



UvA-DARE (Digital Academic Repository)

Censorship-Resistance and Compliance Behavior in the Ethereum Consensus Mechanism

Boss, S.; Bodó, B.

DOI

[10.1109/ICBC64466.2025.11114701](https://doi.org/10.1109/ICBC64466.2025.11114701)

Publication date

2025

Document Version

Final published version

Published in

2025 IEEE International Conference on Blockchain and Cryptocurrency (ICBC 2025)

License

Article 25fa Dutch Copyright Act (<https://www.openaccess.nl/en/policies/open-access-in-dutch-copyright-law-taverne-amendment>)

[Link to publication](#)

Citation for published version (APA):

Boss, S., & Bodó, B. (2025). Censorship-Resistance and Compliance Behavior in the Ethereum Consensus Mechanism. In *2025 IEEE International Conference on Blockchain and Cryptocurrency (ICBC 2025): Pisa, Italy, 2-6 June 2025* (pp. 109-113). IEEE. <https://doi.org/10.1109/ICBC64466.2025.11114701>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (<https://dare.uva.nl>)

Censorship-Resistance and Compliance Behavior in the Ethereum Consensus Mechanism

Stefanie Boss

*Institute for Information Law, Institute for Informatics, Data
Science Center
University of Amsterdam
Amsterdam, the Netherlands
s.boss@uva.nl*

Balázs Bodó

*Institute for Information Law
University of Amsterdam
Amsterdam, the Netherlands
b.bodo@uva.nl*

Abstract— This paper examines Ethereum’s Proof-of-Stake (PoS) consensus mechanism and the factors shaping compliance behavior through statistical analysis and anomaly detection. Although PoS was designed to uphold credible neutrality and decentralization, the results show diverse behavior among builders, relays, and validators, driven by their roles, incentives, and the system’s design. Features like proposer-builder separation (PBS) and Maximal Extractable Value (MEV) enhance the capacity of builders and relays to influence transaction inclusion, while validators’ influence is mostly limited to their proposer tasks. The paper further shows that partial enforcement of sanctions is insufficient to eliminate sanctioned transactions from the network, which demonstrates the challenge of balancing regulatory compliance with decentralization. In the current state, there is an inherent tension within Ethereum’s consensus mechanism, where both credible neutrality and compliance seem compromised.

Keywords— *Ethereum, Blockchain, Consensus, Proof-of-Stake, Compliance, Censorship-Resistance*

I. INTRODUCTION

Ethereum’s efficacy of the credible neutrality principle, which ought to be executed by the consensus mechanism, currently fails to function. Credible neutrality refers to a mechanism that does not discriminate for or against any specific people or outcomes to the furthest possible extent [1]. When applying this principle to the consensus mechanism, which is responsible for validating and ordering transactions, the idea would be to have an objective and clear set of rules that decides upon ordering, inclusion and exclusion. In reality, however, the actors in that consensus mechanism turn out to take different factors into account when executing this task [2]. The heterogeneity in operation of their gatekeeping role facilitates a reality where there is no unilaterally agreed upon – manner to approach the task. It has become difficult to find actual consensus about whether a transaction merits to be added to the blockchain, and in which place in the order it belongs. At this point, transaction inclusion has become a matter of timing and luck with having the ‘right’ builder, relay or validator to decide upon a transaction request, rather than an absolute distinction between inclusion and exclusion.

Several factors have contributed to this reality where the consensus mechanism has become both more subjective and obscure. Two main factors that will be discussed in this paper are application of censorship for compliance purposes, and maximum extractable value (MEV) as an incentivization mechanism that drives validators to approach their tasks in more opportunistic, profit-driven, and thus arbitrary yet predictable manner.¹

