoin Core and could have resulted in Denial-of-Service (DoS) attacks [72] and token forgery [48]. Implementation bugs could also threaten *privacy*, if nodes reveal information about message origin (e.g., IP addresses). Similar threats arise when code gets reused across different projects. Often, a new blockchain is only a "derivative", that is a project that started by forking an existing codebase, including copyrighted information [146]. Such projects often remain unpatched, so bugs in the initial implementation tend to spill over [44]. Vulnerabilities may also arise when adapting an existing implementation in a new setting, e.g., copying Bitcoin's code and replacing PoW with PoS [94]. Thus, widespread usage of multiple node implementations, developed by different teams, is a hallmark of a secure and reliant ecosystem.

Asset Management. Distributed ledgers are mostly used for bookkeeping of digital asset transactions, so securely managing and transferring assets is a core necessity. Digital assets are typically managed by private keys and represented via addresses. The software responsible for managing keys and addresses is the wallet [98] and its principal functionalities are: i) store the user's keys; ii) prove ownership of the assets (managed by the keys); iii) issue transactions that transfer assets to other accounts; iv) retrieve the user's (keys') balance and history information. Therefore, the *resource* of interest is the set of all assets managed via the ledger and the *relevant parties* are, like before, software developers.

The wallet is a major point of security consideration, so multiple properties rely on it. A bug which e.g., corrupts the user's keys could not only prohibit a

---
[7] Appendix B also explores software testing.

user from transacting with their assets (*liveness* hazard), but forever lose access to them (*stability* hazard). This was demonstrated in 2017, when the "Parity" Ethereum wallet saw a vulnerability that allowed a user to take ownership of multiple assets and then lock them [86]; as a result, 300m worth of Ethereum tokens were forever lost. In another example, some Bitcoin wallets possibly displayed incorrect balance, effectively enabling a double spending attack [114]. Consequently, if a few implementations are predominantly used, a vulnerability could result in assets being unusable or stolen. Such vulnerability could also turn into a systemic point of failure, with the whole ledger becoming unusable.[8]

5 Network

Blockchain nodes communicate over a peer-to-peer (P2P) network, the decentralization properties of which have been of great interest to researchers [133,5,79,73]. Systems often implement a message *diffusion* mechanism [75] via a gossip protocol that avoids full graph connectivity [54]. Users typically use the Internet to access the P2P overlay, though some efforts try to introduce an independent infrastructure [47]. Here we explore networking aspects which present single points of failure, in terms of topology and bootstrapping. A notable research question that arises organically from our analysis, and touches upon both following subsections, is creating a P2P network that is both permissionless and Byzantine resilient, e.g., as explored in [49,123].

Topology. The first networking aspect of interest is the network's topology. Evaluating a real-world network's clustering properties is a well-known problem, traditionally done by generating random graphs and comparing the expected with the observed values [63,134,1]. In blockchain systems, every node maintains a list of peer connections. Crucially, message provenance is not typically provided, so no party can know the network origin of an incoming message. Therefore, the *resource* of interest are "bridges" (single nodes or small cluster of nodes) between the network graph's components and the *relevant parties* are the bridges' owners or operators. Here, a component is a single node or a cluster of tightly interconnected nodes.

A distributed network, in the tradition of Baran [17], is key in maintaining *safety* and *liveness*. Under the CAP theorem [30], any networked system can satisfy at most two of the following properties: i) a consistent data copy; ii) data availability; iii) network partition tolerance. Ledger systems are no exception. Each node needs to maintain at least one connection to an honest party to receive all messages and avoid eclipse attacks [84].

If some parties cannot communicate with the rest of the network, either they halt and any transactions they submit are dropped (violating *liveness*) or they

[8] A prime such example was "The DAO", which in 2016 attracted nearly 14% of all Ethereum tokens and, when hacked, instigated a change in Ethereum's consensus layer and a hard fork which split the network [160].

produce separate ledger versions (violating *safety*).[9] By preventing communication between a node and the rest of the network, an attacker reduces the honest computational power and isolates that node, making a 51% or a double spending attack (against the isolated node) easier to deploy. Additionally, an adversary can violate *liveness* by blocking the node's transactions from reaching the rest of the network. Third, an adversary that controls all of a node's connections can link transactions to the specific node, correlate the user's addresses, and associate them with real-world information, such as an IP address, thus compromising *privacy*[10]. This also holds on a macroscopic level, i.e., a node that acts as a central communication hub between two clusters can obtain information and even deanonymize some participants (violating *privacy*).

Node Bootstrapping and Peer Discovery. Joining a ledger's network and synchronizing with it is the so-called "bootstrapping" process. All real-world ledgers rely on an initial (trusted) setup, the first ("genesis") block.[11] Obtaining the correct genesis block is done in an out-of-band, typically secure mechanism, since it is often well-known and easy to validate from various sources.

Initially, the node needs to connect to some peers and receive all available chains. Therefore, the *resource* here is bootstrapping nodes and the *relevant parties* are these nodes' operators. Then, using the ledger protocol's chain resolution mechanism, it decides which chain to adopt. There are two points of interest here.

First, it should be guaranteed that the node connects to at least one honest peer, to retrieve all available information and avoid getting eclipsed. As explained above, if nodes gets eclipsed, the system's *liveness*, *safety* and *privacy* properties may get compromised. However, connecting to honest peers is not straightforward, given that the node has no knowledge about the network participants, a standard problem in P2P networks [57]. Blockchain systems predominantly use either hardcoded peer lists or DNS seeding [118], although both techniques are censorship-prone. Notably, [118] showed that more secure alternatives, such as ZMap [61], cannot be feasibly used in existing blockchain systems.

Second, even if the node connects to some honest peers, catching up by using only genesis ("bootstrapping from genesis") is not always feasible. Especially in PoS systems, an attacker can effortlessly assemble an arbitrarily-long, seemingly correct chain (violating *safety*) [78,89,116]. To counter such attacks, some ledgers employ checkpoints [105,52,50], which are often issued centrally and are either hardcoded or received from the peers. Other solutions do exist, e.g., analyzing block density and relying on key erasures [13] or using VDFs [27,53], but

[9] In longest-chain protocols, like Bitcoin, miners keep producing blocks in isolation, ending up with different ledger versions. In BFT-style protocols, like Algorand, the ledger cannot be updated if a large number of block producers become unreachable and thus do not adopt new transactions.

[10] Note that protecting privacy in the network layer can be a particularly challenging problem, even in the setting where no single adversary controls all of a node's connections.

[11] Some proposals rely on computational assumptions instead of a trusted setup [77], but their real-world performance and applicability is still untested.

enforcing and/or relaxing such assumptions still poses an interesting research problem.

6 Consensus

A key element of any distributed ledger is its consensus protocol, and a lot of the research behind blockchain decentralization has been dedicated to this layer [79,80,173,117]. Protocols in our context are "resource-based", i.e., they are executed by parties possessing units of an underlying resource (e.g., hashing power or stake). To guarantee safety and liveness, at least a majority (or in some cases a supermajority of $\frac{2}{3}$) of the ledger's participating resources should be honestly controlled [74]. Therefore, when a handful of actors control enough resources to break one of the properties, a direct point of failure arises.

Protocols like Bitcoin [75], Algorand [40], or Ouroboros Praos [75], enable resource holders to engage in the protocol directly with (essentially) whatever amount of resources they have. In these protocols, block producers can, even though they do not have to, form coalitions called *pools*. In PoW, a pool "leader" validates transactions, and organizes them in a candidate block, while each "member" executes the PoW puzzle for the leader-made block. If a member is successful, the leader collects the block's reward and distributes it, proportionately to each member's power. In PoS, the leader has full control over the block's creation, while the members only pay fees to delegate their staking rights to the leader and collect rewards. Pooling behavior is also driven by temporal discounting [145], i.e., the tendency to disfavor rare or delayed rewards. In essence, a small miner may prefer small frequent payments, at the cost of some fee, over rare large payments, when producing a block.

Other systems, like Cosmos [112] and EOS [90], impose restrictions on which parties can participate in consensus and require the rest to delegate their resources to a representative or "validator" node. This "barrier to entry" means that any party without enough stake, i.e., below the system's threshold or less than its competitors, is required to delegate their staking rights to a validator. At every "epoch", a committee of (a fixed number of) parties is elected to run the protocol. The election mechanism is voting-based, with resource delegation acting as the voting process.

In both types of systems, there are two *resources* of interest: i) owned participating power, e.g., computational or stake; ii) delegated participating power, including the power to choose a block's content. Accordingly, the *relevant parties* are: i) miners and stakeholders, who own hashing power and stake respectively; ii) pool leaders and delegates, who control how the resources are used.

Typically, the security of a ledger is guaranteed if the parties that represent an aggregate majority of the participating power are honest (i.e., they follow the protocol as prescribed) [75]. Therefore, the concentration of participating power around few entities poses a threat to the system. This hazard is well-known and blockchain users and participants have actively tried to avoid it since at least 2014 [83]. Those controlling a power majority can hurt *liveness*, by refusing to publish or accept certain transactions, as well as *safety* by launching a long-range attack. Both types of attacks also indirectly hurt *stability*, since the system's trustworthiness is challenged.

A second concern revolves around block proposers. A proposer is a party that maintains a mempool and chooses which transactions are added to a block and in what order. Initially, a single party acted as both block proposer and builder. With the increase in hardware requirements needed to run a full node and the formation of pools, the two roles of proposer and builder were separated.

In PoW ledgers, the leader of the pool typically proposes the block's content, whereas the pool members only run the PoW algorithm. Therefore, pool members are not involved in a block's construction and often do not even validate its contents. Therefore, the leader may censor transactions (*liveness* hazard), steal member rewards (*stability* hazard), or possibly link the user's resources with information like IP addresses (*privacy* hazard).

In addition, smart contracts enable MEV-type attacks [187], which might hurt stability. Here, block builders have the ability to observe transactions before publication and choose their order in a block, which they can exploit to extract value from honest transactions. A countermeasure that has been introduced is the proposer-builder separation (PBS) model, wherein a trusted party maintains a mempool and proposes a block, whereas validators sign it without ever observing its content (thus not being able to exploit its MEV) [34]. Still, the current implementation of PBS in Ethereum has been criticized for facilitating censorship and centralization, hence its usefulness remains unclear [85].

Finally, a threat arises due to the lack of self-healing, i.e., the inability to recover from a temporary adversarial takeover. In PoW, even if a majority gets corrupted, honest users can increase their own power and, eventually, overthrow the adversary and restore the ledger's security [11,14]. In PoS though, power shift takes place on the ledger, by transferring stake. If an adversary temporarily obtains a majority, they can prohibit transactions that shift power away from them, thus retaining control indefinitely (for example, a large centralized cryptocurrency exchange can make it hard to issue outgoing payments and withdrawals, while enabling payments between different users of the exchange). Consequently, a diverse stake distribution (cf. Section 7) is vital to protect against takeovers.

7 Cryptocurrency Economics

A core component of ledger systems is their native token. Tokens compensate system maintenance and accommodate value transfers. They are treated as currency or assets by their users, thus forming a market economy. To record data on the ledger, e.g., payments or interactions with applications, users obtain tokens to pay the corresponding fees. System maintainers get compensated in tokens to offset their costs. Several studies have considered the distribution of tokens and their availability (e.g., on exchanges) as integral parts of a blockchain and its eventual degree of decentralization [157,42,152,128]. Accordingly, in this section, we explore decentralization in blockchain-based economies, in terms of initial token distribution, token ownership, and secondary markets.

Initial Token Distribution. To bootstrap the system, a blockchain protocol defines two parameters: i) the distribution of tokens at the system's launch, and ii) how new tokens are generated and distributed as the system evolves. Thus, the generated tokens form the *resource* of interest, while the *relevant parties* are the token holders.

As with other aspects, Bitcoin led the way and other systems explored alternatives. In Bitcoin, no coins existed prior to its beginning, i.e., there was no "pre-mine." Starting from genesis, each block creates a predetermined amount of coins, based on a rate that converges to 21 million tokens in existence [130]. New coins, along with transaction fees, are awarded to the miner that produces each block. Therefore, to acquire new tokens a user gathers enough computing power to produce a block. In other blockchain systems, some tokens were sold via traditional markets before the blockchain was deployed. This approach, termed "Initial Coin Offering" (ICO), enabled funding the project with the future proceeds of the token investment. In return, investors acquired a pre-launch amount of tokens, which was codified in the chain's first block. In terms of token generation, most systems employ a variation of Bitcoin's mechanism, e.g., Ethereum blocks yield 2 new tokens, while others, like Cardano, employ elaborate mechanisms to incentivize pooling around a target number of pools [31].

