Collusion-Resilience in Transaction Fee Mechanism Design

Hao Chung*
Carnegie Mellon University
haochung@andrew.cmu.edu

Tim Roughgarden†
Columbia University and a16z crypto
tim.roughgarden@gmail.com

Elaine Shi*
Carnegie Mellon University
runting@cs.cmu.edu

Abstract

Users bid in a transaction fee mechanism (TFM) to get their transactions included and confirmed by a blockchain protocol. Roughgarden (EC'21) initiated the formal treatment of TFMs and proposed three requirements: user incentive compatibility (UIC), miner incentive compatibility (MIC), and a form of collusion-resilience called OCA-proofness. Ethereum's EIP-1559 mechanism satisfies all three properties simultaneously when there is no contention between transactions, but loses the UIC property when there are too many eligible transactions to fit in a single block. Chung and Shi (SODA'23) considered an alternative notion of collusion-resilience, called c-side-contract-proofness (c-SCP), and showed that, when there is contention between transactions, no TFM can satisfy UIC, MIC, and c-SCP for any $c \geq 1$. OCA-proofness asserts that the users and a miner should not be able to "steal from the protocol." On the other hand, the c-SCP condition requires that a coalition of a miner and a subset of users should not be able to profit through strategic deviations (whether at the expense of the protocol or of the users outside the coalition).

Our main result is the first proof that, when there is contention between transactions, no (possibly randomized) TFM in which users are expected to bid truthfully satisfies UIC, MIC, and OCA-proofness. This result resolves the main open question in Roughgarden (EC'21). We also suggest several relaxations of the basic model that allow our impossibility result to be circumvented.

*Supported by NSF awards 2212746, 2044679, 1704788, a Packard Fellowship, a generous gift from the late Nikolai Mushegian, a gift from Google, and an ACE center grant from Algorand Foundation.

†Author's research at Columbia University supported in part by NSF awards CCF-2006737 and CNS-2212745, and research awards from the Briger Family Digital Finance Lab and the Center for Digital Finance and Technologies.

Contents

1 Introduction **1**
- 1.1 Our Contributions . . . 2
- 1.2 TFM Incentive-Compatibility Notions: A Cheat Sheet . . . 4

2 Definitions **5**
- 2.1 Transaction Fee Mechanism in the Plain Model . . . 5
- 2.2 Incentive Compatibility Notions . . . 7

3 Preliminary: Myerson's Lemma **8**

4 Warmup: Impossibility of UIC + MIC + Global SCP for Deterministic Mechanisms **9**

5 Impossibility of UIC + MIC + Global SCP for Randomized Mechanisms **10**
- 5.1 Proof Roadmap . . . 10
- 5.2 Formal Proofs . . . 11

6 Feasibility and Impossibility of UIC + MIC + OCA-Proof **13**
- 6.1 A Non-Truthful Mechanism with UIC + MIC + OCA-Proof . . . 14
- 6.2 Impossibility of UIC + MIC + OCA-Proof for Truthful Mechanisms . . . 14

7 How to Circumvent the Impossibilities **19**
- 7.1 Allowing the Globally Optimal Strategy to Coordinate . . . 19
- 7.2 Allowing the Globally Optimal Strategy to Output Multiple Bids . . . 20
- 7.3 Inclusion-Rule-Respecting . . . 22

8 Transaction Fee Mechanisms in the Trusted Hardware Model **22**
- 8.1 The Trusted Hardware Model . . . 22
- 8.2 Defining Bayesian Incentive Compatibility . . . 23
- 8.3 Second-Price Auction Achieves UIC + Bayesian MIC + Global SCP . . . 24
- 8.4 Impossibility of Bayesian UIC + Bayesian MIC + Bayesian 1-SCP + Global SCP . . . 27

9 Static Revelation Principle for Transaction Fee Mechanisms **30**
- 9.1 Static Revelation Principle: Bidding Rules That Output Single Bid . . . 30
- 9.2 Static Revelation Principle: Allowing Bidding Rules that Output Multiple Bids . . . 31

A Comparison of Collusion-Resilience Notions **37**

B Alternative Proof of Theorem 5.6 **40**

1 Introduction

Real estate on the blockchain is scarce, and blockchain users bid in an auction called the transaction fee mechanism (TFM) to have their transactions included and confirmed on the blockchain. The original Bitcoin protocol adopted a simple first-price auction, where the top k bids win and they each pay their bid. However, such first-price auctions are known to incentivize untruthful bidding. Therefore, a line of subsequent works [LSZ19, Yao, BEOS19, BCD+, Rou20, Rou21, FMPS21, CS23, SCW23, WSC24, GY22, ZCZ22, BGR23, TY23, KKLP23, XFP23, CMW23, LRMP23, Ndi23] explored what is the "dream TFM" for blockchains. Most works [Rou20, Rou21, CS23, SCW23, WSC24, GY22, ZCZ22, BGR23, TY23] agree on roughly the same set of desiderata, that is, a dream TFM should provide incentive compatibility not just for an individual user, but also for the miner of the block. Further, a dream TFM should provide resilience against miner-user collusion.

Roughgarden [Rou21] was the first to formally define the aforementioned requirements for a TFM, which he referred to as *user incentive compatibility*[1], (myopic) *miner incentive compatibility*, and *OCA-proofness*, where OCA stands for "off-chain agreement" and refers to colluding strategies between the miner and a set of users that allow offchain transfers. Roughgarden [Rou21] also showed that the simple "posted price auction with all fees burnt" mechanism, which corresponds to the behavior of Ethereum's EIP-1559 TFM [BCD+] when there is no congestion, satisfies all three properties. However, the posted price auction with all fees burnt does not satisfy all three properties when there is congestion. In practice, congestion does occur when there are major events such as an NFT mint or price fluctuations — for example, in Ethereum, roughly 2.3% of the blocks experience congestion.[2] When congestion arises, approximately speaking, Ethereum's EIP-1559 mechanism falls back to the first-price auction, violating user incentive compatibility. Therefore, an interesting question is whether we can design a dream TFM satisfying all three properties for finite block sizes.

Chung and Shi [CS23] considered an alternative notion of collusion-resilience, called side-contract-proofness. Unfortunately, they proved that no (even randomized) TFM can simultaneously satisfy user incentive compatibility and side-contract-proofness. Because side-contract-proofness is a more demanding property than OCA-proofness, the question raised by Roughgarden [Rou21], of whether there is a dream TFM satisfying all three properties under his collusion-resilience notion, had remained open.

Two notions of miner-user collusion-resilience Multiple natural notions of collusion-resilience can and have been studied in the context of TFM design. Here we clarify informally the key differences between the notions proposed by Roughgarden [Rou21] and Chung and Shi [CS23]. These notions are defined formally in Definitions 2.4–2.6 (see Section 2) and compared further via examples in Appendix A.

- **OCA-proofness**: Roughgarden's notion, henceforth referred to as OCA-proofness, asserts that there should exist a "reference strategy" for a miner and all users that is guaranteed to maximize their joint surplus. In this reference strategy, the miner is expected to follow the inclusion rule intended by the TFM. For users, the definition requires only that users follow some fixed bidding strategy σ (i.e., a mapping from a private user valuation to a user bid)

[1] User incentive compatibility (UIC) is usually called dominant-strategy incentive compatible (DSIC) in the mechanism design literature. In general, we allow UIC TFMs to make use of non-truthful (but dominant) bidding strategies (see Definition 2.2).

[2] From Jan 1, 2024 to Feb 5, 2024, 256595 blocks have been produced on Ethereum, and 5840 blocks among them were full (meaning more than 99.9% of the gas limit (30M) was used).

that is individually rational (i.e., $\sigma(v) \leq v$ for all $v \geq 0$). In particular, in the reference strategy, users are expected to bid independently (with a user's bid independent of other users' valuations and bids), and expected to submit a single bid (with no additional fake bids injected). One example of such a bidding strategy is the truth-telling strategy (with $\sigma(v) = v$). Because Roughgarden [Rou21] wished to discuss the OCA-proofness properties of non-UIC TFMs like first-price auctions, the definition also allows the reference strategy to be defined by a non-truthful bidding strategy (e.g., $\sigma(v) = v/2$). As a consequence, to prove that a TFM is both UIC and OCA-proof, it is sufficient to prove that it is UIC under one bidding strategy and OCA-proof under a possibly different bidding strategy (as in the example in Section 6.1).

- **c-SCP**: Chung and Shi's notion [CS23], henceforth called c-SCP (where SCP stands for side-contract-proofness), requires that the honest strategy (i.e., all users follow the honest bidding rule and the miner honestly implements the inclusion rule) is the profit-maximizing strategy for any coalition consisting of the miner of the present block and at most c users. For truthful mechanisms, the honest bidding rule is the truthful one, while for non-truthful mechanisms, the bidding rule can be more general (see Section 2.1 for the formal definition). Chung and Shi's notion aligns with standard notions used in a line of work at the intersection of game theory and cryptography [HT04, KN08, ADGH06, OPRV09, AL11, ACH11, GKM+13, GKTZ15, GTZ15, Kat08, DR07, GLR10, CGL+18, WAS22, CCWS21, PS17, KMSW22, FW20, EFW22].