II. BACKGROUND

A. Ethereum Consensus, Validation and PBS

The Ethereum blockchain is an open-source, permissionless, decentralized blockchain platform. It uses a peer-to-peer network that consists of nodes to securely execute and verify transactions [2], [3], [4]. In the consensus mechanism, the aim is to find consensus on the inclusion and order of transactions that have been requested to be added to the blockchain. The system works as follows [4], [5]. Users request transactions, and bid the gas that they are willing to pay for the transaction processing. These transaction requests end up in the mempool. Builders construct full blocks with the transactions in the mempool and their private order flow, and submit bids of these blocks to relays. Some builders ask searchers to search a fitting mixture of transactions for them to include in the block. Then, the relay verifies the validity of the blocks it received from its builders, and selects the most profitable block. The relay sends this block to the MEV-Boost, which receives multiple blocks from multiple relays, picks the most profitable block, and sends this block to a proposing validator. The proposer receives the blind blocks, chooses the most profitable block that it received from the relays that it is registered to, signs the block, and sends it back to the relay that it received the block from. The relay verifies the proposers’ signature, and sends the full block to the proposer, who propagates it to the validators in the peer-to-peer network to attest to.

B. Content Agnosticism and Censorship Resistance

The Ethereum consensus mechanism is designed to be content agnostic and censorship resistant. This means that filtering is not prescribed or encouraged, and all transactions that comply with the technical standards of the network and are consistent with the transaction history shall be included. The main drivers of inclusion and ordering should therefore be, for example, the alignment of transactions with the networks’ rules, and whether the transactor has sufficient funds to complete the transaction [2].

However, recent events are pointing towards some more elaborate moderation practices, where (financial) incentives and regulation rise as additional factors in the decision-making process for consensus layer participants. The Merge, where Ethereum moved to PoS, has led to a change in the reward structure of the consensus mechanism, shifting incentives to an even more profit-driven system [6]. Another event is the decision of the U.S. Office of Foreign Asset Control (OFAC) to sanction several Ethereum wallet addresses. Addresses on this list are considered illegal to interact with. Much focus lied on the Tornado Cash sanctions

¹ Research artifacts and data can be found at:

https://osf.io/4qtfr/?view_only=ea8888de0274a8fa447ccabf37e5baf

of August 2022, a service that had been subject to multiple lawsuits and enforcement actions, because it was found to facilitate money laundering and terrorist financing [5], [7], [8]. Whereas the Tornado Cash sanctions have been overturned in November 2024 [9], they had started discussions about censorship among consensus layer actors, and to what extent censorship is induced by compliance-related reasons.

III. RELATED WORK

Ramos and Ellul [10] and Barcentewicz [11] have found that MEV on the consensus layer may lead to undesired censorship that is not just affecting free entry of transactions into the network, but also allows for problematic market manipulation practices. Kraner et al. [12] have identified similar risks. Wahrstätter et al. [4] have attempted to unravel the block construction market, the roles and interactions of actors on the market, and how the market deals with MEV and the accompanied incentives.

Wang et al. [5] identified the relevance of consensus level censorship, as a result of the OFAC sanction lists. They point out that in the Ethereum ecosystem, several actors have shown efforts to comply with these sanction lists by censoring the sanctioned addresses, despite it being easy to circumvent. They also identify that these censorship practices make them targets for denial of service (DoS) attacks, compromising their security. Wahrstätter et al. [13] have identified similar security concerns, and have found that censorship does not only result in exclusion, but may also show in the form of delayed inclusion. This occurs when not every builder, proposer or validator is censoring, so that inclusion becomes a matter of time and luck, instead of it being a full ban. Kraner et al. [12] have modeled Ethereum's consensus mechanism, and found that the mechanism is vulnerable for different kinds of attacks, such as reorganization attacks and balancing attacks. They have also found that external factors, such as latency, affect the efficiency and inclusion in the consensus process. Öz et al. [6] have identified that inclusion delays may also stem from the monetary incentivization to wait with inclusion of certain transactions. This has similarly been found by Wahrstätter et al. [4], who further highlight that certain staking consortia have attempted to mitigate such delays, whereas others are more flexible regarding the timing of proposing within a slot. Grandjean et al. [14] have identified that innovations such as liquid staking and pooled staking may affect incentives in the consensus system, as there are fewer participants that receive more spread-out funds, rather than absolute and direct rewards or punishments for their actions. Lastly, similar problems surrounding adverse and selfish behavior have shown in other blockchains, such as in the Bitcoin blockchain [15], [16].