The initial token distribution is particularly important in PoS systems, where Sybil resilience relies on it (cf. Section 6). If centralized around a few parties, e.g., via pre-mining (or "pre-minting"), early investors have to maintain the system in its early stages, while also receiving the early blocks' rewards. Fewer consensus participants during this time lowers the threshold for adversarial takeover, threatening the system's *safety* and *liveness*. In both PoW and PoS systems, new users are onboarded if early investors sell tokens on secondary markets. Consequently, early investors control the system's expansion and valuation, impacting its *stability*.

Finally, the process of distributing tokens might be elemental for privacy-oriented systems. Typically, such projects employ zero-knowledge protocols that rely on a secure construction of a common reference string (CRS). If the CRS's construction is centralized, then the party that creates it can deanonymize all transactions or violate their correctness. To avoid such hazards, various ceremony protocols have been proposed in order to construct the CRS in a distributed manner [140,101,135,108].

Token Ownership. Diverse token ownership plays a central role in the usability and security of a blockchain. Hence, the system's circulating tokens are the *resource* of interest, while the *relevant parties* are: i) addresses; ii) key managers; iii) legal asset owners. This distinction arises due to the existence of custodians, who control assets on behalf of other stakeholders, and users controlling multiple addresses.

If most tokens are owned by a few parties, many hazards arise. First, PoS systems' security, i.e., *safety* and *liveness*, relies directly on diverse token ownership, which makes corrupting enough parties to control a majority of tokens more threatening. Second, the token's price may be manipulated, posing a risk on the system's *stability* and, indirectly, security, in both PoS and PoW systems. Specifically, participation cost, e.g., for mining equipment or electricity, is denominated in fiat currency. However, miner income from block rewards comes in tokens. Thus, miners need to sell part of the rewards (for fiat) to pay for their operational costs. If the market is volatile, profitability is more precarious and miners are possibly less inclined to participate, which can impact the *safety* or *liveness* of the system by reducing the threshold for conducting a 51% attack.

Various factors drive token ownership centralization. Initial tokens are often allocated centrally (see above). System incentives, e.g., fixed token supply, generally favor hoarding tokens instead of spending them. Finally, rich participants may accumulate capital faster than small ones, an inevitability in pseudonymous systems where downwards wealth redistribution is impossible [97].

Secondary Markets. Distributing the tokens to a wide population is predominantly made on secondary markets. The rate of token production is typically slow, depending on block production, and the new tokens are often distributed to existing users. Therefore, new users are onboarded via centralized exchanges and, to a lesser extent, face-to-face transactions. The tokens that are bought and sold through these markets constitute the *resource* of interest, when it comes to measuring the decentralization of secondary markets, while the *relevant parties* are i) the assets for which they are bought and sold ("trading pairs"),[12], and ii) the exchanges that host these trades.

Many hazards arise when tokens are available on limited markets. First, exchanges offer little *privacy* guarantees, so their operators have full access of user data, following KYC regulations. Second, exchanges are largely *unregulated* by financial authorities and may engage in market manipulation. Third, few marketplaces often result in lower liquidity. Thus, the threshold for manipulating the token's price by some percentage, via selling or buying tokens, also lowers. Similarly, if most of the token's liquidity is allocated to a few trading pairs, then it becomes exposed to the problems of the tokens at the other end of the pairs (e.g., the collapse of one of these systems might trigger a huge liquidity loss). All such events threaten the system's *stability*, while also, when mining profitability drops due to the token's devaluation, *safety* and *liveness* are indirectly hurt.

8 Client API

To join a blockchain system, full nodes need to download and parse the entire ledger, which often amounts to hundreds of GBs.[13] The ledger's state, which is usually stored in memory, is also large[14] and often poorly maintained [96]. Consequently, maintaining a full node requires significant computational and storage capacity and, eventually, becomes impossible to host on home equipment.

[12] In our context the liquidity of a trading pair is measured across *all* exchanges.
[13] Bitcoin: 485 GBs; Ethereum: 819 GBs. [bitinfocharts, Etherscan; August 2022]
[14] Bitcoin's UTxO set is 4.71 GBs. [Satoshi info; August 2022]

This concern is well-known and ongoing research tries to resolve it via ledger compression [103,29,99]. In practice though, users often employ third-party services that offer an interface to the ledger [122]. Given the widespread use and variety of applications that rely on such services, the "client API" layer can be susceptible to centralization [39]. The *resources* we identify in this case are tokens, which are stored in wallets without ledger verification capabilities, and the *relevant parties* the full node operators that service them. We note that solutions like Simplified Payment Verification (SPV) [130] or succinct verification proofs [51,104,103,32], which are not full nodes but do validate the ledger to some extent, are not considered here. Instead, we focus on wallets that rely entirely on a trusted node for access to the ledger's content.

Many properties are at risk for such wallets by corrupted full node services. For example, the service could perform a double-spending attack (*safety* violation), by presenting to the wallet a transaction that is not in fact published on the ledger; observe that, without any access to the ledger itself, the wallet needs to trust the data presented by the node. Similarly, since the wallet relies on the full node for transaction processing and balance computations, *liveness*, *privacy*, and *stability* hazards arise, as the node can block, de-anonymize or, depending on the implementation, divert a user's funds and transactions.

9 Governance

Governance in blockchain systems is a broad topic [102,141,19] and of high interest among studies of blockchain decentralization [80,12,156]. Here, we focus on two aspects: i) improvements and conflict resolution; ii) fund allocation for research and development (R&D).

Improvements and Conflict Resolution. Decision-making mainly concerns conflicts that arise regarding potential blockchain modifications and improvement proposals. Proposals may affect *mining*, e.g., changing the PoW function [127] or switching to PoS [65], the *consensus protocol*, e.g., changing block structure [147], or *token ownership*, e.g., denylisting [33]. In theory, anyone can propose changes in blockchain systems and respond in some way, depending on their role. In essence, full nodes assume executive, legislative, and judicial powers by operating the ledger and choosing its rules, while other actors voice their opinions by affecting the token's market price [3]. The governance *resource* is decision-making power, which may take various forms, and the *relevant parties* are all active entities in the system.

If the other relevant layers are centralized, governance follows suit. For instance, if mining is concentrated around a few operators, they might force a choice by mining on one ledger. Where a voting mechanism is employed to reach a decision, voting power typically corresponds to each participant's wealth, with each token granted one vote (due to the pseudonymous nature of these systems). If ownership is concentrated around a small number of stakeholders, the decisions might aim at benefiting these few parties in the short term, at the expense of the system's long-term benefit. If a disagreement turns into a stalemate, systems may split into distinct ledgers, that share the same history up to a point but diverge thereupon [142,33]. These outcomes harm the system's *stability*, and indirectly threaten its *safety* and *liveness*.

An effective governance process should prevent such harmful events. However, it is not always possible to make sound decisions in a decentralized manner, as demonstrated by theoretical results in social choice such as Arrow's impossibility theorem [7]. Additionally, when agents act in a selfish manner, as is presumed in distributed ledgers, efficiency can degrade (cf. "Price of Anarchy" [110]). Therefore, decentralized decision-making processes face a challenge, as they need to address various social choice theory (e.g., Arrow's theorem) and game-theoretic (e.g., rational ignorance [60]) considerations.

R&D Funding. Funding for research and development can cover the maintenance of legacy codebase, research in features like privacy and scalability, market incentives, e.g., stabilizing the token's price at times of high volatility, and more. Thus, the *resource* of interest here is capital and the *relevant parties* are the active researchers and developers in a ledger's ecosystem.

Ledgers typically make no funding provisions, besides allocating rewards from coin issuance and transaction fees. R&D is conducted via corporate vehicles which rely on traditional funding models, such as venture capital. However, since designing and implementing hardware and software for distributed ledgers is particularly expensive, this model can lead to centralization, as discussed in Sections 3 and 4. In addition, lack of funding or concentration around a few teams may delay crucial updates or new features, thus hindering *stability*.

A common alternative to traditional financing is ledger "self-funding." Here, the system defines a *treasury*, i.e., a pot which accumulates funds over time that are allocated for R&D [2,184]. A treasury is typically managed collectively by the ecosystem, often via an open and inclusive process where anyone can submit proposals and the system's stakeholders vote for funding allocation. Therefore, a treasury can help nurture a diverse ecosystem of development teams, albeit it is not, on its own, a sufficient condition for decentralized R&D funding.

10 Geography

Geographic decentralization is a key point of interest [121,165,156], and it touches upon all layers covered in the previous sections. Accordingly, it involves all resources described so far, e.g., hashing power or tokens. Nonetheless, it constitutes a dimension on its own, as parts of a system may be well distributed w.r.t. one dimension but geographically concentrated.[15]

The tendency to centralize in certain areas arises due to economical, technological, or sociopolitical factors. For example, miners often set up their operations in countries with low electricity costs, hardware companies operate in countries with small production costs, nodes are hosted in areas with high internet speed,

[15] For example, independent actors may participate in mining within a single country.

and tokens are accumulated by residents of countries with low taxes and where many exchanges operate. Geographical centralization poses two main threats to the properties of a ledger: i) physical hazards and ii) legal impediments.

Physical safety. Physical hazards could threaten a system's infrastructure. If part of a system is located in a small area, connectivity failures or outages could destroy or split the ledger's network.[16] This concern is particularly relevant in PoW, where equipment is hard to relocate. All *resources* examined above can be impacted (e.g., via drops in hashing power or token loss when mining equipment or cold storage is damaged), while the *relevant parties* are the regions of resource concentration. Single points of failure may arise when geographically-concentrated nodes act as central hubs, harming either *safety* or *liveness* (cf. Section 5) and, indirectly, *stability* (e.g., due to increased market volatility).

Legal compliance. Failures can also possibly occur due to legal pressures. If some layer is concentrated in a specific jurisdiction, authorities can possibly restrict or subvert it. Again, this touches upon all *resources* examined so far, as all are influenced by the law, with the *relevant parties* being legal jurisdictions. Depending on the occasion, different properties of the system are impacted. For example, if a country bans Bitcoin mining, the power drop could decrease the threshold for controlling a majority (*safety* hazard), while blocks are produced at a slower pace until the PoW difficulty is recalculated (*liveness* hazard). *Stability* could also be hurt, if some part of the system, e.g., mining, software access, or asset ownership is restricted. Additionally, exchanges can be legally bound to follow KYC procedures to comply with AML regulations [126], linking the users' identities to their activity something that may lead to compromising their *privacy*. Arguably, a system is more likely to uphold its properties by falling under many jurisdictions, such that violating the properties requires the coordinated efforts of multiple authorities.

11 Case Studies

We now apply our methodology to a number of case studies. First, we review Bitcoin's status w.r.t. each identified layer, showcasing how a project can be analyzed across all strata. Second, we apply the Minimum Decentralization Test (MDT) (cf. Definition 1) on an array of projects and show that they fall short due to centralization some layer — specifically the governance layer.

Cross-layer Study: Bitcoin
Hardware. No concrete data could be found on the distribution of hashing power across PoW mining products. Although hundreds of ASICs are available, on top of generic hardware (e.g., GPUs), as of 2022 the market appears centralized

[16] Although exogenous hazards, like natural disasters, are outside the scope of this work, concentration in a small area can enable weaker adversaries to disrupt a ledger's execution. For example, an adversary could isolate a building or a single computing center from the rest of the network, although assuming an adversarial disruption of the grid of an entire country, or continent, can be considered unrealistic.

around 4 ASIC manufacturers [186]. Interestingly, only some ASICs are profitable, so, unless the token's price increases without an increase in PoW difficulty, mining should be expected to concentrate around these products.[17]

Software. An overwhelming majority (99.15%) of full nodes run Bitcoin Core,[18] so Bitcoin is completely centralized around (releases of) this product. Regarding the distribution of tokens across wallets, no data could be found.

Network. Regarding peer discovery, Bitcoin Core sets 8 outgoing and 125 incoming connections, chosen randomly from known and/or hardcoded peers. Most Bitcoin nodes communicate over Tor, making topology analyses particularly hard.[19] Nonetheless, it is estimated that the network is evenly spread across multiple Autonomous Systems, thus presenting high levels of decentralization [5].