Discussion The two notions of collusion-resilience address different issues. OCA-proofness captures the intuitive requirement that the users and miners *should not be able to steal from the protocol* through strategic deviations — for this reason, it considers *only the global coalition* consisting of the miner and all users. By contrast, the c-SCP notion captures the intuitive idea that a miner-user coalition's best response is to act honestly, and that no strategic deviations can allow the coalition to *steal from other users or steal from the protocol*. For further discussion, see the end of this section and Section A.

1.1 Our Contributions

As explained above, both Roughgarden's and Chung and Shi's collusion-resilience notions capture meaningful incentive compatibility considerations. Recognizing their differences, one natural question is: does Chung and Shi's finite-block impossibility result still hold if we adopt the original OCA-proofness notion of Roughgarden in lieu of c-SCP? Notably, no existing TFM construction [LSZ19, Yao, BEOS19, BCD+, Rou20, Rou21, FMPS21, CS23, SCW23, WSC24, GY22, ZCZ22, BGR23, TY23, KKLP23, XFP23, CMW23, LRMP23, Ndi23] simultaneously satisfies user incentive compatibility, miner incentive compatibility, and OCA-proofness under finite block size.

Main impossibility result In our work, we give an affirmative answer to the above question. We show that, indeed, an analog of Chung and Shi's finite-block impossibility result still holds when we replace the c-SCP requirement with OCA-proofness. Specifically, we prove the following theorem.

Theorem 1.1. *Suppose the block size is finite. Then, no possibly randomized, truthful TFM can simultaneously satisfy user incentive compatibility (UIC), miner incentive compatibility (MIC), and OCA-proofness. Further, this impossibility holds even when the globally optimal strategy σ need not be individually rational.*

In a truthful TFM, a user is expected to bid truthfully, so if the mechanism satisfies UIC, a user's utility is maximized when it just reports its true value. However, OCA-proofness allows the global coalition to adopt a non-truthful bidding strategy σ even for truthful mechanisms.

Our Theorem 1.1 is intuitively stronger but technically incomparable in comparison with Chung and Shi's impossibility, which shows that no TFM can simultaneously satisfy UIC and 1-SCP for finite block sizes. The reason is that Chung and Shi's impossibility does not rely on MIC; however, MIC is necessary for our Theorem 1.1 to hold. Specifically, a simple second-price auction with no burning (see Theorem 6.11) satisfies both UIC and OCA-proofness, but does not satisfy MIC since the miner may benefit by injecting a fake $(t+1)$-th bid where t is the number of confirmed bids, since the $(t+1)$-th bid sets the price for confirmed bids.

Global SCP We suggest a simpler version of OCA-proofness that we call **global SCP**, which also intuitively captures the requirement that strategic users and miners cannot steal from the protocol, and is perhaps more appropriate when focusing on UIC TFMs (as we do in this paper). In our work, global SCP is not only a technical *stepping stone* towards proving Theorem 1.1, but also of *independent interest* as we explain below. Specifically, global SCP is almost the same as OCA-proofness, except for requiring σ to be the honest bidding strategy indicated by the mechanism (i.e., the same bidding strategy used to establish UIC). In other words, a mechanism satisfies global SCP if and only if the honest strategy is surplus-maximizing for the global coalition. It is easy to see that for a truthful mechanism, c-SCP for any c implies global SCP, which in turn implies OCA-proofness. To prove Theorem 1.1, we first prove the following theorem:

Theorem 1.2. *Suppose that the block size is finite. Then no possibly randomized TFM can simultaneously satisfy user incentive compatibility (UIC), miner incentive compatibility (MIC), and global SCP. Further, the impossibility holds even for non-truthful mechanisms.*

We now explain why the global SCP notion is of independent interest. One advantage of global SCP is that the *revelation principle* holds for any TFM that satisfies UIC, MIC, and global SCP, which we formally prove in Section 9. In other words, given any TFM that is UIC, MIC, and global SCP, there is an equivalent truthful mechanism that simulates it. For this reason, Theorem 1.2 rules out even non-truthful TFMs that simultaneously satisfy UIC, MIC, and global SCP.[3]

By contrast, Theorem 1.1 holds only for truthful mechanisms. In particular, in Section 6.1, we show a non-truthful mechanism that simultaneously satisfies UIC, MIC, and OCA-proof. The mechanism is contrived, but it demonstrates the subtlety and the technical challenges when modeling the notion of collusion-resilience. This also suggests that the revelation principle does not hold for mechanisms that satisfy UIC, MIC, and OCA-proofness, partly because in such a mechanism, the bidding strategies used to establish UIC and OCA-proofness may be different.

Ways to circumvent the impossibilities We show in Section 7 that the impossibility of Theorem 1.1 can be circumvented by allowing non-truthful mechanisms or by allowing users to coordinate in bidding in the globally optimal strategy σ. Additionally, in Section 8, we show that the impossibility can also be circumvented by adopting a Bayesian notion of incentive compatibility. Bayesian incentive compatibility is particularly justified when the strategic players must submit their bids without having seen other honest users' bids—for example, when honest users' bids are encrypted and block production is performed by trusted hardware or through secure multi-party computation [SCW23].

[3]Simultaneously with and independently of this paper, Gafni and Yaish [GY24] proved, among other results, a version of Theorem 1.2 for the special case of deterministic mechanisms and a block size of 1.

1.2 TFM Incentive-Compatibility Notions: A Cheat Sheet

We gather here informal definitions and comparisons of the key incentive-compatibility notions used in this paper. First, a transaction fee mechanism specifies how a user is supposed to bid (as a function of its private valuation), which transactions a miner is supposed to include (as a function of the transactions it knows about and their bids), and the resulting outcome (the subset of included transactions that get confirmed, and the payments made by the users and received by the miner). If the bidding strategy suggested by the TFM is the identity, then we additionally call the TFM *truthful*. In this paper, as in the rest of the TFM literature, we consider only static mechanisms.

- **UIC.** (Definition 2.2) Provided that the miner follows the suggested inclusion rule, the bidding strategy suggested by the TFM is a dominant strategy for users.
- **MIC.** (Definition 2.3) The inclusion rule suggested by the TFM is always revenue-maximizing for the miner regardless of users' bids; moreover, the miner cannot increase its revenue through the injection of fake transactions.
- **Global SCP.** (Definition 2.5) If the miner follows the inclusion rule suggested by the TFM and all users follow the bidding rule suggested by the TFM, then their joint surplus is at least as large as it would be from any coordinated deviation.
- **c-SCP.** (Definition 2.4) For every coalition of the miner and at most c users, if the miner follows the inclusion rule suggested by the TFM and the users in the coalition follow the bidding rule suggested by the TFM, then the joint surplus of the coalition is at least as large as it would be from any coordinated deviation (holding fixed the bids of users outside the coalition).
- **OCA-proof.** (Definition 2.6) If the miner follows the inclusion rule suggested by the TFM and all users follow a suitably chosen individually rational bidding rule σ (possibly different from the one suggested in the TFM description), then their joint surplus is as large as it would be from any coordinated deviation.

For example, in [Rou21] it was shown that Ethereum's EIP-1559 TFM and a variant called the "tipless mechanism" satisfy UIC, MIC, and OCA-proofness when there is no contention between transactions; in fact, in this case, these TFMs satisfy the c-SCP condition for every $c \geq 1$. When there is contention between transactions for inclusion in a block, the EIP-1559 TFM loses its UIC property and the tipless mechanism loses (all three notions of) collusion-resilience.

As mentioned above:

- (Theorem 9.1) A relevation principle holds for the global SCP and c-SCP notions: any UIC and MIC TFM that satisfies one of these properties can be simulated by a truthful UIC and MIC TFM that satisfies the same property.
- A relevation principle does not in general hold for the OCA-proof notion: while there are non-truthful TFMs that satisfy UIC, MIC, and OCA-proofness (Section 6.1), there are no truthful TFMs that satisfy UIC, MIC, and OCA-proofness (Theorem 6.10).

The main result in Chung and Shi [CS23] states that, even among randomized TFMs, no TFM satisfies UIC and c-SCP for any $c \geq 1$. Our Theorem 5.6 proves that, even among randomized TFMs, no TFM satisfies UIC, MIC, and global SCP. (Due to the revelation principle mentioned

above, these impossibility results apply to both truthful and non-truthful TFMs.) Our Theorem 6.10 proves the stronger statement that, even among randomized TFMs, no truthful TFM satisfies UIC, MIC, and OCA-proofness.