This paper contributes to scholarship, by analyzing and providing insights into consensus layer censorship from a compliance standpoint. We use a granular approach, where we analyze which actors in the consensus mechanism can truly exercise censorship, and to what extent they engage in such practices. It will provide first steps to analyzing factors that come into play in that decision-making process.

IV. METHODOLOGY AND DATA

This paper follows a mixed methodology approach. Hypotheses are drafted from a scoping literature review, and descriptive statistics and machine learning mechanisms for anomaly detection will be used to test the hypotheses and retrieve results. We have used the following datasets:

- *Blacklisted addresses* (13.09.2019 – 01.05.2024): list of addresses that are sanctioned by the OFAC, with the official sanction dates.
- *Gas* (07.08.2015 – 27.05.2024): daily gas information (gas limit mean, gas limit median, gas used mean, and the gas used median).
- *Transactions* (13.10.2017/ block 4362139 – 13.06.2024/ block 20084671): Ethereum transactions that interact with a sanctioned address, with information such as transaction dates, value transferred in ETH, and the gas fee.
- *Block info* (15.09.2022/ block 15538721 – 18.10.2023/ block 18373769) [17]: information about Ethereum blocks, including the number, the builder, the relay, the validator, and the MEV value.

V. PROBLEM STATEMENTS AND HYPOTHESES

A. Incentive-related Problems

The protocols governing the validation process in Ethereum do not prescribe or encourage the filtering of transactions. Consensus layer actors may, however, choose to voluntarily comply with state regulation: a growing number of such actors comply with the OFAC sanction list, even if they are not obliged to do so – or at least, not clearly obliged by law to do so [18]. This leads to the first hypothesis: *some actors in the Ethereum consensus layer, but not all, adhere to legal norms (hypothesis 1).*

B. System Design Related problems

Apart from intrinsic motivations, the consensus mechanism design also influences censorship behavior. A relevant aspect is the block-content knowledge of actors. Builders and relays can see the transactions in the blocks that they deal with and can use that knowledge for strategic choices around profit maximization and compliance. Such conscious decision-making is more difficult for proposers, as they receive blind blocks, and can only influence the inclusion through choosing specific relays they want to work with. Attesting validators have the least influence on transaction inclusion, as refusing to attest is the only way to censor and carries slashing risks. This leads to the second hypothesis: *Builders and relays provide better prediction accuracy, as they have more active control over censorship than proposers and attestors, due to their access to transaction details and fewer punishment risks (hypothesis 2.1)*

Further, actors may be more willing to accept transactions in their blocks that provide higher rewards. Sanctioned addresses may share this knowledge, and offer higher gas, to make inclusion seem more profitable. This exploits their profit-driven nature and speeds up the process of sanctioned transactions being included in a block. This leads to the last hypotheses: *Illegal transactions are accompanied with higher gas (hypothesis 2.2), and actors that are related to high MEV-rewards include more illegal transactions (hypothesis 2.3).*

VI. RESULTS

A. Descriptive Statistics

While sanctioned addresses appear less active after their sanction date, there still is some continued activity (fig. 1). To better understand this pattern, we performed a windowed analysis around the sanction date, comparing transaction volumes in five time intervals (5, 10, 15, 20, and 50 days).

Because the data is not normally distributed, we used a wilcoxon test to assess statistically significant (table 1). We observe a consistent and statistically significant increase in transaction volume after sanctioning across all windows, except for the 50-day window, which showed no significant difference. These findings suggest that sanctions may not immediately suppress activity; instead, they may trigger short-term behavioral responses. A possible explanation for higher activity may relate to money fleeing before the sanctions are technically implemented, but this requires further research.

Further, we observe a statistically significant, upward trend for gas prices after the sanction date through a similar wilcoxon test (table 2). We also confirm with a t-test (fig. 2) that there is a significant difference in gas limits with all gas on average, with the sanctioned address typically carrying a higher gas limit. We furthermore see that in some cases, a sanction is shortly followed by a change in gas (fig. 2).