Consensus. On the consensus layer, Bitcoin presents mixed results regarding decentralization (cf. Section 6). Hashing power is distributed across thousands of machines. Although no concrete data could be found, folklore evidence suggests that these machines are owned by a highly diverse set of users. However, Bitcoin also observes high levels of centralization around pools, i.e., w.r.t. block formation and the input to the PoW module; specifically, at the time of writing, 4 pools control more than 75% of the whole network's mining power.[20]

Tokenomics. At its onset, no Bitcoin tokens existed. They were generated and allocated as the system progressed. Early participants were disproportionately favored, as half of all tokens were created within the first two years, when consensus participation was sparse and mining was conducted by only a few parties. As more transactions were issued, the tokens were distributed more widely, albeit wealth is still highly centralized, compared to real-world economies (cf. Table 2). Specifically, approx. 43M addresses own some amount of tokens, with the top 100 addresses controlling 14.01% of all wealth. Nonetheless, tokens are traded on more than 100 marketplaces at volumes of approx. $53B (cf. Table 3).

Client API. Most of the available Bitcoin wallet software is either SPV or explorer-based [98]. In the first case, the wallet downloads only the block headers, so it does not validate each block's transactions, while in the second case the wallet relies entirely on a server. However, no data could be found on the ownership of Bitcoin tokens w.r.t. wallet types, therefore Bitcoin's decentralization w.r.t. the client API layer is inconclusive.

Governance. Deciding on improvement proposals and conflict resolution in Bitcoin is somewhat centralized, but not entirely. Specifically, decisions, which are made by accepting suggestions via GitHub, are typically taken by a small set of developers, who are often the ones to comment during the relevant discussions [80,12]. In terms of development funding, Bitcoin makes no provisions. Therefore, the available data are inconclusive on how many sources of funding exist, e.g., companies and foundations, and how much influence each has.

[17] A profitability calculator is available at nicehash.com.
[18] Source: blockchair.com
[19] 53.3% of Bitcoin's nodes operate over Tor. [bitnodes; October 2022]
[20] Source: statista.com

Geography. As mentioned earlier, most Bitcoin nodes communicate over Tor, which makes analyzing the network's topology difficult. Nonetheless, Bitcoin miners, although fairly well distributed with a presence in 95 different countries, tend to cluster in certain areas. At the time of writing, more than $\frac{1}{3}$ of mining is located in the USA, with Kazakhstan and Russia following with 18% and 11% respectively [4]. In terms of full nodes (which may not participate in mining), USA and Germany see roughly equivalent shares, with other countries hosting far fewer nodes (cf. Table 4), although still a majority communicates anonymously.

MDT Studies. We now turn our attention to projects that fail the Minimum Decentralization Test (MDT), showcasing how the MDT can be used to identify points of centralization in blockchain systems.
Fiat-backed stablecoins. A prime example of projects that fail the MDT is fiat-backed stablecoins. Briefly, in a USD-backed stablecoin system, for each token that is live on the ledger there exists $1 which is held in escrow in a company's bank account, s.t. at any point in time, a token holder can exchange their tokens for the equivalent number of USD. Therefore, the main selling point is that the token should always (in theory) be valued by the market at $1. Such projects include Tether (USDT), USDC, Gemini Dollar (GUSD), TrueUSD (TUSD), and Binance USD (BUSD).[21] In all these systems, there exists a single legal entity which is responsible for issuing tokens when receiving USD and redeeming tokens in exchange for the USD held in escrow. Therefore, these entities are single points of control within the *governance* layer of each system.
Wrapped tokens. Another family of systems for which the MDT often fails is bridges. A bridge enables transferring assets of one ledger to another, e.g., Bitcoin to Ethereum. This is achieved by creating a "wrapped" version of the original token on the destination, with each wrapped token corresponding to a (frozen) token on the source side. These systems can also be centralized in the governance layer, as a single custodian is typically responsible for the creation and destruction of the wrapped tokens. Such bridge examples include Wrapped Bitcoin (WBTC), with BitGo being solely responsible for minting new tokens, and Huobi Bitcoin (HBTC), with Huobi being the custodian.[22]

12 Discussion

Our main contribution is the systematization of the rather fragmented body of literature related to decentralization into a unified framework, under which the decentralization of any distributed ledger can be analyzed. Table 1 summarizes our methodology, i.e., the layers that comprise a ledger and the relevant resources and parties that guarantee its core properties. Removing single points of failure via a diverse distribution of resources across independent parties is critical in guaranteeing security, privacy, and stability and can also have legal implications (cf. Appendix A). Our work also opens various research threads.

[21] USDT: tether.to, USDC: circle.com, GUSD: gemini.com, TUSD: tusd.io, BUSD: binance.com
[22] WBTC: wbtc.network, HBTC: htokens.finance

First, even though we do not treat off-chain "Layer 2" protocols [93] as a distinct stratum, our methodology can be readily applied to a combination of Layer 1 (main chain) and Layer 2 protocols. For example, one could apply our layering methodology on a combination of Bitcoin and the Lightning Network [144] and assess the decentralization of the combined system as a payment network. Similarly, our methodology can assess blockchain-based "decentralized applications" (DApps), e.g., DeFi systems.[23] Investigating the decentralization of Layer 2 protools or DApps via our methodology is thus a promising research thread.

Second, exploring the relationship between decentralization and fault tolerance, as well as the settings where decentralization is beneficial and those where it is not, is another interesting topic of future research.[24]

Third, we offer a framework for analyzing blockchain decentralization, but not specific quantitative metrics. A compelling direction for future work is systematically identifying the right metrics for each layer to capture all relevant aspects of blockchain decentralization.[25] Building on this, a natural end-goal of all questions posed above is producing a quantitative blockchain decentralization index. Historically, it has been observed that increasing decentralization on one axis coincides with, or even results in, centralization on another [158]. Future work should further explore the dynamics between all layers, resolve possibly inescapable trade-offs, determine the importance (weight) that should be assigned to each layer and each metric, and efficiently combine them into an index.

References

1. Albert, R., Barabási, A.L.: Statistical mechanics of complex networks. Reviews of modern physics **74**(1), 47 (2002)
2. Allen, D.W., Berg, C., Lane, A.M.: Trust and governance in collective blockchain treasuries. Available at SSRN (2021)
3. Alston, E.: Constitutions and blockchains: Competitive governance of fundamental rule sets. Case Western Reserve Journal of Law, Technology & the Internet **11**(4) (2019)
4. for Alternative Finance, C.C.: Cambridge bitcoin electricity consumption index (cbeci): Bitcoin mining map (2022), `https://ccaf.io/cbeci/mining_map`
5. Apostolaki, M., Zohar, A., Vanbever, L.: Hijacking bitcoin: Routing attacks on cryptocurrencies. In: 2017 IEEE Symposium on Security and Privacy. pp. 375–392. IEEE Computer Society Press, San Jose, CA, USA (May 22–26, 2017). `https://doi.org/10.1109/SP.2017.29`
6. Araya, D.: The promise of bitcoin: An interview with andreas m. antonopoulos (2016), `https://futurism.com/promise-bitcoin-interview-andreas-m-antonopoulos`
7. Arrow, K.J.: A difficulty in the concept of social welfare. Journal of political economy **58**(4), 328–346 (1950)

[23] In such cases, smart contract development would naturally fall in the software layer.

[24] Appendix E expands on the relationship between fault tolerant systems and decentralization.

[25] Appendix D offers an overview of metrics that have been used so far in this setting.

8. Ateniese, G., Kiayias, A., Magri, B., Tselekounis, Y., Venturi, D.: Secure outsourcing of cryptographic circuits manufacturing. In: Baek, J., Susilo, W., Kim, J. (eds.) ProvSec 2018: 12th International Conference on Provable Security. Lecture Notes in Computer Science, vol. 11192, pp. 75–93. Springer, Heidelberg, Germany, Jeju, South Korea (Oct 25–28, 2018). https://doi.org/10.1007/978-3-030-01446-9_5
9. Atkinson, A.B., et al.: On the measurement of inequality. Journal of economic theory **2**(3), 244–263 (1970)
10. Avan-Nomayo, O.: Bitcoin.org blocks access to bitcoin software download in the uk (2021), https://cointelegraph.com/news/bitcoin-org-blocks-access-to-bitcoin-software-download-in-the-uk
11. Avarikioti, G., Käppeli, L., Wang, Y., Wattenhofer, R.: Bitcoin security under temporary dishonest majority. In: Goldberg, I., Moore, T. (eds.) Financial Cryptography and Data Security - 23rd International Conference, FC 2019, Frigate Bay, St. Kitts and Nevis, February 18-22, 2019, Revised Selected Papers. Lecture Notes in Computer Science, vol. 11598, pp. 466–483. Springer (2019). https://doi.org/10.1007/978-3-030-32101-7_28, https://doi.org/10.1007/978-3-030-32101-7_28
12. Azouvi, S., Maller, M., Meiklejohn, S.: Egalitarian society or benevolent dictatorship: The state of cryptocurrency governance. In: Zohar, A., Eyal, I., Teague, V., Clark, J., Bracciali, A., Pintore, F., Sala, M. (eds.) FC 2018 Workshops. Lecture Notes in Computer Science, vol. 10958, pp. 127–143. Springer, Heidelberg, Germany, Nieuwpoort, Curaçao (Mar 2, 2019). https://doi.org/10.1007/978-3-662-58820-8_10
13. Badertscher, C., Gazi, P., Kiayias, A., Russell, A., Zikas, V.: Ouroboros genesis: Composable proof-of-stake blockchains with dynamic availability. In: Lie, D., Mannan, M., Backes, M., Wang, X. (eds.) ACM CCS 2018: 25th Conference on Computer and Communications Security. pp. 913–930. ACM Press, Toronto, ON, Canada (Oct 15–19, 2018). https://doi.org/10.1145/3243734.3243848
14. Badertscher, C., Gaži, P., Kiayias, A., Russell, A., Zikas, V.: Consensus redux: Distributed ledgers in the face of adversarial supremacy. Cryptology ePrint Archive, Report 2020/1021 (2020), https://eprint.iacr.org/2020/1021
15. Bank, W.: Gini index (world bank estimate) (2022), https://data.worldbank.org/indicator/SI.POV.GINI
16. Bao, Z., Wang, Q., Shi, W., Wang, L., Lei, H., Chen, B.: When blockchain meets sgx: An overview, challenges, and open issues. IEEE Access **8**, 170404–170420 (2020)
17. Baran, P.: On distributed communications networks. IEEE transactions on Communications Systems **12**(1), 1–9 (1964)
18. Bastiat, F.: Popular Fallacies regarding General Interests; being a translation of the "Sophismes économiques"... With notes by GR Porter [the translator]. John Murray (1846)
19. Beck, R., Müller-Bloch, C., King, J.L.: Governance in the blockchain economy: A framework and research agenda. Journal of the Association for Information Systems **19**(10), 1 (2018)
20. Bendiksen, C.S., Gibbons, S., Lim, E.: The bitcoin mining network-trends , composition , marginal creation cost , electricity consumption & sources (2018), https://api.semanticscholar.org/CorpusID:149451309
21. Biryukov, A., Dinu, D., Khovratovich, D.: Argon2: new generation of memory-hard functions for password hashing and other applications. In: 2016 IEEE European Symposium on Security and Privacy (EuroS&P). pp. 292–302. IEEE (2016)