Reflecting on the competing notions of collusion-resilience, we can observe the following. The c-SCP condition may be particularly appropriate in scenarios where the primary concern is deviations by small coalitions, or in scenarios where users may wish to deviate in ways that exploit other users. The c-SCP condition is also notable in that, together with the UIC condition, it already triggers the impossibility result in [CS23] (without any appeal to MIC). The OCA-proofness condition is distinguished by being the weakest of the three notions (thus leading to the strongest impossibility results) and by allowing the discussion of non-UIC mechanisms.[4] For TFMs that are UIC and MIC, like those studied in this paper, global SCP is arguably the "right" definition—capturing the spirit of OCA-proofness, without any additional technical complications arising from users using different bidding strategies to satisfy UIC and collusion-resilience. Put differently, the UIC and MIC conditions imply that the miner and the users following their intended strategies constitutes a Nash equilibrium; the global SCP condition asserts that this Nash equilibrium is also robust to deviations by the grand coalition, while OCA-proofness only asserts such robustness for a possibly different strategy profile (defined by the intended inclusion rule and some individually rational bidding strategy). From this vantage point, one might view Theorem 5.6 as the main impossibility result in this paper, with Theorem 6.10 serving as a technically challenging extension of the result under still weaker incentive-compatibility conditions.

2 Definitions

2.1 Transaction Fee Mechanism in the Plain Model

In this work, we assume each user has a true value $v \in \mathbb{R}_{\geq 0}$ for its transaction being confirmed on the blockchain, where $\mathbb{R}_{\geq 0}$ denotes non-negative reals. We assume all transactions have an equal size, and we define the *block size* to be the maximal number of transactions that can be included in a block. If there exists a finite number k such that a block can include at most k transactions, we say the block size is *finite*;[5] otherwise, we say the block size is *infinite*. When the block size is finite, we always assume $k \geq 1$. Additionally, we assume that the users' utilities are independent of the order of the transactions in a block. In practice, transactions may have different sizes and the order of the transactions matters. For example, the size is measured in "gas" in Ethereum, and the prices for trading may depend on the order of the transactions. However, for our impossibility results, these assumptions only make them stronger.

A transaction fee mechanism (TFM) consists of the following possibly randomized algorithms:

- **Bidding rule** (executed by the user): takes in the user's true valuation $v \in \mathbb{R}_{\geq 0}$ for getting its transaction confirmed, and outputs a vector of a non-negative number of real-valued bids. Without loss of generality, we assume that at most one bid in the output vector may correspond to the user's actual transaction[6] which has a true value of v (henceforth called the *primary* bid), and the remaining bids are *fake bids* with a true value of 0. We say a

[4] For example, in a first-price auction, the "reference outcome" might be defined by a (non-truthful) bidding strategy that would constitute a Bayes-Nash equilibrium with respect to some prior over user valuations (cf., Corollary 5.12 in [Rou21]).

[5] The finite block size regime in this work and [CS23] corresponds to the case in [Rou21] where the base fee in the EIP-1559 or tipless mechanisms is *excessively low*, i.e. the number of transactions willing to pay the base fee exceeds the maximum block size (cf., Definition 5.6 in [Rou21]).

[6] The blockchain protocol can always suppress conflicting or double-spending transactions.

mechanism is *truthful* if the honest bidding rule asks the user to submit a single primary bid that reports its true value. Otherwise, if the honest bidding rule may ask the user to submit bids other than its true value, the mechanism is said to be non-truthful. Our formulation is more general and admits non-truthful mechanisms, too.[7]

- **Inclusion rule** (executed by the miner): takes a bid vector $\mathbf{b} = (b_1, \ldots, b_t) \in \mathbb{R}^t$ as input, and outputs a subset of $S \subseteq \mathbf{b}$ indicating which bids to be included in the block. When the block size k is *finite*, it must be that $|S| \leq k$.

- **Confirmation rule** (executed by the blockchain): takes as input a block $\mathbf{B}$, i.e., a vector of included bids, and outputs a subset of $\mathbf{B}$ to be confirmed. In general, not all transactions included in the block must be confirmed, and only the confirmed transactions are executed.[8]

- **Payment rule** (executed by the blockchain): given a block $\mathbf{B}$ as input, outputs the payment of each bid. We require *individual rationality*, i.e., for any vector of true values $\mathbf{v} = (v_1, \ldots, v_t)$, if all users execute the honest bidding rule, and we then execute the honest inclusion, confirmation, and payment rules, then the following holds with probability 1:[9] for every user $i \in [t]$, if its primary bid is not confirmed, then its total payment is zero; else if its primary bid is confirmed, then its total payment is at most its true value v_i. Here, a user's total payment is the sum of the payment of all bids it has submitted.

- **Miner revenue rule** (executed by the blockchain): given a block $\mathbf{B}$ as input, outputs how much revenue the miner gets. We require *budget feasibility*, i.e., the miner's revenue is at most the total payment collected from all confirmed bids. When the miner revenue is strictly smaller than the total payment, the difference between the total payment and the miner revenue is said to be *burnt*, i.e. eliminated from circulation.

We say a TFM is *trivial* if the confirmation probability of all transactions is zero for any bid vector assuming the miner honestly follows the inclusion rule; otherwise, it is called *non-trivial*.

A strategic miner or miner-user coalition may deviate from the honest inclusion rule. On the other hand, since the confirmation, payment, and miner revenue rules are executed by the blockchain, they are always implemented honestly.

We focus on mechanisms that are *weakly symmetric*, i.e., mechanisms that do not make use of the bidders' identities or other auxiliary information (e.g., timestamp, transaction metadata), except for tie-breaking among equal bids. More formally, we define weak symmetry as below.

Definition 2.1 (Weak symmetry)**.** A mechanism is called *weakly symmetric* if the mechanism can always be equivalently described in the following manner: given a bid vector $\mathbf{b}$ where each bid may carry some extra information such as identity or timestamp, the honest mechanism always sorts the vector $\mathbf{b}$ by the bid amount first. During the sorting step, if multiple bids have the same amount, then arbitrary tie-breaking rules may be applied, and the tie-breaking can depend on extra information such as timestamp, identity, or random coins. After this sorting step, the inclusion rule and the confirmation rules should depend only on the amount of the bids and their relative position in the sorted bid vector.

[7] Throughout the paper except Section 9, we only focus on bidding rules that output a single bid. In Section 9, we consider general bidding rules that may output multiple bids.

[8] Roughgarden [Rou21] assumes that all included transactions are confirmed. However, Chung and Shi [CS23] show that allowing unconfirmed transactions in a block enlarges the design space. For example, some mechanisms require a block to contain some unconfirmed transactions (see Section 7 in [CS23]).

[9] We can also relax the requirement such that individual rationality holds in expectation. Both the impossibility results (Sections 4, 5 and 6.2) and the revelation principle result (Section 9) continue to hold.

The above weak symmetry definition is equivalent to the following: for two bid vectors $\mathbf{b}$ and $\mathbf{b}'$ of length n that are equivalent up to a permutation, then the distribution of the mechanism's outcomes on $\mathbf{b}$ and $\mathbf{b}'$ are identically distributed — assuming that the outcome is encoded in the form of $\left(\mu, \mathsf{sorted}(\{(v_i, x_i, p_i)\}_{i\in[n]})\right)$ sorted in descending order of v_i, where μ encodes the miner's revenue, each tuple (v_i, x_i, p_i) indicates that some user with bid v_i had the confirmation outcome $x_i \in \{0, 1\}$, and paid p_i.

If we require that the tie-breaking cannot depend on the extra information, we can define *strong symmetry* as follows: for any bid vector $\mathbf{b} = (b_1, \ldots, b_t)$, and for any $i \neq j$ such that $b_i = b_j$, the random variables (x_i, p_i) and (x_j, p_j) must be identically distributed. Throughout the paper, we focus on weakly symmetric mechanisms. Since a strongly symmetric mechanism must also be weakly symmetric, our impossibility results also rule out the existence of strongly symmetric mechanisms.

Strategy space A strategic user can deviate from the honest bidding rule and post an arbitrary bid vector with zero to multiple bids. Without loss of generality, we may assume that in the strategic bid vector, at most one bid can correspond to the user's actual transaction which has a non-zero true value; all other bids must be fake bids with zero true value. A strategic miner can deviate from the honest inclusion rule, and instead create an arbitrary block (subject to the block size limit) that includes any subset of the bid vector as well as any number of fake bids that it chooses to inject. A strategic miner-user coalition can adopt a combination of the above strategies.

Utility and social welfare For a user with true value v, let $x \in \{0, 1\}$ be the indicator of whether its primary bid is confirmed or not, let p denote its total payment, then the user's utility is $x \cdot v - p$. The miner's utility is simply its revenue. The social welfare is defined to be the sum of the utilities of all users and the miner (i.e., the total value of the confirmed transactions, less any burned payments).