B. Consensus Analysis

Next, we analyze the compliance rates of specific actors in Ethereum’s consensus mechanism, namely the builders, relays and validators. On average, relays include sanctioned transactions in 2.95% of their blocks, builders in 1.08% of their blocks, and validators in 1.58% of their blocks. We observe that compliant relays, such as BloXroute, Eden and Flashbots have fairly low numbers of sanctioned transactions in their blocks, although they are not zero. For BloXroute Max Profit, this may be explained by their late switch towards a compliant attitude in December 2023. Ethereum-wide, we see a trend where less active actors have relatively high numbers of sanctioned addresses in their blocks (fig. 3, 4)

C. Anomaly Detection Experiment

We performed an anomaly detection exercise to identify anomalous decision-making patterns with Isolation Forest (IF) and CatBoost Classifier (CB). We selected these models for their ability to handle high-dimensional, imbalanced data and to account for categorical and non-linear features.

The dataset was preprocessed through label encoding, feature selection, and standardization. Rows with missing values were removed to ensure a clean training set. Key features included block timing (slot, block nr.), actor roles (relay, builder, proposer, validator), revenue (MEV value), and legality. IF, an unsupervised algorithm, was used to detect rare patterns without labeled training data. Configured with a contamination rate of 1%, which reflects the proportion of illicit blocks identified in the previous sections, it flagged blocks that deviated significantly from the norm. In addition, CB was trained on labeled data using the same features, enabling a supervised perspective. The CB model achieved high accuracy (99%) and precision (94.8%), indicating that it was effective at identifying true anomalies with few false positives. However, the recall (29%) and F1 score (44.3%) indicate a limited ability to detect all relevant anomalies, which is a common challenge in rare-event classification. When trained exclusively on labeled data, recall improved to 48%, but at the cost of reduced precision (34%) and accuracy (98%), highlighting a trade-off between sensitivity and reliability (a F1 score of 40%). In a future study, we suggest experimenting with a hybrid approach that integrates the precision of supervised models with the generalization ability of unsupervised models to enhance the anomaly detection.

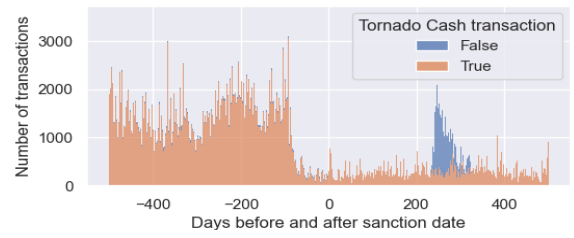


Fig. 1. Plot of sanctioned transactions, relative to their sanctioning date, and based on whether it contains a Tornado Cash element.

TABLE I. WILCOXON RESULTS FOR TRANSACTION VOLUME

Window	Mean before	Mean after	Wilcoxon
5 days	11.658333	25.891667	0.000341
10 days	19.350000	35.958333	0.000262
15 days	30.416667	41.300000	0.002427
20 days	38.150000	46.408333	0.002404
50 days	100.158333	104.116667	0.068172

TABLE III. WILCOXON RESULTS FOR GAS

Window	Mean before	Mean after	Wilcoxon
5 days	44819.021609	439492.488374	2.369128207519189e-08
10 days	49581.019238	471117.543112	2.0275377998200918e-08
15 days	52698.147046	469023.992073	3.696324226119073e-08
20 days	52361.658032	455556.246045	2.7754215046155825e-08
50 days	58448.949369	448363.190173	3.123590717931469e-08

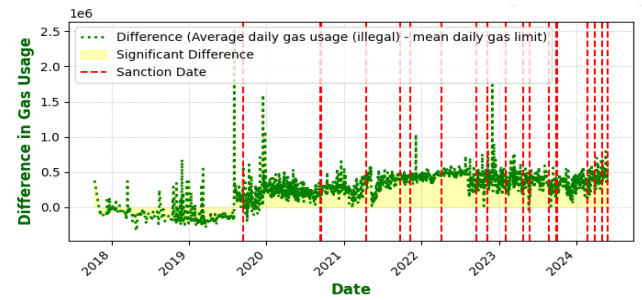


Fig. 2. differences in gas for sanctioned transactions over time, compared to regular gas limits, plotted next to relevant sanction dates, with a t-statistic of 53.609653053299695 and a p-value of 0.0.