22. Biryukov, A., Khovratovich, D., Pustogarov, I.: Deanonymisation of clients in bitcoin P2P network. In: Ahn, G.J., Yung, M., Li, N. (eds.) ACM CCS 2014: 21st Conference on Computer and Communications Security. pp. 15–29. ACM Press, Scottsdale, AZ, USA (Nov 3–7, 2014). https://doi.org/10.1145/2660267.2660379
23. Board, F.S.: Crypto-asset markets. Potential channels for future financial stability implications **10** (2018), https://www.fsb.org/wp-content/uploads/P101018.pdf
24. Board, F.S.: Decentralised financial technologies: Report on financial stability, regulatory and governance implications. June **6**, 2019 (2019), https://www.fsb.org/wp-content/uploads/P060619.pdf
25. Bodó, B., Brekke, J.K., Hoepman, J.H.: Decentralisation: a multidisciplinary perspective. Internet Policy Review **10**(2), 1–21 (2021)
26. Bodó, B., Giannopoulou, A.: The logics of technology decentralization–the case of distributed ledger technologies. Blockchain and web **3** (2019)
27. Boneh, D., Bonneau, J., Bünz, B., Fisch, B.: Verifiable delay functions. In: Shacham, H., Boldyreva, A. (eds.) Advances in Cryptology – CRYPTO 2018, Part I. Lecture Notes in Computer Science, vol. 10991, pp. 757–788. Springer, Heidelberg, Germany, Santa Barbara, CA, USA (Aug 19–23, 2018). https://doi.org/10.1007/978-3-319-96884-1_25
28. Bonneau, J., Heninger, N. (eds.): FC 2020: 24th International Conference on Financial Cryptography and Data Security, Lecture Notes in Computer Science, vol. 12059. Springer, Heidelberg, Germany, Kota Kinabalu, Malaysia (Feb 10–14, 2020)
29. Bonneau, J., Meckler, I., Rao, V., Shapiro, E.: Coda: Decentralized cryptocurrency at scale. Cryptology ePrint Archive, Report 2020/352 (2020), https://eprint.iacr.org/2020/352
30. Brewer, E.: Cap twelve years later: How the" rules" have changed. Computer **45**(2), 23–29 (2012)
31. Brünjes, L., Kiayias, A., Koutsoupias, E., Stouka, A.: Reward sharing schemes for stake pools. In: IEEE European Symposium on Security and Privacy, EuroS&P 2020, Genoa, Italy, September 7-11, 2020. pp. 256–275. IEEE (2020). https://doi.org/10.1109/EuroSP48549.2020.00024, https://doi.org/10.1109/EuroSP48549.2020.00024
32. Bünz, B., Kiffer, L., Luu, L., Zamani, M.: FlyClient: Super-light clients for cryptocurrencies. In: IEEE S&P 2020 [88], pp. 928–946. https://doi.org/10.1109/SP40000.2020.00049
33. Buterin, V.: Hard fork completed (2016), https://blog.ethereum.org/2016/07/20/hard-fork-completed/
34. Buterin, V.: Two-slot proposer/builder separation (2021), https://ethresear.ch/t/two-slot-proposer-builder-separation/10980
35. Cao, T., Yu, J., Decouchant, J., Luo, X., Veríssimo, P.: Exploring the monero peer-to-peer network. In: Bonneau and Heninger [28], pp. 578–594. https://doi.org/10.1007/978-3-030-51280-4_31
36. Cavallaro, L., Kinder, J., Wang, X., Katz, J. (eds.): ACM CCS 2019: 26th Conference on Computer and Communications Security. ACM Press (Nov 11–15, 2019)
37. cdetrio: Ropsten revival (2017), https://github.com/ethereum/ropsten/blob/master/revival.md
38. Cevallos, A., Stewart, A.: A verifiably secure and proportional committee election rule. In: Baldimtsi, F., Roughgarden, T. (eds.) AFT '21: 3rd ACM Conference

on Advances in Financial Technologies, Arlington, Virginia, USA, September 26 - 28, 2021. pp. 29–42. ACM (2021). https://doi.org/10.1145/3479722.3480988, https://doi.org/10.1145/3479722.3480988
39. Chatzigiannis, P., Baldimtsi, F., Chalkias, K.: Sok: Blockchain light clients. In: International Conference on Financial Cryptography and Data Security. pp. 615–641. Springer (2022)
40. Chen, J., Gorbunov, S., Micali, S., Vlachos, G.: ALGORAND AGREEMENT: Super fast and partition resilient byzantine agreement. Cryptology ePrint Archive, Report 2018/377 (2018), https://eprint.iacr.org/2018/377
41. Chen, T., Li, Z., Zhu, Y., Chen, J., Luo, X., Lui, J.C.S., Lin, X., Zhang, X.: Understanding ethereum via graph analysis. ACM Transactions on Internet Technology (TOIT) **20**(2), 1–32 (2020)
42. Cheng, L., Zhu, F., Liu, H., Miao, C.: On decentralization of bitcoin: An asset perspective (2021)
43. Cho, H.: Asic-resistance of multi-hash proof-of-work mechanisms for blockchain consensus protocols. IEEE Access **6**, 66210–66222 (2018)
44. Choi, J., Choi, W., Aiken, W., Kim, H., Huh, J.H., Kim, T., Kim, Y., Anderson, R.: Attack of the clones: Measuring the maintainability, originality and security of bitcoin'forks' in the wild. arXiv preprint arXiv:2201.08678 (2022)
45. CoinYuppie: It's time to formally face the centralization problem that infura brings (2022)
46. Commission, C.F.T.: Cftc orders coinbase inc. to pay $6.5 million for false, misleading, or inaccurate reporting and wash trading (2021), https://www.cftc.gov/media/5796/enfcoinbaseorder031921/download
47. Cook, C.: Introducing the blockstream satellite base station (2021)
48. Core, B.: Cve-2018-17144 full disclosure (2018), https://bitcoincore.org/en/2018/09/20/notice/
49. Coretti, S., Kiayias, A., Moore, C., Russell, A.: The generalsâ€™ scuttlebutt: Byzantine-resilient gossip protocols. IACR Cryptol. ePrint Arch. p. 541 (2022), https://eprint.iacr.org/2022/541
50. Daian, P., Pass, R., Shi, E.: Snow white: Robustly reconfigurable consensus and applications to provably secure proof of stake. In: Goldberg and Moore [81], pp. 23–41. https://doi.org/10.1007/978-3-030-32101-7_2
51. Daveas, S., Karantias, K., Kiayias, A., Zindros, D.: A gas-efficient superlight bitcoin client in solidity. Cryptology ePrint Archive, Report 2020/927 (2020), https://eprint.iacr.org/2020/927
52. David, B., Gazi, P., Kiayias, A., Russell, A.: Ouroboros praos: An adaptively-secure, semi-synchronous proof-of-stake blockchain. In: Nielsen, J.B., Rijmen, V. (eds.) Advances in Cryptology – EUROCRYPT 2018, Part II. Lecture Notes in Computer Science, vol. 10821, pp. 66–98. Springer, Heidelberg, Germany, Tel Aviv, Israel (Apr 29 – May 3, 2018). https://doi.org/10.1007/978-3-319-78375-8_3
53. Deb, S., Kannan, S., Tse, D.: Posat: Proof-of-work availability and unpredictability, without the work. In: Borisov, N., Díaz, C. (eds.) Financial Cryptography and Data Security - 25th International Conference, FC 2021, Virtual Event, March 1-5, 2021, Revised Selected Papers, Part II. Lecture Notes in Computer Science, vol. 12675, pp. 104–128. Springer (2021). https://doi.org/10.1007/978-3-662-64331-0_6, https://doi.org/10.1007/978-3-662-64331-0_6
54. Decker, C., Wattenhofer, R.: Information propagation in the bitcoin network. In: 13th IEEE International Conference on Peer-to-Peer Computing, IEEE P2P 2013,

Trento, Italy, September 9-11, 2013, Proceedings. pp. 1–10. IEEE (2013). https://doi.org/10.1109/P2P.2013.6688704, https://doi.org/10.1109/P2P.2013.6688704
55. Delgado-Segura, S., Bakshi, S., Pérez-Solà, C., Litton, J., Pachulski, A., Miller, A., Bhattacharjee, B.: TxProbe: Discovering bitcoin's network topology using orphan transactions. In: Goldberg and Moore [81], pp. 550–566. https://doi.org/10.1007/978-3-030-32101-7_32
56. Dfinity: Node provider machine hardware guide - internet computer wiki, https://wiki.internetcomputer.org/wiki/Node_Provider_Machine_Hardware_Guide
57. Dickey, C.G., Grothoff, C.: Bootstrapping of peer-to-peer networks. In: 2008 International Symposium on Applications and the Internet. pp. 205–208. IEEE (2008)
58. Donet, J.A.D., Pérez-Solà, C., Herrera-Joancomartí, J.: The bitcoin P2P network. In: Böhme, R., Brenner, M., Moore, T., Smith, M. (eds.) Financial Cryptography and Data Security - FC 2014 Workshops, BITCOIN and WAHC 2014, Christ Church, Barbados, March 7, 2014, Revised Selected Papers. Lecture Notes in Computer Science, vol. 8438, pp. 87–102. Springer (2014). https://doi.org/10.1007/978-3-662-44774-1_7, https://doi.org/10.1007/978-3-662-44774-1_7
59. Dotan, M., Pignolet, Y.A., Schmid, S., Tochner, S., Zohar, A.: Survey on blockchain networking: Context, state-of-the-art, challenges. ACM Comput. Surv. **54**(5) (may 2021). https://doi.org/10.1145/3453161, https://doi.org/10.1145/3453161
60. Downs, A., et al.: An economic theory of democracy (1957)
61. Durumeric, Z., Wustrow, E., Halderman, J.A.: ZMap: Fast internet-wide scanning and its security applications. In: King, S.T. (ed.) USENIX Security 2013: 22nd USENIX Security Symposium. pp. 605–620. USENIX Association, Washington, DC, USA (Aug 14–16, 2013)
62. Ekblaw, A., Barabas, C., Harvey-Buschel, J., Lippman, A.: Bitcoin and the myth of decentralization: Socio-technical proposals for restoring network integrity. In: 2016 IEEE 1st international workshops on foundations and applications of self* systems (FAS* W). pp. 18–23. IEEE (2016)
63. Erdos, P., Rényi, A., et al.: On the evolution of random graphs. Publ. Math. Inst. Hung. Acad. Sci **5**(1), 17–60 (1960)
64. Ethereum: Ethash (2022), https://ethereum.org/en/developers/docs/consensus-mechanisms/pow/mining-algorithms/ethash/
65. Ethereum: The merge (2022), https://ethereum.org/en/upgrades/merge
66. of the European Union, C.: Proposal for a regulation of the european parliament and of the council on markets in crypto-assets, and amending directive (eu) 2019/1937 (mica) (2022), https://data.consilium.europa.eu/doc/document/ST-13198-2022-INIT/en/pdf
67. Eyal, I., Sirer, E.G.: Majority is not enough: bitcoin mining is vulnerable. Commun. ACM **61**(7), 95–102 (2018). https://doi.org/10.1145/3212998, https://doi.org/10.1145/3212998
68. Feld, S., Schönfeld, M., Werner, M.: Analyzing the deployment of bitcoin's P2P network under an as-level perspective. In: Shakshuki, E.M., Yasar, A. (eds.) Proceedings of the 5th International Conference on Ambient Systems, Networks and Technologies (ANT 2014), the 4th International Conference on Sustainable Energy Information Technology (SEIT-2014), Hasselt, Belgium, June 2-5, 2014. Procedia Computer Science, vol. 32, pp. 1121–1126. Elsevier (2014). https://doi.org/10.1016/j.procs.2014.05.542, https://doi.org/10.1016/j.procs.2014.05.542