Notice that we allow the miner revenue to be smaller than the sum of users' payment, since the coins can be burnt. When calculating the social welfare, the payments among the users and the miner are canceled out, so the social welfare is independent of the payment; however, the amount of burnt coins decreases the social welfare. For example, suppose there is only one user, and let p be the user's payment and q be the amount of burnt coins. In this case, the user's utility is $x \cdot v - p$, the miner revenue is $p - q$, and the social welfare is $(x \cdot v - p) + (p - q) = x \cdot v - q$.

2.2 Incentive Compatibility Notions

Definition 2.2 (User incentive compatible (UIC))**.** A TFM is said to be *user incentive compatible (UIC)*, iff the following holds: for any user i, for any bid vector $\mathbf{b}_{-i}$ that corresponds to the bids posted by all other users, user i's expected utility is always maximized when it follows the honest bidding rule (assuming that the miner executes the inclusion rule honestly). Notice that for a truthful mechanism, the bidding rule is just truth-telling, so following the honest bidding rule simply means submitting the value truthfully; for a non-truthful mechanism, a bidding rule may ask a user to submit a bid other than its true value.

Definition 2.3 (Miner incentive compatible (MIC))**.** A TFM is said to be *miner incentive compatible (MIC)*, iff given any bid vector $\mathbf{b}$, the miner's expected utility is maximized when the miner does not inject any fake bid and creates a block indicated by the honest inclusion rule.

Definition 2.4 (c-side-contract-proof (c-SCP))**.** A TFM is said to be c-side-contract-proof (c-SCP), iff for any coalition C consisting of the miner and between 1 and at most c users, and for any bid vector $\mathbf{b}_{-C}$ that corresponds to the bids posted by all users not in C, the coalition's joint

utility is always maximized when it adopts the honest strategy, i.e., all users in the coalition bid according to the honest bidding rule, and the miner follows the honest inclusion rule.

Definition 2.5 (Global side-contract-proof (global SCP))**.** A TFM is said to be *global side-contract-proof (global SCP)*, iff given any vector of true values $\mathbf{v}$, the expected social welfare is maximized when all the users bid according to the honest bidding rule, and the miner follows the honest inclusion rule, where the maximization is taken over all the coordinated strategies that the coalition consisting of the miner and all users can adopt.

Definition 2.6 (OCA-proof)**.** A TFM is said to be *OCA-proof* iff there exists an individually rational bidding strategy $\sigma : \mathbb{R}_{\geq 0} \to \mathbb{R}$ such that the expected social welfare is maximized when all the users bid according to σ, and the miner creates a block indicated by the inclusion rule, where the maximization is taken over all the coordinated strategies that the coalition consisting of the miner and all users can adopt.

In the definitions above, the expectation is taken over the randomness of the TFM. More explicitly, in Definition 2.2, the expectation is taken over the randomness of the inclusion/confirmation/payment rules; in Definitions 2.3 to 2.6, the expectation is taken over the randomness of the inclusion/confirmation/payment/miner revenue rules.

Note that in the OCA-proofness definition, σ is required to output a single real-valued bid. A canonical example of σ is scaling; that is, $\sigma(v) = \gamma v$ for some $\gamma \in [0, 1]$ (cf., Corollary 5.12 and 5.14 in [Rou21]).

A detailed comparison between c-SCP, global SCP, and OCA-proofness is given in Section A.

3 Preliminary: Myerson's Lemma

If a TFM satisfies UIC, then for any user, its confirmation probability and the payment must satisfy the famous Myerson's lemma [Mye81]. It is an important ingredient for proving our impossibilities, and we review it in this section. Formally, given a bid vector $\mathbf{b} = (b_1, \ldots, b_t)$, let $x_i(\mathbf{b}) \in [0, 1]$ denote the probability that user i's bid is confirmed, and let $p_i(\mathbf{b})$ denote user i's expected payment assuming the miner implements the mechanism honestly. For any vector $\mathbf{b}$, let $\mathbf{b}_{-i}$ denote a vector when we remove i-th coordinate in $\mathbf{b}$. With these notations, Myerson's lemma can be stated as follows.

Lemma 3.1 (Myerson's lemma [Mye81])**.** *For any TFM that satisfies UIC if and only if*

- *For any user i, bid vector $\mathbf{b}_{-i}$, and b_i, b_i' such that $b_i' > b_i$, it must be $x_i(\mathbf{b}_{-i}, b_i') \geq x_i(\mathbf{b}_{-i}, b_i)$.*
- *For any user i, bid vector $\mathbf{b}_{-i}$ from other users, and bid b_i from user i, user i's expected payment must be*

$$p_i(\mathbf{b}_{-i}, b_i) = b_i \cdot x_i(\mathbf{b}_{-i}, b_i) - \int_0^{b_i} x_i(\mathbf{b}_{-i}, t)dt,$$

  with respect to the normalization condition: $p_i(\mathbf{b}_{-i}, 0) = 0$, i.e. user i's payment must be zero when $b_i = 0$.

When the mechanism is deterministic, the confirmation probability x_i is either 0 or 1. In this case, user i's payment can be simplified as

$$p_i(\mathbf{b}_{-i}, b_i) = \begin{cases} \inf\{z \in [0, b_i] : x_i(\mathbf{b}_{-i}, z) = 1\}, & \text{if } x_i(\mathbf{b}_{-i}, b_i) = 1; \\ 0, & \text{if } x_i(\mathbf{b}_{-i}, b_i) = 0. \end{cases}$$

Conceptually, user i must pay the minimal price which makes its bid confirmed.

4 Warmup: Impossibility of UIC + MIC + Global SCP for Deterministic Mechanisms

As a warmup, we first show a finite-block impossibility for UIC + MIC + global SCP for *deterministic* mechanisms. Recall that a TFM is said to be trivial if everyone's confirmation probability is zero for any bid vector assuming the miner follows the inclusion rule. In this case, everyone's utility is always zero in an honest execution. We will show that no non-trivial mechanism can satisfy all three properties simultaneously. Later in Section 5, we extend the impossibility to randomized mechanisms. Due to the revelation principle that we prove in Section 9, if we can prove the impossibility for truthful mechanisms, the impossibility immediately extends to non-truthful mechanisms as well. Therefore, in this section, we shall assume truthful mechanisms.

Lemma 4.1. *For any global SCP mechanism, the confirmed bids must correspond to the highest bids.*

Proof. Suppose in some scenario, Alice bids her true value b and Bob bids his true value $b' < b$; however, Bob's bid is confirmed and Alice's is not. Now, we can have Alice and Bob swap their bids. The miner creates the same block as before in which the position originally corresponding to Bob now has Alice's bid of b'. Since the mechanism is weakly symmetric (Definition 2.1), Alice's bid is confirmed. This way, the social welfare increases by $b - b'$ in comparison with the honest case, and this violates global SCP. □

Lemma 4.2. *For any global SCP mechanism, the amount of burnt coins depends only on the number of confirmed bids.*

Proof. Suppose in two different scenarios, when everyone acts honestly, the blocks made are $\mathbf{B}$ and $\mathbf{B}'$ respectively, the confirmed bids are $\mathbf{b} \subseteq \mathbf{B}$ and $\mathbf{b}' \subseteq \mathbf{B}'$ respectively where $\mathbf{b}$ and $\mathbf{b}'$ are of the same length, and the burnt amount in the two scenarios are q and q' respectively, where $q < q'$. Now, suppose we are actually in the second scenario. A global coalition can adopt the following strategy: create a block identical to $\mathbf{B}$ in which the confirmed bids correspond to the users with the highest true values and the rest can be fake bids. Observe that the social welfare is the sum of the true values of all confirmed bids (where fake bids have a true value of 0) minus the total coins burnt. Therefore, the above strategy achieves strictly higher social welfare than the honest case. □

Theorem 4.3. *No non-trivial deterministic TFM can simultaneously satisfy UIC, MIC, and global SCP when the block size is finite.*

Proof. Since the block size is finite, there must exist a scenario that maximizes the number of confirmed bids. Suppose in this scenario, the block created is $\mathbf{B}$ and the confirmed bids contained in $\mathbf{B}$ are $\mathbf{b} = (b_1, \ldots, b_m)$ sorted from high to low. Since the mechanism is non-trivial, we have $m \geq 1$. Given Lemma 4.2, we can use q_i to denote the amount of burnt coins when exactly i bids are confirmed.

First, we show that in any scenario with at least m users whose true values are sufficiently large w.r.t. $q_0, q_1, \ldots, q_m$, if everyone acts honestly, the mechanism must confirm m bids — by Lemma 4.1, they must also be the m highest bids. Suppose this is not the case, say, $m' < m$ bidders are confirmed — by Lemma 4.1, they must be the m' highest bids. Now, the global coalition can create the block $\mathbf{B}$ where the m confirmed bids correspond to the users with the highest true values, and the rest can just be fake bids. In comparison with the honest case, the

increase in social welfare is $\sum_{i=m'+1}^{m}(v_i - (q_i - q_{i-1}))$ which is positive as long as users' true values v_i's are sufficiently large. This violates global SCP.