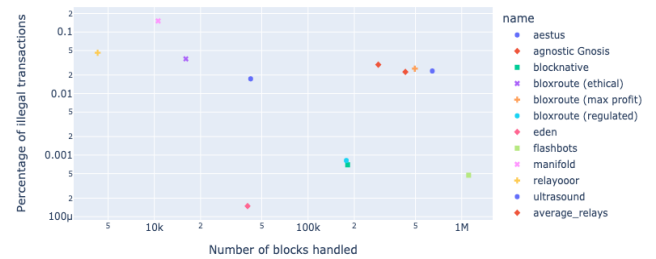


Fig. 3. Relay block volume and percentage incompliant blocks of relays

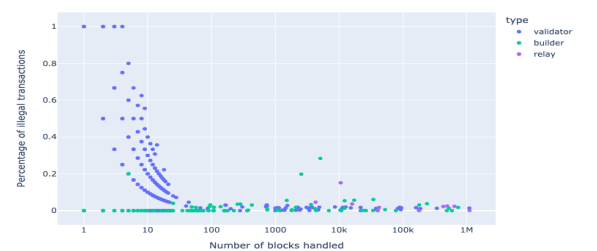


Fig. 4. Ecosystem block volume and percentage incompliant blocks

Across both exercises, the most important predictive features include relay, block number, builder public key, and slot, underscoring the central role of specific actors and timing in shaping inclusion patterns (fig. 5). Notably, MEV value showed low predictive power, suggesting that while MEV is often cited as a key driver of behavior, it may not be a direct indicator of compliance decisions in practice.

D. Indication of Knowledge Points .

Taken all the above into account, we considered the actual influence that each actor has in each stage on transaction inclusion (fig. 6). Key actors do the following.

1) *Users – influence*: By bidding higher gas (fees), a user may be able to push a sanctioned transaction through that is normally not prioritized. The statistics indicate that sanctioned users may be doing this. In the anomaly detection exercise, it nevertheless shows that MEV-value shows to be of low value.

2) *Searchers – influence*: Searchers monitor the mempool and send bundles of transactions to builders. They have discretion to include certain transactions, which may also be sanctioned transactions, for their own benefit.

3) *Builders – high influence*: Builders compose blocks by themselves, or based on the bundles they receive from searchers. They generally aim for maximum profit, but relevant factors may vary individually. They show as a medium to high influence in the anomaly detection exercise.

4) *Relays – high influence*: Relays receive blocks from their builders and verify the execution payloads to ensure that they are valid. Some relays have a policy in which they have solidified that they filter out illegal transactions. They show as a medium to high influence in the anomaly detection exercise.

5) *MEV Boost – no influence*: MEV-Boost connects proposers with the relays that they have selected, and picks the most profitable block from the chosen relays. The protocol itself is neutral and doesn't influence the transaction content.

6) *Proposer – medium influence*: The proposer is blindly selects the most profitable block that is offered by its connected relays. While this suggests a lack of influence, they can choose the relays that they wish to receive blocks from, so that they can consciously choose for compliant relays, and thus influence the inclusion of (illicit) transactions. They show as a medium influence in the anomaly detection exercise.

7) *Validators – medium to low influence*: Validators attest to the correctness of the proposed blocks. They can censor illegal blocks by refusing to attest them. However, they risk being slashed by such refusal. This makes it too late and likely too risky to still have an influence. They show as low influence in the anomaly detection exercise.