69. Franceschi-Bicchierai, L.: People can't see some nfts on twitter, crypto wallets after opensea goes down (2022), https://www.vice.com/en/article/g5qjej/people-cant-see-some-nfts-in-crypto-wallets-after-opensea-goes-down
70. Franzoni, F., Abellan, I., Daza, V.: Leveraging bitcoin testnet for bidirectional botnet command and control systems. In: Bonneau and Heninger [28], pp. 3–19. https://doi.org/10.1007/978-3-030-51280-4_1
71. Franzoni, F., Salleras, X., Daza, V.: Atom: Active topology monitoring for the bitcoin peer-to-peer network. Peer-to-Peer Networking and Applications **15**(1), 408–425 (2022)
72. Fuller, B., Khan, J.: Cve-2018-17145: Bitcoin inventory out-of-memory denial-of-service attack (2020), https://invdos.net/paper/CVE-2018-17145.pdf
73. Gao, Y., Shi, J., Wang, X., Tan, Q., Zhao, C., Yin, Z.: Topology measurement and analysis on ethereum P2P network. In: 2019 IEEE Symposium on Computers and Communications, ISCC 2019, Barcelona, Spain, June 29 - July 3, 2019. pp. 1–7. IEEE (2019). https://doi.org/10.1109/ISCC47284.2019.8969695, https://doi.org/10.1109/ISCC47284.2019.8969695
74. Garay, J.A., Kiayias, A.: SoK: A consensus taxonomy in the blockchain era. In: Jarecki [91], pp. 284–318. https://doi.org/10.1007/978-3-030-40186-3_13
75. Garay, J.A., Kiayias, A., Leonardos, N.: The bitcoin backbone protocol: Analysis and applications. In: Oswald, E., Fischlin, M. (eds.) Advances in Cryptology – EUROCRYPT 2015, Part II. Lecture Notes in Computer Science, vol. 9057, pp. 281–310. Springer, Heidelberg, Germany, Sofia, Bulgaria (Apr 26–30, 2015). https://doi.org/10.1007/978-3-662-46803-6_10
76. Garay, J.A., Kiayias, A., Leonardos, N.: The bitcoin backbone protocol with chains of variable difficulty. In: Katz and Shacham [100], pp. 291–323. https://doi.org/10.1007/978-3-319-63688-7_10
77. Garay, J.A., Kiayias, A., Panagiotakos, G.: Consensus from signatures of work. In: Jarecki [91], pp. 319–344. https://doi.org/10.1007/978-3-030-40186-3_14
78. Gaži, P., Kiayias, A., Russell, A.: Stake-bleeding attacks on proof-of-stake blockchains. Cryptology ePrint Archive, Report 2018/248 (2018), https://eprint.iacr.org/2018/248
79. Gencer, A.E., Basu, S., Eyal, I., van Renesse, R., Sirer, E.G.: Decentralization in bitcoin and Ethereum networks. In: Meiklejohn, S., Sako, K. (eds.) FC 2018: 22nd International Conference on Financial Cryptography and Data Security. Lecture Notes in Computer Science, vol. 10957, pp. 439–457. Springer, Heidelberg, Germany, Nieuwpoort, Curaçao (Feb 26 – Mar 2, 2018). https://doi.org/10.1007/978-3-662-58387-6_24
80. Gervais, A., Karame, G.O., Capkun, V., Capkun, S.: Is bitcoin a decentralized currency? IEEE Secur. Priv. **12**(3), 54–60 (2014). https://doi.org/10.1109/MSP.2014.49, https://doi.org/10.1109/MSP.2014.49
81. Goldberg, I., Moore, T. (eds.): FC 2019: 23rd International Conference on Financial Cryptography and Data Security, Lecture Notes in Computer Science, vol. 11598. Springer, Heidelberg, Germany, Frigate Bay, St. Kitts and Nevis (Feb 18–22, 2019)
82. Goodkind, N.: World's richest people now own 11% of global wealth, marking the biggest leap in recent history (2021), https://fortune.com/2021/12/07/worlds-richest-inequality-richer-during-pandemic/
83. Hajdarbegovic, N.: Bitcoin miners ditch ghash.io pool over fears of 51% attack (2014), https://www.coindesk.com/markets/2014/01/09/bitcoin-miners-ditch-ghashio-pool-over-fears-of-51-attack/

84. Heilman, E., Kendler, A., Zohar, A., Goldberg, S.: Eclipse attacks on bitcoin's peer-to-peer network. In: Jung, J., Holz, T. (eds.) USENIX Security 2015: 24th USENIX Security Symposium. pp. 129–144. USENIX Association, Washington, DC, USA (Aug 12–14, 2015)
85. Heimbach, L., Kiffer, L., Torres, C.F., Wattenhofer, R.: Ethereum's proposer-builder separation: Promises and realities. arXiv preprint arXiv:2305.19037 (2023)
86. Hern, A.: $300m in cryptocurrency accidentally lost forever due to bug (2017), https://www.theguardian.com/technology/2017/nov/08/cryptocurrency-300m-dollars-stolen-bug-ether
87. Hinman, W.: Digital asset transactions: When howey met gary (plastic) (2018), https://www.sec.gov/news/speech/speech-hinman-061418
88. 2020 IEEE Symposium on Security and Privacy. IEEE Computer Society Press, San Francisco, CA, USA (May 18–21, 2020)
89. in, V.: On stake (2014)
90. IO, E.: Eos. io technical white paper v2. EOS, Tech. Rep (2018)
91. Jarecki, S. (ed.): Topics in Cryptology – CT-RSA 2020, Lecture Notes in Computer Science, vol. 12006. Springer, Heidelberg, Germany, San Francisco, CA, USA (Feb 24–28, 2020)
92. Johnson, N.L.: Diversity in decentralized systems: Enabling self-organizing solutions. In: Decentralization II Conference, UCLA. vol. 4. Citeseer (1999)
93. Jourenko, M., Kurazumi, K., Larangeira, M., Tanaka, K.: SoK: A taxonomy for layer-2 scalability related protocols for cryptocurrencies. Cryptology ePrint Archive, Report 2019/352 (2019), https://eprint.iacr.org/2019/352
94. Kanjalkar, S., Kuo, J., Li, Y., Miller, A.: Short paper: I can't believe it's not stake! Resource exhaustion attacks on PoS. In: Goldberg and Moore [81], pp. 62–69. https://doi.org/10.1007/978-3-030-32101-7_4
95. Kappos, G., Yousaf, H., Maller, M., Meiklejohn, S.: An empirical analysis of anonymity in zcash. In: Enck, W., Felt, A.P. (eds.) USENIX Security 2018: 27th USENIX Security Symposium. pp. 463–477. USENIX Association, Baltimore, MD, USA (Aug 15–17, 2018)
96. Karakostas, D., Karayannidis, N., Kiayias, A.: Efficient state management in distributed ledgers. Cryptology ePrint Archive, Report 2021/183 (2021), https://eprint.iacr.org/2021/183
97. Karakostas, D., Kiayias, A., Nasikas, C., Zindros, D.: Cryptocurrency egalitarianism: A quantitative approach. In: Danos, V., Herlihy, M., Potop-Butucaru, M., Prat, J., Piergiovanni, S.T. (eds.) International Conference on Blockchain Economics, Security and Protocols, Tokenomics 2019, May 6-7, 2019, Paris, France. OASIcs, vol. 71, pp. 7:1–7:21. Schloss Dagstuhl - Leibniz-Zentrum für Informatik (2019). https://doi.org/10.4230/OASIcs.Tokenomics.2019.7, https://doi.org/10.4230/OASIcs.Tokenomics.2019.7
98. Karantias, K.: Sok: A taxonomy of cryptocurrency wallets. Cryptology ePrint Archive (2020)
99. Kattis, A., Bonneau, J.: Proof of necessary work: Succinct state verification with fairness guarantees. Cryptology ePrint Archive, Report 2020/190 (2020), https://eprint.iacr.org/2020/190
100. Katz, J., Shacham, H. (eds.): Advances in Cryptology – CRYPTO 2017, Part I, Lecture Notes in Computer Science, vol. 10401. Springer, Heidelberg, Germany, Santa Barbara, CA, USA (Aug 20–24, 2017)
101. Kerber, T., Kiayias, A., Kohlweiss, M.: Mining for privacy: How to bootstrap a snarky blockchain. In: Borisov, N., Díaz, C. (eds.) Financial Cryptography and

Data Security - 25th International Conference, FC 2021, Virtual Event, March 1-5, 2021, Revised Selected Papers, Part I. Lecture Notes in Computer Science, vol. 12674, pp. 497–514. Springer (2021). https://doi.org/10.1007/978-3-662-64322-8_24, https://doi.org/10.1007/978-3-662-64322-8_24
102. Kiayias, A., Lazos, P.: Sok: Blockchain governance. CoRR **abs/2201.07188** (2022), https://arxiv.org/abs/2201.07188
103. Kiayias, A., Miller, A., Zindros, D.: Non-interactive proofs of proof-of-work. Cryptology ePrint Archive, Report 2017/963 (2017), https://eprint.iacr.org/2017/963
104. Kiayias, A., Polydouri, A., Zindros, D.: The velvet path to superlight blockchain clients. Cryptology ePrint Archive, Report 2020/1122 (2020), https://eprint.iacr.org/2020/1122
105. Kiayias, A., Russell, A., David, B., Oliynykov, R.: Ouroboros: A provably secure proof-of-stake blockchain protocol. In: Katz and Shacham [100], pp. 357–388. https://doi.org/10.1007/978-3-319-63688-7_12
106. Kiayias, A., Stouka, A.: Coalition-safe equilibria with virtual payoffs. In: Baldimtsi, F., Roughgarden, T. (eds.) AFT '21: 3rd ACM Conference on Advances in Financial Technologies, Arlington, Virginia, USA, September 26 - 28, 2021. pp. 71–85. ACM (2021). https://doi.org/10.1145/3479722.3480795, https://doi.org/10.1145/3479722.3480795
107. Kohler, S., Pizzol, M.: Life cycle assessment of bitcoin mining. Environmental science & technology **53**(23), 13598–13606 (2019)
108. Kohlweiss, M., Maller, M., Siim, J., Volkhov, M.: Snarky ceremonies. In: Tibouchi, M., Wang, H. (eds.) Advances in Cryptology – ASIACRYPT 2021, Part III. Lecture Notes in Computer Science, vol. 13092, pp. 98–127. Springer, Heidelberg, Germany, Singapore (Dec 6–10, 2021). https://doi.org/10.1007/978-3-030-92078-4_4
109. Konstantinov, E.: The ethereum cloud vs. on-premises nodes conundrum (2019), https://chainstack.com/the-ethereum-cloud-vs-on-premises-nodes-conundrum
110. Koutsoupias, E., Papadimitriou, C.: Worst-case equilibria. In: Annual symposium on theoretical aspects of computer science. pp. 404–413. Springer (1999)
111. Kroll, J.A., Davey, I.C., Felten, E.W.: The economics of bitcoin mining, or bitcoin in the presence of adversaries. In: Proceedings of WEIS. vol. 2013. Washington, DC (2013)
112. Kwon, J., Buchman, E.: Cosmos whitepaper. A Netw. Distrib. Ledgers **27** (2019)
113. Lamport, L., Shostak, R., Pease, M.: The byzantine generals problem. ACM Transactions on Programming Languages and Systems (TOPLAS) **4**(3), 382–401 (1982)
114. Leiba, O.: Unveiling bigspender: Double (and multiple) spend vulnerability in major bitcoin wallets (2020), https://zengo.com/bigspender-double-spend-vulnerability-in-bitcoin-wallets/
115. Li, C., Palanisamy, B.: Comparison of decentralization in dpos and pow blockchains. CoRR **abs/2002.02082** (2020), https://arxiv.org/abs/2002.02082
116. Li, W., Andreina, S., Bohli, J.M., Karame, G.: Securing proof-of-stake blockchain protocols. In: Data Privacy Management, Cryptocurrencies and Blockchain Technology, pp. 297–315. Springer (2017)
117. Lin, Q., Li, C., Zhao, X., Chen, X.: Measuring decentralization in bitcoin and ethereum using multiple metrics and granularities. In: 2021 IEEE 37th Interna-

tional Conference on Data Engineering Workshops (ICDEW). pp. 80–87. IEEE (2021)
118. Loe, A.F., Quaglia, E.A.: You shall not join: A measurement study of cryptocurrency peer-to-peer bootstrapping techniques. In: Cavallaro et al. [36], pp. 2231–2247. https://doi.org/10.1145/3319535.3345649
119. Magaki, I., Khazraee, M., Gutierrez, L.V., Taylor, M.B.: Asic clouds: Specializing the datacenter. ACM SIGARCH Computer Architecture News **44**(3), 178–190 (2016)
120. Management, B.A.: Analysis of real bitcoin trade volume (2019), https://www.sec.gov/comments/sr-nysearca-2019-01/srnysearca201901-5164833-183434.pdf
121. Mariem, S.B., Casas, P., Romiti, M., Donnet, B., Stütz, R., Haslhofer, B.: All that glitters is not bitcoin–unveiling the centralized nature of the btc (ip) network. In: NOMS 2020-2020 IEEE/IFIP Network Operations and Management Symposium. pp. 1–9. IEEE (2020)
122. Marlinspike, M.: My first impressions of web3. Moxie Marlinspike's Website (2022), https://moxie.org/2022/01/07/web3-first-impressions.html
123. Matt, C., Nielsen, J.B., Thomsen, S.E.: Formalizing delayed adaptive corruptions and the security of flooding networks. In: Dodis, Y., Shrimpton, T. (eds.) Advances in Cryptology - CRYPTO 2022 - 42nd Annual International Cryptology Conference, CRYPTO 2022, Santa Barbara, CA, USA, August 15-18, 2022, Proceedings, Part II. Lecture Notes in Computer Science, vol. 13508, pp. 400–430. Springer (2022). https://doi.org/10.1007/978-3-031-15979-4_14, https://doi.org/10.1007/978-3-031-15979-4_14
124. Maymounkov, P., Mazieres, D.: Kademlia: A peer-to-peer information system based on the xor metric. In: International Workshop on Peer-to-Peer Systems. pp. 53–65. Springer (2002)
125. Meiklejohn, S., Pomarole, M., Jordan, G., Levchenko, K., McCoy, D., Voelker, G.M., Savage, S.: A fistful of bitcoins: characterizing payments among men with no names. Commun. ACM **59**(4), 86–93 (2016). https://doi.org/10.1145/2896384, https://doi.org/10.1145/2896384
126. Milewski, K.: Cftc, fincen, and sec warn of crypto aml enforcement (2019), https://www.natlawreview.com/article/cftc-fincen-and-sec-warn-crypto-aml-enforcement
127. Monero: Network upgrade and release 0.15 (2019), https://www.getmonero.org/2019/10/01/announcement-release-0-15.html
128. Moore, T., Christin, N.: Beware the middleman: Empirical analysis of bitcoin-exchange risk. In: Financial Cryptography and Data Security: 17th International Conference, FC 2013, Okinawa, Japan, April 1-5, 2013, Revised Selected Papers 17. pp. 25–33. Springer (2013)
129. Motamed, A.P., Bahrak, B.: Quantitative analysis of cryptocurrencies transaction graph. Applied Network Science **4**(1), 1–21 (2019)
130. Nakamoto, S.: Bitcoin: A peer-to-peer electronic cash system (2008)
131. Naumenko, G., Maxwell, G., Wuille, P., Fedorova, A., Beschastnikh, I.: Erlay: Efficient transaction relay for bitcoin. In: Cavallaro et al. [36], pp. 817–831. https://doi.org/10.1145/3319535.3354237
132. Neudecker, T., Andelfinger, P., Hartenstein, H.: Timing analysis for inferring the topology of the bitcoin peer-to-peer network. In: 2016 Intl IEEE Conferences on Ubiquitous Intelligence & Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Cloud and