Imagine a scenario with at least $m+1$ users whose true values are sufficiently high, and $v_1 \geq \cdots \geq v_m > v_{m+1}$. By Lemma 4.1, it must be $v_1, \ldots, v_m$ are confirmed. Because the mechanism is UIC, by Myerson's lemma, all m confirmed bids must pay v_{m+1}. However, such a mechanism violates MIC, because the miner can pretend that there is a fake bid $v' \in (v_{m+1}, v_m)$. Because the burnt amount is fixed to q_m, the miner will get more revenue if every confirmed bid pays v' instead of v_{m+1}. □

5 Impossibility of UIC + MIC + Global SCP for Randomized Mechanisms

In this section, we extend the finite-block impossibility of UIC + MIC + global SCP to even randomized mechanisms. Recall that a TFM consists of five rules as defined in Section 2.1, and a randomized TFM may use randomness in any of the five rules. Since the confirmation, the payment, and the miner revenue rules are executed by the blockchain, the strategic players can only bias the randomness in and deviate from the bidding rule and the inclusion rule. Again, due to the revelation principle proven in Section 9, it suffices to consider truthful mechanisms.

5.1 Proof Roadmap

A new challenge in the randomized setting is that whether a bid is confirmed becomes probabilistic, and the arguments in Section 4 no longer hold. Here we give a brief roadmap of our new proof. The key idea of the proof is captured in Lemma 5.4, which says that if there are i equal bids $(\underbrace{b, \ldots, b}_{i})$, followed by other bids $(b_{i+1}, \ldots, b_{k+1})$ that are strictly smaller than b, then $b_{i+1}, \ldots, b_{k+1}$ must be unconfirmed with probability 1, and moreover, all the b-bids must pay b when confirmed, i.e., the total user utility is 0 — throughout, we use k to denote the block size. Since Lemma 5.4 imposes a very strong constraint on the TFM, using it to lead to the final impossibility is not too hard (see Theorem 5.6).

We therefore provide some intuition for the proof of Lemma 5.4. Specifically, we use an inductive proof. The base case is the scenario $(\underbrace{b, \ldots, b}_{k}, b_{k+1})$ where $b_{k+1} < b$. Using global SCP and the fact that the block size is only k, we show that the bid b_{k+1} must be unconfirmed with probability 1. Then, we show that the k bids at b must pay b if they are confirmed. To show this, we first argue that in a TFM that is MIC and global SCP, the total user utility is equivalent if b_{k+1} is anything below b (see Lemma 5.3). We then consider a scenario where the $(k+1)$-th bid $b_{k+1} < b$ is arbitrarily close to b, and consider lowering one of the first k bids (say, the first bid) to $b_{k+1} - \delta$ — in this case, we argue that the first bid will become unconfirmed using global SCP. Finally, by Myerson's lemma, we conclude that in the original scenario $(\underbrace{b, \ldots, b}_{k}, b_{k+1} < b)$, the first bid must pay b when confirmed.

Once we prove the base case, we proceed by induction for $i = k-1, k-2, \ldots, 1$. The induction step is similar in structure to the base case, except that we first need to use the induction hypothesis that in the scenario $(\underbrace{b, \ldots, b}_{i}, b_{i+1} < b, \ldots)$, all the b bids pay b when confirmed, to conclude that in the scenario $(\underbrace{b, \ldots, b}_{i-1}, b_i < b, \ldots)$, b_i is unconfirmed with probability 1 — this follows due to

Myerson's lemma. From this point on, the rest of the proof for the induction step is similar to the base case.

5.2 Formal Proofs

In the rest of this section, we present the formal proofs.

Lemma 5.1. *Suppose the mechanism satisfies global SCP. Then, for any bid vector* $\mathbf{b} = (b_1, \ldots, b_t)$*, if* $b_i > b_j$*, it must be that* $x_i(\mathbf{b}) \geq x_j(\mathbf{b})$ *where* $x_i(\mathbf{b})$ *and* $x_j(\mathbf{b})$ *denote user* i *and* j*'s confirmation probabilities, respectively.*

Proof. For the sake of contradiction, suppose there exists two integers i, j such that $b_i > b_j$ and $x_i(\mathbf{b}) < x_j(\mathbf{b})$. The global coalition can swap user i's and user j's bids. Concretely, user i should bid b_j and user j should bid b_i. Since the bid vector is the same as before (though b_i represents user j's bid, and b_j represents user i's bid), the expected burning is the same as before. However, since the mechanism is weakly symmetric, the social welfare increases by $(b_i \cdot x_j(\mathbf{b}) + b_j \cdot x_i(\mathbf{b})) - (b_i \cdot x_i(\mathbf{b}) + b_j \cdot x_j(\mathbf{b})) > 0$, which violates global SCP. □

Lemma 5.2. *Suppose the mechanism satisfies global SCP and let the block size be* k*. Consider any bid vector satisfying* $b_1 \geq b_2 \geq \ldots \geq b_k > b_{k+1}$*. Then, it must be that* b_{k+1} *has 0 probability of confirmation.*

Proof. Suppose for the sake of contradiction that b_{k+1} has non-zero probability of confirmation. This means it must be included in the block with some non-zero probability p, and conditioned on being included, it has a non-zero probability p' of confirmation. A global coalition whose true values are $b_1, \ldots, b_{k+1}$ can play the following strategy. First, run the honest mechanism in its head including flipping any coins required by the inclusion rule assuming everyone bids honestly, and let $\mathbf{B}$ be the resulting block. If $b_{k+1} \notin \mathbf{B}$, just have everyone bid truthfully and build the block $\mathbf{B}$. Else, if $b_{k+1} \in \mathbf{B}$, then let $b_i > b_{k+1}$ be one bid that is left out of the block — such a bid must exist because the block size is only k. The coalition will have the i-th user actually bid b_k and build the same block $\mathbf{B}$ where the b_{k+1} actually corresponds to user i's strategic bid. With this strategy, the coalition's expected gain in utility is $p \cdot p' \cdot (b_i - b_{k+1}) > 0$, which violates global SCP. □

Lemma 5.3. *Suppose the mechanism satisfies MIC and global SCP. Let* $\mathbf{a}$ *be an arbitrary vector of positive length, and let* $\mathbf{b}, \mathbf{b}'$ *arbitrary vectors —* $\mathbf{b}$ *and* $\mathbf{b}'$ *may or may not be of the same length, and their lengths are allowed to be* 0*. Consider two scenarios with bid vectors* $(\mathbf{a}, \mathbf{b})$ *and* $(\mathbf{a}, \mathbf{b}')$ *respectively. Suppose that* $\mathbf{b}$ *and* $\mathbf{b}'$ *have 0 confirmation probability in each of the two scenarios, respectively. Then, both scenarios enjoy the same expected miner utility, total social welfare, and total user utility.*

Proof. We first prove that expected miner utility is the same in both scenarios. Suppose this is not true, and without loss of generality, suppose expected miner utility is higher in scenario 1. Then, the miner can ignore the bids $\mathbf{b}$, inject the fake bids $\mathbf{b}'$, pretend that the bid vector is $(\mathbf{a}, \mathbf{b}')$, and run the honest mechanism. Since the confirmation probability of $\mathbf{b}'$ is 0, the miner need not pay any cost for the fake bids. Therefore, the miner gets higher expected utility by taking the above strategy which violates MIC.

The proof of total social welfare is similar. Suppose, without loss of generality, that the expected total social welfare in scenario 1 is higher. Then, the global coalition can inject fake bids $\mathbf{b}'$ and pretend that the bid vector is $(\mathbf{a}, \mathbf{b}')$, thus allowing it to increase its expected social welfare. This violates global SCP.

The equivalence in total user utility follows directly from the above, since total user utility is the difference between the social welfare and the miner utility. □

Lemma 5.4. *Suppose the mechanism satisfies UIC, MIC, and global SCP, and the block size is k. Let $\mathbf{b} = (b_1, \ldots, b_{k+1})$ be any vector where $b_1 = b_2 = \cdots = b_i > b_{i+1} > \cdots > b_{k+1}$. Then, under the bid vector $\mathbf{b}$, the following hold: 1) users $i+1, \ldots, k+1$ have 0 confirmation probability; and 2) the total user utility is 0 assuming every one is bidding their true value.*

Proof. We prove both statements by induction. Henceforth, given a bid vector $\mathbf{b} = (b_1, \ldots, b_{k+1})$, we use $x_i(\mathbf{b})$ to denote the confirmation probability of the i-th user under $\mathbf{b}$.