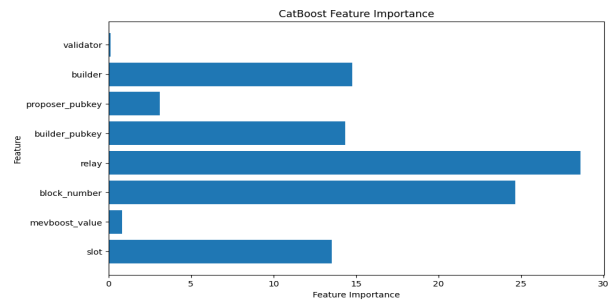


Fig. 5. CatBoost feature importance

VII. CONCLUSION

In sum, the current Ethereum consensus landscape is defined by a tension between neutrality, compliance and profit. We show that while Ethereum's consensus mechanism is theoretically grounded in credible neutrality, it is influenced by a range of actor-specific motivations and system-level constraints. Some consensus layer participants voluntarily comply with legal norms, such as with the OFAC sanction list (hypothesis 1). Relays, for example, that advertise as compliant show lower rates of sanctioned transactions in their blocks. Though, this voluntary compliance is not uniform across actors and appears to be influenced by system design. The roles and block content knowledge of different actors create asymmetries in their ability to censor. The anomaly detection exercise namely confirms that builders and relays, who have direct control over block content and fewer penalty risks, have more influence over transaction inclusion than proposers or validators, who operate under stricter technical and protocol constraints (hypothesis 2.1).

We further find that sanctioned transactions carry higher gas, which may reflect an attempt by users to incentivize inclusion through profit (hypothesis 2.2). Conversely, MEV value does not show amongst the most important features in the anomaly detection exercise, suggesting that high MEV strategies do not necessarily lead to more non-compliant behavior. This challenges hypothesis 2.3, where we assumed a relationship between MEV rewards and non-compliance.

Lastly, we observe the limitation of partial enforcement in the system: isolated censorship efforts cannot fully ensure compliance. Even when some actors attempt to comply, the network's design allows sanctioned transactions to eventually propagate and be included, albeit with delay. Addressing this limitation may require a coordinated, system-wide approach, rather than individual policy shifts. In a future study, we recommend to qualitatively investigate more individual factors that underly an actor's decision-making path, such as the value they attach to the credible neutrality principle, monetary gains, their reputation and their jurisdiction.

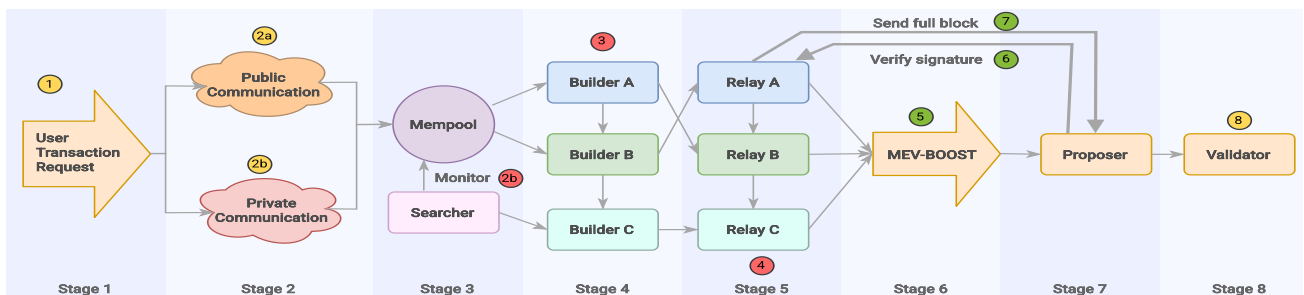


Fig. 6. Sketch of the Ethereum consensus mechanism with knowledge points. Influential actors are indicated with a red dot, other actors have been indicated with a green dot. Validators are marked orange, as they have a minor influence on the inclusion of blocks.