Big Data Computing, Internet of People, and Smart World Congress (UIC/ATC/ScalCom/CBDCom/IoP/SmartWorld). pp. 358–367. IEEE (2016)
133. Neudecker, T., Hartenstein, H.: Network layer aspects of permissionless blockchains. IEEE Communications Surveys & Tutorials **21**(1), 838–857 (2018)
134. Newman, M.E.: The structure and function of complex networks. SIAM review **45**(2), 167–256 (2003)
135. Nikolaenko, V., Ragsdale, S., Bonneau, J., Boneh, D.: Powers-of-tau to the people: Decentralizing setup ceremonies. Cryptology ePrint Archive (2022)
136. Nomics: Nomics' crypto trust index - q1 2021, north america (2021), `https://nomics.com/blog/essays/q1-crypto-trust-index`
137. O'Leary, R.R.: The race is on to replace ethereum's most centralized layer (2018), `https://www.coindesk.com/markets/2018/12/05/the-race-is-on-to-replace-ethereums-most-centralized-layer/`
138. Pass, R., Seeman, L., shelat, a.: Analysis of the blockchain protocol in asynchronous networks. In: Coron, J.S., Nielsen, J.B. (eds.) Advances in Cryptology – EUROCRYPT 2017, Part II. Lecture Notes in Computer Science, vol. 10211, pp. 643–673. Springer, Heidelberg, Germany, Paris, France (Apr 30 – May 4, 2017). `https://doi.org/10.1007/978-3-319-56614-6_22`
139. Pease, M., Shostak, R., Lamport, L.: Reaching agreement in the presence of faults. Journal of the ACM (JACM) **27**(2), 228–234 (1980)
140. Peck, M.: The crazy security behind the birth of zcash, the inside story (2016), `https://spectrum.ieee.org/the-crazy-security-behind-the-birth-of-zcash`
141. Pelt, R.v., Jansen, S., Baars, D., Overbeek, S.: Defining blockchain governance: a framework for analysis and comparison. Information Systems Management **38**(1), 21–41 (2021)
142. Peters, K.: A history of bitcoin hard forks (2021), `https://www.investopedia.com/tech/history-bitcoin-hard-forks/`
143. Platis, M., Sanderford, B.: Nakamoto coefficient (2021), `https://crosstower.com/resources/education/nakamoto-coefficient`
144. Poon, J., Dryja, T.: The bitcoin lightning network: Scalable off-chain instant payments (2016)
145. Reed, D.D., Luiselli, J.K.: Temporal Discounting, pp. 1474–1474. Springer US, Boston, MA (2011). `https://doi.org/10.1007/978-0-387-79061-9_3162`, `https://doi.org/10.1007/978-0-387-79061-9_3162`
146. Reibel, P., Yousaf, H., Meiklejohn, S.: Short paper: An exploration of code diversity in the cryptocurrency landscape. In: Goldberg and Moore [81], pp. 73–83. `https://doi.org/10.1007/978-3-030-32101-7_5`
147. Reiff, N.: All about the bitcoin cash hard fork (2022), `https://www.investopedia.com/news/all-about-bitcoin-cash-hard-fork/`
148. Ren, L., Devadas, S.: Bandwidth hard functions for ASIC resistance. In: Kalai, Y., Reyzin, L. (eds.) TCC 2017: 15th Theory of Cryptography Conference, Part I. Lecture Notes in Computer Science, vol. 10677, pp. 466–492. Springer, Heidelberg, Germany, Baltimore, MD, USA (Nov 12–15, 2017). `https://doi.org/10.1007/978-3-319-70500-2_16`
149. Rényi, A., et al.: On measures of entropy and information. In: Proceedings of the fourth Berkeley symposium on mathematical statistics and probability. vol. 1. Berkeley, California, USA (1961)
150. Rhoades, S.A.: The herfindahl-hirschman index. Fed. Res. Bull. **79**, 188 (1993)

151. Romiti, M., Judmayer, A., Zamyatin, A., Haslhofer, B.: A deep dive into bitcoin mining pools: An empirical analysis of mining shares. CoRR **abs/1905.05999** (2019), http://arxiv.org/abs/1905.05999
152. Ron, D., Shamir, A.: Quantitative analysis of the full Bitcoin transaction graph. In: Sadeghi, A.R. (ed.) FC 2013: 17th International Conference on Financial Cryptography and Data Security. Lecture Notes in Computer Science, vol. 7859, pp. 6–24. Springer, Heidelberg, Germany, Okinawa, Japan (Apr 1–5, 2013). https://doi.org/10.1007/978-3-642-39884-1_2
153. Ron, D., Shamir, A.: Quantitative analysis of the full bitcoin transaction graph. In: Sadeghi, A. (ed.) Financial Cryptography and Data Security - 17th International Conference, FC 2013, Okinawa, Japan, April 1-5, 2013, Revised Selected Papers. Lecture Notes in Computer Science, vol. 7859, pp. 6–24. Springer (2013). https://doi.org/10.1007/978-3-642-39884-1_2, https://doi.org/10.1007/978-3-642-39884-1_2
154. Rosenthal, D.: Gini coefficients of cryptocurrencies (2018), https://blog.dshr.org/2018/10/gini-coefficients-of-cryptocurrencies.html
155. Roubini, N.: The big blockchain lie (2018), https://www.project-syndicate.org/commentary/blockchain-big-lie-by-nouriel-roubini-2018-10
156. Sai, A.R., Buckley, J., Fitzgerald, B., Gear, A.L.: Taxonomy of centralization in public blockchain systems: A systematic literature review. Inf. Process. Manag. **58**(4), 102584 (2021). https://doi.org/10.1016/j.ipm.2021.102584, https://doi.org/10.1016/j.ipm.2021.102584
157. Sai, A.R., Buckley, J., Gear, A.L.: Characterizing wealth inequality in cryptocurrencies. Frontiers Blockchain **4**, 730122 (2021). https://doi.org/10.3389/fbloc.2021.730122, https://doi.org/10.3389/fbloc.2021.730122
158. Schneider, N.: Decentralization: an incomplete ambition. Journal of cultural economy **12**(4), 265–285 (2019)
159. Securities, E., Authority, M.: Initial coin offerings and crypto-assets. Innovation and Products **ESMA50-157-1391** (2019), https://www.esma.europa.eu/sites/default/files/library/esma50-157-1391_crypto_advice.pdf
160. Securities, U., Commission, E., et al.: Report of investigation pursuant to section 21(a) of the securities exchange act of 1934: The dao (2017), https://www.sec.gov/litigation/investreport/34-81207.pdf
161. Securities, U., Commission, E., et al.: Sec issues investigative report concluding dao tokens, a digital asset, were securities. July **25**, 2017–131 (2017), https://www.sec.gov/news/press-release/2017-131
162. Securities, U., Commission, E., et al.: Framework for "investment contract" analysis of digital assets. US Securities and Exchange Commission **3** (2019)
163. Shannon, C.E.: A mathematical theory of communication. ACM SIGMOBILE mobile computing and communications review **5**(1), 3–55 (2001)
164. Stoll, C., Klaaßen, L., Gallersdörfer, U.: The carbon footprint of bitcoin. Joule **3**(7), 1647–1661 (2019)
165. Sun, W., Jin, H., Jin, F., Kong, L., Peng, Y., Dai, Z.: Spatial analysis of global bitcoin mining. Scientific Reports **12**(1), 10694 (2022)
166. Tanenbaum, A.S., Wetherall, D.: Computer networks, 5th Edition. Pearson (2011), https://www.worldcat.org/oclc/698581231
167. Taylor, M.B.: The evolution of bitcoin hardware. Computer **50**(9), 58–66 (2017)
168. Theil, H.: Economics and information theory. Tech. rep. (1967)
169. Tosh, D.K., Shetty, S., Liang, X., Kamhoua, C.A., Kwiat, K.A., Njilla, L.: Security implications of blockchain cloud with analysis of block withholding attack.

In: 2017 17th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGRID). pp. 458–467. IEEE (2017)
170. Tran, M., Choi, I., Moon, G.J., Vu, A.V., Kang, M.S.: A stealthier partitioning attack against bitcoin peer-to-peer network. In: IEEE S&P 2020 [88], pp. 894–909. https://doi.org/10.1109/SP40000.2020.00027
171. Tran, M., Shenoi, A., Kang, M.S.: On the routing-aware peering against network-eclipse attacks in bitcoin. In: Bailey, M., Greenstadt, R. (eds.) USENIX Security 2021: 30th USENIX Security Symposium. pp. 1253–1270. USENIX Association (Aug 11–13, 2021)
172. UNDP, U.: Decentralization: A sampling of definitions. Working Paper (1999)
173. Valdivia, L.J., Del-Valle-Soto, C., Rodriguez, J., Alcaraz, M.: Decentralization: The failed promise of cryptocurrencies. IT Professional **21**(2), 33–40 (2019)
174. Ventures, O.: Blockchain development trends 2021 (2021), https://outlierventures.io/wp-content/uploads/2021/02/OV-Blockchain-Dev-Q1-2021-_v7.pdf
175. Verified, B.: 2020 market data integrity report (2020), https://web.archive.org/web/20211009155412/https://btiverified.com/crypto-market-data-report-2020/
176. Victor, F., Ruppel, P., Kupper, A.: A taxonomy for distributed ledger analytics. Computer **54**(02), 30–38 (2021)
177. Wang, C., Chu, X., Qin, Y.: Measurement and analysis of the bitcoin networks: A view from mining pools. In: 2020 6th International Conference on Big Data Computing and Communications (BIGCOM). pp. 180–188 (2020). https://doi.org/10.1109/BigCom51056.2020.00032
178. Wang, T., Zhao, C., Yang, Q., Zhang, S., Liew, S.C.: Ethna: Analyzing the underlying peer-to-peer network of ethereum blockchain. IEEE Transactions on Network Science and Engineering **8**(3), 2131–2146 (2021)
179. Werner, S.M., Perez, D., Gudgeon, L., Klages-Mundt, A., Harz, D., Knottenbelt, W.J.: Sok: Decentralized finance (defi). CoRR **abs/2101.08778** (2021), https://arxiv.org/abs/2101.08778
180. Williams, C.: Infura outage sparks debate over ethereum's decentralization (2020), https://cryptobriefing.com/infura-outage-sparks-debate-over-ethereums-decentralization/
181. Wüst, K., Gervais, A.: Ethereum eclipse attacks. Tech. rep., ETH Zurich (2016)
182. Young, M.: Opensea's team of 37 staff is currently handling 98% of combined nft volumes (2021), https://cointelegraph.com/news/opensea-s-team-of-37-staff-are-currently-handling-98-of-combined-nft-volumes
183. Zeng, L., Chen, Y., Chen, S., Zhang, X., Guo, Z., Xu, W., Moscibroda, T.: Characterizing ethereum's mining power decentralization at a deeper level. In: 40th IEEE Conference on Computer Communications, INFOCOM 2021, Vancouver, BC, Canada, May 10-13, 2021. pp. 1–10. IEEE (2021). https://doi.org/10.1109/INFOCOM42981.2021.9488812, https://doi.org/10.1109/INFOCOM42981.2021.9488812
184. Zhang, B., Oliynykov, R., Balogun, H.: A treasury system for cryptocurrencies: Enabling better collaborative intelligence. In: ISOC Network and Distributed System Security Symposium – NDSS 2019. The Internet Society, San Diego, CA, USA (Feb 24-27, 2019)
185. Zhang, L., Ma, X., Liu, Y.: Sok: Blockchain decentralization. arXiv preprint arXiv:2205.04256 (2022)