Base case. The base case is when $i = k$; that is, there are $k+1$ users with valuations $\mathbf{b} = (b_1, \ldots, b_{k+1})$ where $b_1 = \cdots = b_k > b_{k+1}$. By Lemma 5.2, we have $x_{k+1}(\mathbf{b}) = 0$. Next, we show that the joint utility of all users must be zero. By Lemma 5.3 and Lemma 5.2, the total user utility under $\mathbf{b} = (\underbrace{b, \ldots, b}_{k}, b_{k+1})$ is the same as under $\mathbf{b}' = (\underbrace{b, b, \ldots, b}_{k}, b - \delta)$ for any arbitrarily small $\delta > 0$ — in both cases, the last user must have 0 confirmation probability. Now, consider the scenario $\mathbf{b}' = (\underbrace{b, b, \ldots, b}_{k}, b - \delta)$. If one of the b-bidders denoted $i \in [k]$ lowered its bid to less than $b - \delta$ (and everyone else's bids stay the same), its confirmation probability would become 0 by Lemma 5.2. Since the mechanism is UIC, by Myerson's lemma, under $(b, b, \ldots, b, b - \delta)$, user i's expected payment must be at least $(b - \delta) \cdot x_i(\mathbf{b}')$. Therefore, user i's expected utility is at most $\delta \cdot x_i(\mathbf{b}')$. Therefore, the expected total user utility under $\mathbf{b}'$ is at most $k \cdot \delta \cdot x_i(\mathbf{b}') \leq k \cdot \delta$. This means that under $\mathbf{b} = (b_1, \ldots, b_{k+1})$, the expected total user utility is at most $k \cdot \delta$ for any arbitrarily small $\delta > 0$, i.e., the expected total user utility is at most 0. Finally, by the individual rationality of the payment rule, each user's utility is non-negative when it bids truthfully. Therefore, the expected total user utility is 0.

Inductive step. Fix any $i \in \{2, \ldots, k\}$, suppose the lemma statement holds. We will show that the lemma also holds for any vector $\mathbf{b} = (b_1, \ldots, b_{k+1})$ where $b_1 = b_2 = \cdots = b_{i-1} > b_i > b_{i+1} > \cdots > b_{k+1}$.

We first show that all users $i, \ldots, k+1$ are unconfirmed. Consider the vector $\underbrace{b, b, \ldots, b}_{i}, b_{i+1}, \ldots, b_{k+1}$. By our induction hypothesis, i users can have non-zero confirmation probability and moreover, the total user utility is 0. Because of individual rationality, each user's utility must be non-negative, so every user's utility is also 0. In this case, whenever a bid is confirmed, the payment is equal to the bid (and hence the true value). By UIC and Myerson's lemma, it means that if the i-th user lowers its bid to $b_i < b$ (which becomes the scenario $\mathbf{b}$ we care about), then user i's confirmation probability becomes 0. By Lemma 5.1, users $i+1, \ldots, k+1$ must have 0 confirmation probability under $\mathbf{b}$ too.

Next, we show that the joint utility of all users is zero under $\mathbf{b}$ and assuming everyone bids truthfully. By Lemma 5.3, the total user utility under $\mathbf{b}$ is the same as under the scenario $\mathbf{b}' = (\underbrace{b, b, \ldots, b}_{i-1}, b - \delta, b - 2\delta, \ldots, b - (k+2-i)\delta)$ for an arbitrarily small $\delta > 0$ — since in both scenarios, users $i, \ldots, k+1$ have 0 confirmation probability. Under $\mathbf{b}'$, if any b-bidders denoted j lowers its bid to less than $b - (k+2-i)\delta$, its confirmation probability becomes 0 by Lemma 5.2. By UIC and Myerson's lemma, j's expected payment under $\mathbf{b}'$ is at least $(b - (k+2-i)\delta) \cdot x_j(\mathbf{b}')$. This means that the total user utility under $\mathbf{b}'$ is at most $(k+2-i)\delta \cdot x_j(\mathbf{b}') \cdot (i-1) \leq k^2\delta$. Therefore, under

$\mathbf{b}$, the total user utility is at most $k^2\delta$ for any arbitrarily small $\delta > 0$. This means that the total user utility under $\mathbf{b}$ is at most 0. Finally, by the individual rationality of the payment rule, each user's utility is non-negative when it bids truthfully. Therefore, the expected total user utility is 0. □

Lemma 5.5. *Suppose the mechanism satisfies UIC, MIC, and global SCP, and the block size is k. Let a be any positive real number. Consider a scenario with only one bid a. Then, the only user's utility is zero assuming it bids its true value.*

Proof. Fix $a_1 > 0$. Consider any vector $\mathbf{a} = (a_1, a_2, \ldots, a_{k+1})$ such that $a_1 > a_2 > \cdots > a_{k+1}$. By Lemma 5.4, under $\mathbf{a}$, all users except the first user must have 0 confirmation probability, and moreover, the first user has utility 0. By Lemma 5.3, the total expected user utility under $\mathbf{a}$ and under a single bid a_1 are the same. Therefore, under a single bid a_1, the user's utility is also 0. □

Theorem 5.6. *No non-trivial, possibly randomized TFM can simultaneously satisfy UIC, MIC, and global SCP when the block size is finite.*

Proof. We will show that under any sufficiently large a, the confirmation probability under a single bid a is non-zero. If we can show this, then we can show a contradiction to UIC. Specifically, consider $b > a$ and both sufficiently large. By Lemma 5.5, if there is only one user with true value b, its utility is zero when it bids truthfully. However, the user can underbid a. Since the confirmation probability is non-zero and the payment is at most a, the user enjoys positive utility, which violates UIC.

We now show that the confirmation probability under a single bid a for a sufficiently large a is non-zero assuming the miner honestly follows the inclusion rule. Suppose this is not true for the sake of a contradiction; that is, when the bid vector is just a and the miner is honest, no one has positive confirmation probability, so the social welfare is zero. Since the mechanism is non-trivial, there must exist a scenario $\mathbf{b} = (b_1, \ldots, b_t)$ such that some user i has positive confirmation probability $p(\mathbf{b})$. Let $a > \frac{1}{p(\mathbf{b})} \cdot \sum_{j=1}^{t} b_j$. Then, imagine that there is only one user with true value a. The global coalition can replace the user's primary bid with b_i, inject the fake bids $\mathbf{b}_{-i}$, and pretend that the bid vector is $\mathbf{b}$. This way, the user would get positive confirmation probability $p(\mathbf{b})$. Since its payment is at most $\sum_{j=1}^{t} b_j$, the user's expected utility (and thus the social welfare) is positive. This violates global SCP.

□

6 Feasibility and Impossibility of UIC + MIC + OCA-Proof

We can generalize the proof in Section 5, and rule out UIC, MIC, and OCA-proof (rather than global SCP) for truthful mechanisms. Recall that for a truthful mechanism, the difference between OCA-proof and global SCP is that global SCP insists that the optimal strategy of the global coalition is the truthful strategy, whereas OCA-proofness allows it to be some other strategy in which each user acts independently and bids the outcome of some function $\sigma(\cdot)$.

Interestingly, if we allow the bidding rule to be not truth-telling, i.e. considering non-truthful mechanisms, we can have a mechanism that satisfies UIC, MIC, and OCA-proof. We present the feasibility for non-truthful mechanisms in Section 6.1, and we prove the impossibility of UIC + MIC + OCA-proof for truthful mechanisms in Section 6.2. Notice that because of the feasibility in Section 6.1, we must require the bidding rule to be truth-telling to reach an impossibility in Section 6.2.

6.1 A Non-Truthful Mechanism with UIC + MIC + OCA-Proof

The rationale of the design is to signal to the mechanism when everyone is adopting the globally optimal strategy σ (as opposed to the bidding rule used to establish UIC). When the mechanism detects that everyone is behaving according to σ, it adopts a different behavior to optimize social welfare. We use the range $[0, 1)$ to encode the actual bid, and use the range $[1, \infty)$ for signalling. While the resulting mechanism is somewhat contrived and not necessarily meaningful from a practical point of view, it clarifies which notions of collusion-resilience most accurately capture the intended modeling goals and illustrates some technical challenges involved in the proof in Section 6.2. Consider the following TFM:

- **Globally optimal strategy** $\sigma(v)$: Given a true value v, output a bid $v + 1$.
- **Bidding rule**: Given a true value v, output a bid $1/(v + 2)$.
- **Inclusion rule**: Let S be the set of all pending bids that are in $[0, 1)$. If $|S| > k$, then randomly select k bids from S to include. If $1 \leq |S| \leq k$, then include all bids in S. If $|S| = 0$, choose the top up to k bids to include.
- **Confirmation, payment, and miner revenue rules**: All included bids are confirmed. Each confirmed bid pays nothing, and the miner gets nothing.

Obviously, this mechanism is non-trivial.