REFERENCES

- [1] V. Buterin, "Credible Neutrality As A Guiding Principle", Nakamoto. [Online]. Available: <https://nakamoto.com/credible-neutrality/>
- [2] M. Nofer, P. Gomber, O. Hinz, and D. Schiereck, "Blockchain", *Business & Information Systems Engineering*, vol. 59, pp. 183–187, 2017
- [3] D. Mancino, A. Leporati, M. Viviani, and G. Denaro, "Exploiting Ethereum after" The Merge": The Interplay between PoS and MEV Strategies.", *ITASEC*, 2023
- [4] A. Wahrstätter, L. Zhou, K. Qin, D. Svetinovic, and A. Gervais, "Time to bribe: Measuring block construction market", 2023, arXiv:2305.16468. [Online]. Available: <https://arxiv.org/abs/2305.16468>
- [5] Z. Wang, X. Xiong, and W. J. Knottenbelt, "Blockchain Transaction Censorship:(In) secure and (In) efficient?", 2023, Cryptology ePrint Archive. [Online]. Available: <https://eprint.iacr.org/2023/786>
- [6] B. Öz, B. Kraner, N. Vallarano, B. S. Kruger, F. Matthes, and C. J. Tessone, "Time moves faster when there is nothing you anticipate: The role of time in mev rewards", *Proceedings of the 2023 Workshop on Decentralized Finance and Security*, pp. 1–8, 2023.
- [7] "U.S. Treasury Sanctions Notorious Virtual Currency Mixer Tornado Cash", U.S. Department of the Treasury News, Washington, Aug. 2022. [Online]. Available: <https://home.treasury.gov/news/press-releases/jy0916#:~:text=WASHINGTON%20%E2%80%93%20Today%2C%20the%20U.S.%20Department,since%20its%20creation%20in%202019.>
- [8] C. Shumba, "Dutch Prosecutors Seek 64-Month Jail Sentence for Tornado Cash Dev Alexey Pertsev", Coindesk, Mar. 2024. [Online]. Available: <https://www.coindesk.com/policy/2024/03/27/dutch-prosecutors-seek-64-month-jail-sentence-for-tornado-cash-dev-alexey-pertsev/>
- [9] N. Raymond, "Court overturns US sanctions against cryptocurrency mixer Tornado Cash", *Reuters*, Nov. 2024. [Online]. Available: <https://www.reuters.com/legal/court-overturns-us-sanctions-against-cryptocurrency-mixer-tornado-cash-2024-11-27/>
- [10] S. Ramos and J. Ellul, "The MEV Saga: Can Regulation Illuminate the Dark Forest?", *the International Conference on Advanced Information Systems Engineering*, Springer, pp. 186–196, 2023
- [11] M. Barczentewicz, A. Sarch, and N. Vasan, "Blockchain transaction ordering as market manipulation", *Ohio St. Tech. LJ*, vol. 20, p. 1, 2023
- [12] B. Kraner, N. Vallarano, C. Schwarz-Schilling, and C. J. Tessone, "Agent-based modelling of ethereum consensus", *the 2023 IEEE International Conference on Blockchain and Cryptocurrency (ICBC)*, IEEE, pp. 1–8, 2023
- [13] A. Wahrstätter *et al.*, "Blockchain Censorship", *Proceedings of the ACM Web Conference*, pp. 1632-1643, 2024
- [14] D. Grandjean, L. Heimbach, and R. Wattenhofer, "Ethereum proof-of-stake consensus layer: Participation and decentralization", *International Conference on Financial Cryptography and Data Security, Springer Nature Switzerland*, pp. 253-280, 2024
- [15] S.-N. Li, C. Campajola, and C. J. Tessone, "Statistical detection of selfish mining in proof-of-work blockchain systems", *Scientific Reports*, vol. 14, no. 1, p. 6251, 2024.
- [16] C. Schwarz-Schilling, S.-N. Li, and C. J. Tessone, "Stochastic modelling of selfish mining in proof-of-work protocols", *Journal of Cybersecurity and Privacy*, vol. 2, no. 2, pp. 292–310, 2022.
- [17] A. Wahrstätter, "eth_data". [Online]. Available: <https://mevboost.pics/data.html>
- [18] Labrys, "MEV-Watch". [Online]. Available: <https://www.mevwatch.info/>