186. Zhou, F., Song, N.: 2020 q2 cryptocurrency mining industry report (2020), https://medium.com/coinmonks/2020-q2-cryptocurrency-mining-industry-report-50a908773b17
187. Zhou, L., Xiong, X., Ernstberger, J., Chaliasos, S., Wang, Z., Wang, Y., Qin, K., Wattenhofer, R., Song, D., Gervais, A.: Sok: Decentralized finance (defi) attacks. In: 44th IEEE Symposium on Security and Privacy, SP 2023, San Francisco, CA, USA, May 21-25, 2023. pp. 2444–2461. IEEE (2023). https://doi.org/10.1109/SP46215.2023.10179435, https://doi.org/10.1109/SP46215.2023.10179435
188. Zhou, Q.: Irregular-program-based hash algorithms. In: 2019 IEEE International Conference on Decentralized Applications and Infrastructures (DAPPCON). pp. 125–128. IEEE (2019)

A Decentralization and Policymaking

In the US, an early policymaking decision was the SEC's ruling of "DAO" tokens as securities [161]. It argued that the DAO was predominantly controlled by its creators, who handpicked parties with decisive operational capabilities. In 2018, it was posited that characterization as a security depended on whether the token's network is sufficiently decentralized [87]. This was highlighted in SEC's guidance [162], where the existence of "active participants" (APs) that undertake essential responsibilities was named as a deciding factor for the classification of a digital asset as a security — particularly an "investment contract." Our methodology, which measures decentralization as the distance from single points of failure in different strata, could thus aid law enforcement in making such decisions. In essence, an AP, as described by the SEC, is also a single point of failure from a (cyber-) security perspective so, if a system is centralized under our methodology, the sale of its digital asset is possibly constituting an investment contract, as opposed to a system proclaimed to be decentralized. Our minimum decentralization test (Definition 1) formalizes this perspective and could serve as a litmus test.

In the UK, the Financial Stability Board (FSB) identified the possibility of power concentration around a small set of parties, when it comes to ownership and operation of key infrastructure, as a major concern in the risks of the application of decentralized technologies [24]. Nonetheless, it also warned that decentralization in conjunction with inadequate governance "makes it difficult to resolve technological limitations or errors and may lead to uncertainty" [23].

Decentralization is also a matter of interest in the European Union. EU member states have suggested that legislation should take into account the decentralized nature of the technology on which various businesses operate [159]. In late 2022, the Markets in Crypto-assets regulation (MiCA) was approved by the EU council and the Parliament Committee on Economic and Monetary Affairs [66].[26] This regulation makes a specific mention of decentralization as *the* distinguishing factor on whether a system falls within its scope. Interestingly,

[26] A final vote in a full parliament session is expected by the end of 2022.

systems might fall under this regulation even if some parts of them are decentralized, but they are not fully decentralized, i.e., across all relevant strata.[27]

B Software Testing

Testing is a core part of software development. In blockchains, a major means of testing new applications or ledger features is testnets. A testnet is a separate chain, identical to the main chain in terms of offered functionalities. To transact, a user acquires testnet tokens for free, so the native testnet tokens have no real-world value. Testnets offer multiple functionalities. Users can test features without risking losing funds. Developers test new features and applications in a scale that closely resembles the main chain. Adversaries evaluate the efficacy of attacks [37] or exploit the zero-cost nature of testnet transactions [70]. Therefore, testnets indirectly safeguard all ledger properties. Fewer testnets increase centralization around specific full node software products, while testnets maintained by diverse teams may collect richer data. Hence, the *resource* is testnets and the *relevant parties* are their operators.

Bitcoin offers a single primary testnet; the same holds for alternative cryptocurrencies like Zcash and Monero.[28] In Ethereum, although seemingly multiple testnets exist, most are deprecated due to the system's transition to PoS, and only one of the recommended networks is expected to be maintained in the long term.[29] PoS testnets, e.g., in Cardano and Solana, are also highly centralized.[30]

C Brief Evaluations per Layer

In this section we provide brief evaluations of various systems for each subcategory of the layers covered in Sections 3 - 10. In doing so, we identify various questions that require further research across two broad axes. First, from a measurement perspective, many systems and dimensions lack pertinent data or, to make matters more interesting, it is unclear how to even conduct robust measurements for the data under question. Second, from a design perspective, a relevant thread of research would focus on enabling or incentivizing protocol designers to implement accurate data collection mechanisms as a part of the systems themselves.

[27] "This Regulation applies to natural, legal persons and other undertakings and the activities and services performed, provided or controlled, directly or indirectly, by them, including when part of such activity or services is performed in a decentralized way. Where crypto-asset services as defined in this Regulation are provided in a fully decentralised manner without any intermediary they do not fall within the scope of this Regulation." [66]

[28] Sources: Bitcoin Wiki, Zcash Docs, Monero Docs [August 2022]

[29] The deprecation of Ropsten, Rinkeby, Kiln and Kovan was announced in 2022 [Ethereum blog]. Goerli will be maintained in the long run, while the future of Sepolia is undecided. [ethereum.org; August 2022]

[30] Cardano and Solana testnets are run by Input Output and the Solana Foundation resp. [Cardano Testnets, Solana Docs; August 2022]

C.1 Hardware: Physical Hardware

Centralization around specialized hardware has been documented [62], although no academic research could be found on mining hardware usage in real-world systems. Interestingly, it is unclear how to even measure the usage of hardware equipment in PoW mining via public data, as well as how to develop PoW algorithms that promote diversity, thus future research could aim at answering these questions. Nonetheless, there exist some reports, though they often present conflicting assessments. In Bitcoin, between 2017−2019 a single mining hardware provider accounted for either $65-75\%$ [20] or $46-58\%$ [186] of the network's hashrate, with 98% of the market controlled by 4 firms.

C.2 Hardware: Virtual Hardware

A comprehensive evaluation of centralization in terms of hardware hosting, and how to incentivize hosting diversity is a promising thread of future research. Here, we consider two examples of highly-valued PoS systems, Solana and Avalanche.[31] Solana's validators predominantly operate cloud-based nodes; of the 1873 nodes, more than half are hosted in two services, with more than 50% and more than 66% of participating stake hosted by 3 and 5 providers respectively.[32] Avalanche observes similar concentration issues; 731 out of 1254 validators, who control 71.84% of all stake, are hosted by a single company.[33]

C.3 Software: Protocol Participation

The literature is lacking formal analyses on the usage of full node ledger software, so a rigorous evaluation of the dynamics in software development and usage could highlight various centralization tendencies. Various community and commercial projects do keep track of statistics though. In most systems, a single client software is predominantly used by the participating nodes in the network. In Bitcoin 99% use Bitcoin Core (aka Satoshi), in Ethereum 78% use geth, in Litecoin 95% use LitecoinCore, while systems like Zcash are completely centralized with all nodes using one software (MagicBean); a notable exception is Bitcoin Cash, where usage is split between BCH Unlimited (33%), Bitcoin Cash Node (51%), and Bitcoin ABC (12%).[34]

Some projects are actively managed by a wide network of developers, e.g., more than 200 contribute to Ethereum [174], while others are particularly centralized. As of 2018, 7% of all Bitcoin Core files were written by the same person, while 30% of all files had a single author. In Ethereum, these figures rise to 20% and 55% respectively [12]. Comments observe similar centralization patterns, with 8 (0.3%) and 18 (0.6%) people contributing half of all comments in Bitcoin and Ethereum respectively [12].

---

[31] Solana and Avalanche are #9 and #14 respectively w.r.t. market capitalization. [CoinMarketCap; August 2022]

[32] validators.app. [August 2022]

[33] Data obtained from avascan.info. [August 2022]

[34] Sources: blockchair, ethernodes [August 2022]

C.4 Software: Asset Management

As keys and addresses are wallet agnostic, it is impossible to identify if two addresses are generated by the same wallet implementation, unless it purposely reveals such information. Consequently, it is unclear how to evaluate the wallet market's diversity and how widespread wallet usage is from public data. To our knowledge, no rigorous investigation has been conducted on this topic, either analyzing historical data patterns or conducting usability studies.

C.5 Network: Topology

Bitcoin is notoriously vigilant in hiding its network topology [55,71]. Various works analyze it by inferring a node's neighborhood [22], timing analysis [132], or conflicting transaction propagation [55]. In 2014, it was found that more than half of Bitcoin nodes resided in 40 autonomous systems (ASs), with 30% in just 10 ASs [68]. In 2017, Bitcoin's and Ethereum's P2P networks observed similar sizes (3390 nodes for Bitcoin, 4302 for Ethereum). Bitcoin offered lower latency and higher bandwidth, with nodes being closer geographically and 56% of them hosted on dedicated hosting services (vs. 28% for Ethereum) [79]. In addition, 68% of the mining power was hosted on 10 transit networks, while 3 transit networks saw more than 60% of all connections [5]. In 2019, Ethereum's network presented a large degree of centralization around clusters, forming a "small world network" [73] with 10 cloud hosting providers accounting for 57% of all nodes and one hosting almost a quarter [109]. This was reaffirmed in 2020, as Ethereum messages could be sent to most nodes within 6 hops [178]. In 2020, Monero's topology also observed a high level of centralization, as 13.2% of nodes maintained 82.86% of all connections [35]. No analysis of PoS systems' networks could be found; given their non-reliance on specialized hardware and ease of relocation, a PoS-PoW comparison would be of interest.

Bitcoin, as the first blockchain system, has also seen multiple eclipse attacks and defenses [170,84,171]. Some works attempt to increase the number of connections without reaching prohibitive levels of bandwidth usage [131]. Ethereum was also found vulnerable to eclipse attacks that do not require monopolizing a node's connections, but relied on message propagation [181].

C.6 Network: Node Bootstrapping and Peer Discovery

Bitcoin Core defines 8 outgoing connections, selected randomly from a known list of identities, and up to 125 incoming [59]. When (re)joining the network, a node attempts to connect to previously-known identities and, if unsuccessful, employs a (hardcoded) list of DNS seeds. Other systems, like Ethereum and Cardano, employ more complex, DHT-based mechanisms [124] that require further analysis. Cardano is also an interesting implementation, as it assumes two node types: (a) core nodes that participate in consensus, and (b) relays that intermediate between core and edge nodes (e.g., wallets); in the default configuration relays are operated by only a small committee [59].

C.7 Consensus

In (game) theory, Bitcoin's resistance to centralization has been both supported [111,106] and refuted [67], depending on the economic model assumed for the participants' utilities. In practice, mining pools have been observed as early as 2013 [80]. Between $2016-2020$, pools created 98.6% of Bitcoin blocks [115], with 5 pools consistently contributing between $65-85\%$ of the eventual blocks and 25 controlling more than 94% of all hashing power [177]. Centralization has also been observed within mining pools. Between $2017-2018$, no entity controlled more than 21% [79] of hashing power, but three pools controlled a majority; within these pools, a few participants (≤ 20) received over 50% of rewards [151]. Miners often participate in multiple pools at the same time, a behavior also observed in Ethereum [183]. Although centralization around pools is high in (PoW-based) Ethereum (in 2019, 3 pools controlled a majority of mining power [117]), power within the pools is spread across hundreds of addresses [183], albeit some possibly owned by the same parties.

Some systems use a committee-based approach, as opposed to Bitcoin's open participation model. Here, at each time there exists a known designated party which proposes a block and a committee of participants that vote for it. The following are examples of such systems, where each employs their own consensus protocol and defines a different number of participants per epoch using an on-chain process:[35] i) Cosmos: 175; ii) Polkadot: 297; iii) EOS: 21; iv) Harmony: 800; v) NEAR: 100. In all these systems, well-known exchanges are among the top elected validators.[36] Interestingly, the stake controlled by the elected validators is mostly delegated, instead of self-owned. Also, organizations often control multiple validators, so the number of real actors is often even smaller than the nominal number of participants (nevertheless some systems, e.g., Polkadot, go to greater lengths to ensure the representative participation satisfies desirable properties such as proportionality, cf. [38]). Consequently, identifying the participation distribution among *real-world users* and the refreshment rate of the elected committee across multiple epochs is an interesting research question. Similarly for investigating all the desiderata of representative participation from a social choice perspective.