Claim 6.1. *The above mechanism satisfies UIC, MIC, and OCA-proofness.*

Proof. For UIC, notice that if a user follows the bidding rule, its bid is always in $[0, 1)$. If there is no bid in $[0, 1)$ before a user submits its bid, then bidding $1/(v + 2)$ always guarantees user's bid to be included and confirmed, where v denote the true value. If there is already some bids in $[0, 1)$ before a user submits its bid, then bidding $1/(v + 2)$ is a dominant strategy since it guarantees the user's bid is added to S, the set of all bids in $[0, 1)$, which is the best a user can do. Next, MIC holds since the miner revenue is always zero. Finally, if all users follow the globally optimal strategy σ, everyone's bid is at least 1. The honest inclusion rule will include the top up to k bids, which maximizes the social welfare. Thus, OCA-proofness holds. □

Remark 6.2. We can try to apply revelation principle, and bake the bidding rule into the mechanism so that the resulting mechanism is truthful. For example, whenever seeing a bid b, the miner and the mechanism view it as $1/(b+2)$. The modified mechanism, however, does not satisfy OCA-proofness anymore when the number of users is larger than the block size, since the miner should choose k users with the highest true values instead of the random selection as indicated by the inclusion rule. This is not a coincidence: in the next section, we show that it is impossible to have a non-trivial truthful mechanism satisfying UIC, MIC, and OCA-proofness.

6.2 Impossibility of UIC + MIC + OCA-Proof for Truthful Mechanisms

The structure of the proof in this section is similar to the proof in Section 5. In fact, Lemmas 6.3, 6.4 and 6.7 to 6.9 and Theorem 6.10 are the analogs of Lemmas 5.1 to 5.5 and Theorem 5.6, respectively, except that we need to work on the images of σ in order to apply OCA-proofness. Before diving into the proof, we make a few remarks.

- A key step in Section 5 is to prove a user with true value v must have zero utility, and we prove it by making an unconfirmed bid arbitrarily close to v (see Lemma 5.4). To extend the similar idea to OCA-proofness, we require σ to be strictly increasing and continuous. In fact, we show that being strictly increasing is a consequence of UIC and OCA-proofness (Lemma 6.5), and being strictly increasing implies that there exists a point at which σ is continuous (Theorem 6.6).
- As we have seen a feasibility for non-truthful mechanisms in Section 6.1, we must require the bidding rule to be truth-telling to reach an impossibility. Indeed, the proofs of Lemmas 6.5, 6.8 and 6.9 and Theorem 6.10 rely on Myerson's lemma, and Myerson's lemma requires the mechanism to be truthful. Notice that the mechanism in Section 6.1 does not satisfy the requirements of Myerson's lemma. If we apply the revelation principle to make the mechanism become truthful so that Myerson's lemma could apply, it breaks OCA-proofness (see Theorem 6.2).
- Notice that OCA-proofness requires σ to output a single real-valued bid. In fact, if we allow σ to output multiple bids, we can have a (somewhat contrived) mechanism that satisfies UIC, MIC, and OCA-proof (see Section 7.2). In particular, Lemma 6.5 requires σ to only output a single real number.
- The original definition of OCA-proofness in [Rou21] also requires σ to be individually rational, but this will not be needed for the impossibility proof.

Given any vector $\mathbf{b} = (b_1, \ldots, b_t)$ of the length t, let $\sigma(\mathbf{b})$ denote the element-wise application of σ, i.e. $\sigma(\mathbf{b}) = (\sigma(b_1), \ldots, \sigma(b_t))$.

Lemma 6.3. *Suppose the mechanism satisfies OCA-proofness with the globally optimal strategy σ. Then, for any bid vector $\sigma(\mathbf{b}) = (\sigma(b_1), \ldots, \sigma(b_t))$, if $b_i > b_j$, it must be that the confirmation probability of the bid $\sigma(b_i)$ is at least as high as that of $\sigma(b_j)$.*

Proof. Imagine that t players each have true values $b_1, \ldots, b_t$, and they all bid according to the globally optimal strategy σ. Let x_i, x_j denote the confirmation probabilities of $\sigma(b_i)$ and $\sigma(b_j)$ respectively. Suppose for the sake of contradiction that $b_i > b_j$, and $x_j > x_i$. The global coalition can engage in the following OCA: have users i and j swap their bids, that is, user i who has true value b_i bids $\sigma(b_j)$, and user j who has true value b_j bids $\sigma(b_i)$. The set of bids is the same as before, therefore, the expected burning amount does not change. The expected social welfare thus increases by $b_i x_j + b_j x_i - (b_i x_i + b_j x_j) = (b_i - b_j)(x_j - x_i) > 0$. This violates OCA-proofness and the fact that σ is the social-welfare-maximizing-strategy of the global coalition. □

Lemma 6.4. *Suppose the mechanism satisfies OCA-proofness with the globally optimal strategy σ, and let the block size be k. Given $b_1 \geq b_2 \geq \ldots \geq b_k > b_{k+1}$, then, under the bid vector $\sigma(b_1), \sigma(b_2), \ldots, \sigma(b_k), \sigma(b_{k+1})$, it must be that $\sigma(b_{k+1})$ has 0 probability of confirmation.*

Proof. Imagine that $k+1$ players each have true values $b_1, \ldots, b_{k+1}$, and they all bid according to the globally optimal strategy σ. It suffices to show the following: suppose $\sigma(b_{k+1})$ has a non-zero probability p of being included in the block, then conditioned on being included, its probability of confirmation must be 0. Suppose this is not true, and conditioned on being included, $\sigma(b_{k+1})$ has probability $p' > 0$ of being confirmed. Then, the global coalition can adopt the following OCA: run the honest inclusion rule in its head, including flipping any coins required by the inclusion rule, and let $\mathbf{B}$ denote the resulting block. If $\sigma(b_{k+1})$ is not included in $\mathbf{B}$, simply output the block $\mathbf{B}$. Else if $\sigma(b_{k+1})$ is included in $\mathbf{B}$, then let $\sigma(b_i)$ be some bid that is left out — such a bid must exist

because the block size is only k. Now, have the i-th user bid $\sigma(b_{k+1})$ and make an actual block $\mathbf{B}$ where the bid $\sigma(b_{k+1})$ corresponds to user i's bid. The expected increase in social welfare relative to everyone adopting σ is $p \cdot p' \cdot (b_i - b_{k+1}) > 0$, which violates OCA-proofness and the fact that σ is the social-welfare-maximizing strategy. □

Lemma 6.5. *Suppose a truthful mechanism satisfies UIC and OCA-proofness with the globally optimal strategy σ, and let the block size be k. Then, there exists a constant c^* such that for any real numbers z, z' satisfying $z > z' > c^*$, it must be $\sigma(z) > \sigma(z')$. Conceptually, σ must be strictly increasing for large inputs.*

Proof. Since the mechanism is non-trivial, there exists a bid vector $(b_1, \ldots, b_t)$ such that the total confirmation probability is positive. Let b_i be the bid with the highest confirmation probability among $b_1, \ldots, b_t$. Let x be the confirmation probability of b_i and let c^* be any real number larger than $\sum_{j=1}^{t} b_j/x$. We will show that for any $z > z' > c^*$, $\sigma(z) > \sigma(z')$.

Consider a scenario where there are $k+1$ users all with the true value z. Suppose they form a global coalition and follow the bidding strategy σ so that the bid vector is $\mathbf{s} = (\underbrace{\sigma(z), \ldots, \sigma(z)}_{k+1})$. There are two possible cases.

First, there exists a user j with positive confirmation probability assuming the miner honestly follows the inclusion rule and select from the bid vector $\mathbf{s}$. Because the mechanism satisfies UIC, by Myerson's lemma, if user j increases its bid, the confirmation probability must be non-decreasing. Thus, if $\sigma(z) \leq \sigma(z')$, and user j increases its bid from $\sigma(z)$ to $\sigma(z')$, its confirmation probability is still positive. However, if user j bids $\sigma(z')$, the resulting bid vector will become $(\underbrace{\sigma(z), \ldots, \sigma(z)}_{k}, \sigma(z'))$. By Lemma 6.4, the bid $\sigma(z')$ has zero confirmation probability, which leads to a contradiction. Consequently, it must be $\sigma(z) > \sigma(z')$.

Second, if everyone's confirmation probability is zero assuming the miner honestly follows the inclusion rule and select from the bid vector $\mathbf{s}$, following σ will lead to zero social welfare. We argue that this case is impossible by showing an alternative strategy of the global coalition that achieves positive social welfare, contradicting the fact that σ is social-welfare-maximizing. Specifically, the alternative strategy works as follows: the global coalition pretends that the bid vector is $(b_1, \ldots, b_t)$ where the bid b_i corresponds to the primary bid of any user, say, the first user. Recall that the first user's true value is z. Thus, social welfare is at least $z \cdot x - p$, where p is the expected payment of all users. Since b_i has the highest confirmation probability among $b_1, \ldots, b_t$, p is at most $x \cdot \sum_{j=1}^{t} b_j$. Because $z > c^* > \sum_{j=1}^{t} b_j/x$, the social welfare is positive.