C.8 Cryptocurrency Economics: (Initial) Token Distribution

PoS systems like Cardano, NEO, and Algorand tried to reduce early-stage risks via a two-phase launch. At first, the ledger was controlled by either the core development company or foundation or a committee numbering a small number of entities. After token ownership was sufficiently distributed, participation opened widely to all stakeholders. Beyond the obvious issues in maintaining a permissioned database, the first phase typically takes years to conclude. Early

[35] Sources: hub.cosmos.network, wiki.polkadot.network, developers.eos.io, docs.harmony.one, near.org [August 2022]

[36] For example, Binance is a validator in all mentioned systems. [Cosmos, Polkadot, EOS, Harmony, NEAR; August 2022]

users often tend to either not participate or transfer their tokens to the few exchanges that support these new tokens [176]. Therefore, an interesting question is the relationship of the delay between launch and full decentralization and the diversity of early investors.

C.9 Cryptocurrency Economics: Token Ownership

Bitcoin's wealth ownership and transaction graph has been analyzed since at least 2012 [152]. Over time, it demonstrated a three-phase history of distinct (de)centralization patterns, where 100 addresses possess a high centralization degree of assets and wealth flow in the network [42,157]. Similar analyses exist for Ethereum [41], Zcash [95], and other cryptocurrencies [129].

As of 2022, cryptocurrency wealth concentration is particularly extreme (Table 2). To establish some context, the income Gini coefficient of the 10 lowest-performing countries ranges between $0.63 - 0.512$ [15]. Bitcoin has a Gini coefficient of 0.514, considering only the $10,000$ richest addresses, and a staggering 0.955 w.r.t. all addresses. In the arguably deeply unequal global real-world economy, the richest 0.01% of individuals ($520,000$ people) hold 11% of all wealth [82]. Bitcoin manages to beat that figure, with 100 addresses holding 14.01% of all tokens.

System	**Addresses**	**Top-100**	**Gini (10K)**	**Gini**
Bitcoin	42,943,534	14.01 %	0.5145	0.956
Ethereum	193,673,067	39.75 %	0.6757	0.978
Dogecoin	4,839,762	68.49 %	0.8297	0.986
Zcash	345,766	32.92 %	0.7796	0.974
Bitcoin Cash	17,123,166	28.4 %	0.672	0.97
Litecoin	5,442,751	36.8 %	0.6757	0.978
Ethereum Classic	511,491	41.37 %	0.8289	0.988

Table 2. Cryptocurrency wealth distribution: i) addresses that control assets; ii) percentage of wealth controlled by the top 100 wealthiest addresses; iii, iv) Gini coefficient of 10K wealthiest and all addresses resp. (Sources: Blockchain ETL, Google BigQuery, CoinCarp. [April 2022])

A complexity in measuring wealth decentralization in cryptocurrencies arises due to their pseudonymous (or even anonymous) nature. Specifically, the number of addresses often does not correspond to individual people or entities, cf. [153,125]. A user may control multiple addresses, e.g., each with a small balance. When interpreting the Gini coefficient, this artificially enlarges the population and possibly biases the results towards decentralization. In addition, an address's assets may be owned by many users (e.g., exchange addresses), which biases Gini towards centralization. Thus, developing tools to compute wealth inequality in blockchain systems, without sacrificing core features like anonymity and privacy, is a crucial problem for exploration.

C.10 Cryptocurrency Economics: Secondary Markets

Table 3 summarizes secondary blockchain market data across 121 exchanges. Many systems (Bitcoin, Ethereum, Litecoin, XRP) are traded on all but a few small exchanges. Tether is by far the most available, in terms of market pairs, and used, in terms of volume. Interestingly, for all systems, except perhaps Bitcoin, the majority of volume is not of the highest transparency.[37] This is consistent with reports that show market manipulation is endemic in cryptocurrency markets, with multiple cases of wash trading, fake trading volumes, and other fraudulent behavior [46,175,120]. Market transactions are primarily conducted in a handful of exchanges. By far the most used is Binance (20% of the total daily volume),[38] although Coinbase is the most recognized in North America [136].

System	**Exchanges**	**Pairs**	**Transparent Volume**
Tether	80	21691	$52.86B (44%)
Bitcoin	100	9622	$41.31B (56%)
Ethereum	103	5680	$14.24B (42%)
Dogecoin	84	1856	$175.5M (29%)
Cardano	78	437	$1.03B (41%)
Solana	75	383	$1.25B (39%)
Avalanche	64	318	$574.29M (49%)
Monero	41	101	$39.9M (22%)

Table 3. Secondary market data across 121 exchanges (June 2022): i) exchanges and trading pairs (*CoinMarketCap*); ii) transaction volume rated as "transparent" by *nomics*.

C.11 Client API

In Bitcoin, most wallets are either SPV or explorer-based [98]. In the first case, the wallet obtains the chain's headers and, to verify that a transaction is pub-

[37] For "transparency" see the methodology and data of Nomics: `https://nomics.com/blog/essays/transparency-ratings`.

[38] CoinMarketCap [August 2022]

lished, requests a proof from full nodes. Although SPV does mitigate safety attacks, it also hurts the user's privacy, as their transaction information is leaked to the full node operators. Explorer-based wallets instead rely entirely on a single explorer service and its full nodes, which are trusted completely. In 2018, $5-10\%$ of all Ethereum nodes reportedly relied on a centralized blockchain API service, Infura [137]. This reliance continued throughout the years. In 2020, a service outage demonstrated in practice the hazards of such centralization [180]. In 2022, a misconfiguration on Infura's part resulted in wallets (and, thus, user funds) being inaccessible [45]. In terms of applications, OpenSea is the leading hosting service for NFTs. As of 2021, it reportedly handled 98% of all NFT volume [182], charging a 2.5% commission on all sales. As expected, an OpenSea outage in 2022 also resulted in the NFT market being practically unusable [69].

C.12 Governance: Improvements and Conflict Resolution

Most systems employ an *Improvement Proposal* mechanism, where proposals are posed as issues in Github, a (centralized) system that is extensively used for software development. If a change gathers enough support, it is incorporated in the codebase. To voice approval for proposals, miners often include encoded messages in blocks. From early on, proposals in Bitcoin and Ethereum have been made by a handful of developers [80,12]. In the discussion phase, many people participate but again only a few actors contribute most comments, while in cases like Bitcoin the groups of developers and commenters largely overlap.

C.13 Governance: Development Funding

Most existing blockchain systems follow the first approach, i.e., not making funding provisions. In many cases, funding is channeled through a few foundations and companies.[39] Treasuries are present in some ledgers, like Decred, Cardano, and Dash. Despite their potential though, widespread funding has yet to be demonstrated for most systems.[40]

C.14 Geography

In 2014, 37% of Bitcoin nodes resided in the US and China [58]. In 2019, Bitcoin mining hardware was mostly located in China (particularly Sichuan) and the

[39] The first usually take the name of the token, e.g., the {Bitcoin, Ethereum, Cardano} Foundations. Examples of the second are the ASIC companies discussed in Section 3 or software companies like Blockstream (Bitcoin), Consensys (Ethereum), Input Output (Cardano), *etc.*

[40] Decred's treasury holds \$23.8M, and has allocated \$250K over the past year. Cardano's treasury holds approx. \$500M and has distributed \$17.2M across 939 projects. Dash, one of the first systems to set a treasury, allocated \$500,000 over 2018, but it appears non-functional as of 2022. [dcrdata.decred.org, cardano.ideascale.com, dashvotetracker; August 2022]

US [20,107]. Notably, the mining pools then-located in China accounted for 68% of all hashrate [164]. As of 2022, a large fraction of nodes communicates over Tor,[41], thus analyzing the network's topology is often hard. Nonetheless, more than $\frac{1}{3}$ of Bitcoin mining is presumably located in the USA, with Kazakhstan and Russia following with 18% and 11% respectively [4]. In terms of full nodes, USA and Germany see roughly equivalent shares, with other countries hosting far fewer; still, a majority communicates over Tor. Until 2021 China hosted as high as 70.9% [4] of Bitcoin mining power; following its ban that year, Bitcoin's hashrate dropped from 197 to 68 Ehash/s in one month.[42] Ethereum (pre PoS) observed similar concentration patterns; by far the most nodes are located in the USA (37%) and, secondarily, Germany (16.66%) [109]. Finally, Monero nodes are mostly located in the US and, to a lesser extent, elsewhere [35]. Table 4 shows various systems' geographical distribution.

Bitcoin (*Total:* 7670)	US	2862	**Ethereum** (*Total:* 5884)	US	2781
	DE	1164		DE	678
	FR	505		SG	283
	CA	377		UK	241
	NL	309		FI	236
Bitcoin Cash (*Total:* 987)	US	419	**Dogecoin** (*Total:* 1096)	US	475
	DE	148		DE	187
	FR	69		FR	71
	NL	35		CA	47
	CA	35		CN	33

Table 4. Geographical distribution of full nodes in various ledger systems. (blockchair, ethernodes; June 2022)

In terms of legal jurisdiction, different aspects are centralized in different countries. In Bitcoin, the 4 companies that predominantly produce mining hardware[43] are all based in China. Regarding secondary markets, many exchanges operate in multiple countries (Table 3); 20 of 121 operate in USA, thus falling under US jurisdiction, 17 in China, and 10 in Japan, with the rest spread across the world. However, only 8 are based in the US, with most registered in the Seychelles (13) and other "offshore" locations. Many ICOs also exclude US investors, following their US classification as securities [161]. Finally, an interesting

[41] 46.5% of Bitcoin's nodes operate over Tor. [bitnodes; August 2022]
[42] bitinfocharts.com
[43] Bitmain, MicroBT, Canaan, Ebang.

case concerns the Bitcoin Core software, which is not available via bitcoin.org in the UK, following a related court ruling [10].

D Measuring decentralization

Our work offers a framework for analyzing blockchain decentralization, but not specific metrics to quantitatively measure it. For example, a metric could assign a single number to reflect how close a system is to a single point of failure, given a distribution of resources over a set of relevant parties. Here, we briefly review some metrics, at a high level, and leave for future work the exploration of alternatives and the computations over real-world data. A first option is Shannon entropy [163]. Briefly, a random variable's entropy measures the uncertainty of its possible outcomes. In our setting, the more bits of entropy in the resource distribution, the more diverse it is, thus the more decentralized the measured component is. Min-entropy, i.e., the smallest of the Rényi family of entropies [149] can be also used instead since it also offers a lower bound. An alternative is the Gini coefficient [154]. Gini expresses the percentage of space between the 45^o line and the curve that plots the cumulative wealth y owned by the bottom x of the population. Intuitively, a Gini value of 0 implies perfect equality, where each person owns the same amount of resources, while 1 reveals extreme inequality. Alternative metrics could also help evaluate different aspects of decentralization. Examples from traditional economics are the Theil [168], Atkinson [9], and Herfindahl-Hirschman [150] indices. Drawing from the blockchain space, an often-used metric is the Nakamoto coefficient [143], which measures the minimum number of parties that control a majority of resources. Nonetheless, a systematic comparison of all alternatives is an interesting question for future research.

E Fault Tolerance and Decentralization

Decentralization disperses control across a large set of parties. This is seemingly beneficial for Byzantine Fault Tolerant (BFT) systems. On the other hand, it may be counterproductive for other notions of faults. Specifically, the goal of BFT systems is to sustain corruptions of some (bounded) number of participants. Therefore, avoiding single points of failure and distributing the system's operation is particularly useful in this context. The more decentralized a BFT system is, the more parties an adversary needs to corrupt. Non-BFT systems, which are e.g., crash fault tolerant, may not be able to sustain the corruption of any participant. In other words, even if a single participant behaves in a Byzantine manner, the system's properties cannot be guaranteed. Thus, the more decentralized a non-BFT system is, the larger its attack surface. Therefore, its security relies on the security of the weakest participant. For larger numbers of participants, i.e., if the system is more decentralized, the likelihood that an adversary can corrupt any one participant typically increases, since participants often do not have the same level of security. Therefore, exploring the relationship between decentralization and fault tolerance, as well as the settings where decentralization is beneficial and those where it is not, is another interesting topic of future research.