□

Corollary 6.6. *Suppose a truthful mechanism satisfies UIC and OCA-proofness with the globally optimal strategy σ, and let the block size be k. Then, there exists a constant c^* such that for any real number $a > c^*$, there exists a real number $a^* \geq a$ at which σ is continuous.*

Proof. A well-known fact is that any function that is monotone must have only countably many discontinuities within any interval $[x_1, x_2]$ where $x_2 > x_1$; further, all such discontinuities must be jump discontinuities. By Lemma 6.5, there exists a constant c^* such that σ is monotone for all inputs larger than c^*. Thus, for any $a > c^*$, within the interval $[a, a+1]$, σ is monotone and there must exist a point a^* such that σ is continuous at a^*. □

Lemma 6.7. *Suppose the mechanism satisfies MIC and OCA-proofness with the globally optimal strategy σ. Let $\mathbf{a}, \mathbf{b}, \mathbf{b}'$ be arbitrary vectors. Consider two scenarios $(\sigma(\mathbf{a}), \sigma(\mathbf{b}))$ and $(\sigma(\mathbf{a}), \sigma(\mathbf{b}'))$.*

Suppose that $\sigma(\mathbf{b})$ and $\sigma(\mathbf{b}')$ have 0 confirmation probability in each of the two scenarios, respectively. Then, both scenarios enjoy the same expected miner utility, social welfare, and total user utility.

Proof. The equivalence in expected miner utility follows directly from MIC and the fact that the miner can change from $\mathbf{b}$ to $\mathbf{b}'$ for free, or vice versa, since these bids are unconfirmed.

The equivalence in social welfare follows from OCA-proofness, and the fact that σ maximizes social welfare. Suppose, for example, the social welfare under $(\sigma(\mathbf{a}), \sigma(\mathbf{b}))$ is strictly greater than $(\sigma(\mathbf{a}), \sigma(\mathbf{b}'))$. Then, under the scenario where the users' true values are $(\mathbf{a}, \mathbf{b}')$, everyone following σ will not maximize social welfare, since there is an OCA that have the users bid $(\sigma(\mathbf{a}), \sigma(\mathbf{b}))$ instead which strictly increases the social welfare.

The equivalence in total user utility follows directly from the above, since total user utility is the difference between the social welfare and the miner utility. □

Lemma 6.8. *Suppose a truthful mechanism satisfies UIC, MIC and OCA-proofness with the globally optimal strategy σ, and the block size is k. Then, there exists a constant c^* such that for any $i \in [k]$, the following holds.*

- *Consider a scenario with the arbitrary bid vector $\mathbf{s} = \underbrace{\sigma(b), \ldots, \sigma(b)}_{i}, \sigma(b_{i+1}), \ldots, \sigma(b_{k+1})$, where $b > b_{i+1} > b_{i+2} > \ldots > b_{k+1} > c^*$. Then, if σ is continuous at b, the bids $\sigma(b_{i+1}), \ldots, \sigma(b_{k+1})$ have 0 confirmation probability.*
- *Moreover, if there are $k+1$ users with the true value $\mathbf{s}$ and σ is continuous at b, the expected total user utility is 0 when all users bid truthfully.*

Proof. By Lemma 6.5, there exists a constant c^* such that for any real numbers z, z' satisfying $z > z' > c^*$, it must be $\sigma(z) > \sigma(z')$. Throughout this proof, we assume $b_{k+1} > c^*$ so that σ is strictly increasing for any input larger than b_{k+1}. By Theorem 6.6, there exists a real number $b > b_{k+1}$ such that σ is continuous at b, and we choose b such that this condition holds. We will prove the lemma by induction.

Base case. The base case is when $i = k$; that is, there are $k+1$ users with the true values $\mathbf{s} = (\underbrace{\sigma(b), \ldots, \sigma(b)}_{k}, \sigma(b_{k+1}))$. Notice the true value is $\sigma(b)$ instead of b for the first k users. By Lemma 6.4, $\sigma(b_{k+1})$ has 0 confirmation probability. By Lemma 6.7, the total user utility under $\mathbf{s}$ is the same as under $\mathbf{s}' = (\underbrace{\sigma(b), \ldots, \sigma(b)}_{k}, \sigma(b-\delta))$ where $\delta > 0$ can be arbitrarily small — since in both scenarios, the last bid is unconfirmed. Now, suppose the bid vector is $\mathbf{s}'$. Fix any user with the true value $\sigma(b)$, say the first user. If it lowered its bid to anything less than $\sigma(b-\delta)$, its confirmation probability must be 0 by Lemma 6.4. Since the mechanism is UIC, by Myerson's lemma, the payment of the first user is at least $\sigma(b-\delta) \cdot x_1(\mathbf{s}')$. In this case, the first user's utility is at most $(\sigma(b) - \sigma(b-\delta)) \cdot x_1(\mathbf{s}') \leq \sigma(b) - \sigma(b-\delta)$. Therefore, the total user utility under $\mathbf{s}'$ (and hence under $\mathbf{s}$) is at most $k \cdot (\sigma(b) - \sigma(b-\delta))$. Since σ is continuous at b, and the above holds for any $\delta > 0$, the total user utility under $\mathbf{s}$ is at most 0. By the individual rationality of the payment rule, each user's utility is non-negative when it bids truthfully. Therefore, the expected total user utility is 0.

Inductive step. Fix any $i \in \{2, \ldots, k\}$, and suppose the lemma holds. We want to show that the lemma also holds for any $\mathbf{s}_{i-1} = (\underbrace{\sigma(b), \ldots, \sigma(b)}_{i-1}, \sigma(b_i), \sigma(b_{i+1}), \ldots, \sigma(b_{k+1}))$. Start from the scenario $\mathbf{s}_i = (\underbrace{\sigma(b), \ldots, \sigma(b)}_{i}, \sigma(b_{i+1}), \ldots, \sigma(b_{k+1}))$, and imagine that the i-th bid lowers from $\sigma(b)$ to $\sigma(b_i)$. Since the i-th user has 0 utility under $\mathbf{s}_i$, by UIC and Myerson's lemma, its confirmation probability must become 0 when it lowers its bid to $\sigma(b_i) < \sigma(b)$. By Lemma 6.3, the bids $\sigma(b_{i+1}), \ldots, \sigma(b_{k+1})$ must have 0 confirmation probability too.

We next show that under $\mathbf{s}_{i-1}$, total user utility is 0. Consider another bid vector $\mathbf{s}' := (\underbrace{\sigma(b), \ldots, \sigma(b)}_{i-1}, \sigma(b-\delta), \ldots, \sigma(b-(k+2-i)\delta))$ for some $\delta > 0$. By the argument above, the bids numbered $i+1, \ldots, k+1$ are not confirmed. By Lemma 6.7, the total user utility under $\mathbf{s}_{i-1}$ is the same as under the scenario $\mathbf{s}'$. In scenario $\mathbf{s}'$, if any of the $\sigma(b)$ bidders, say the first bidder, lowers its bid from $\sigma(b)$ to anything less than $\sigma(b-(k+2-i)\delta)$, its confirmation probability becomes 0 by Lemma 6.4. By UIC and Myerson, the first user's payment under $\mathbf{s}'$ is at least $\sigma(b-(k+2-i)\delta) \cdot x_1(\mathbf{s}')$, and its utility is at most $(\sigma(b) - \sigma(b-(k+2-i)\delta)) \cdot x_1(\mathbf{s}') \leq \sigma(b) - \sigma(b-(k+2-i))\delta$. Therefore, the total user utility under $\mathbf{s}'$ (and also under $\mathbf{s}_{i-1}$) is upper bounded by $k \cdot (\sigma(b) - \sigma(b-(k+2-i)\delta))$. Since σ is continuous at b and the above holds for any arbitrarily small $\delta > 0$, it means that the total user utility under $\mathbf{s}_{i-1}$ is at most 0. By the individual rationality of the payment rule, each user's utility is non-negative when it bids truthfully. Therefore, the expected total user utility is 0. □

Lemma 6.9. *Suppose a truthful mechanism satisfies UIC, MIC and OCA-proofness with the globally optimal strategy σ, and the block size is k. Then, there exists a constant c^* such that for any $a^* > c^*$, if there is only a single user with the true value $\sigma(a^*)$ and σ is continuous at a^*, its utility is zero when it bids truthfully.*

Proof. By Lemma 6.8, there exists a constant c^* such that under the scenario $\mathbf{a} := (\sigma(a^*), \sigma(a_2), \ldots, \sigma(a_{k+1}))$ where $a^* > a_2 > \ldots > a_{k+1} > c^*$, the utility of the first user is 0 assuming it is bidding truthfully. Moreover, all users except the first user must have 0 confirmation probability. By Lemma 6.7, the total expected user utility under $\mathbf{a}$ and under a single bid $\sigma(a^*)$ are the same. Therefore, under a single bid $\sigma(a^*)$, the user's utility is also 0. □

Theorem 6.10. *No non-trivial, possibly randomized truthful TFM can simultaneously satisfy UIC, MIC, and OCA-proofness when the block size is finite.*

Proof. We will show that under any sufficiently large a at which σ is continuous, the confirmation probability under a single bid $\sigma(a)$ is non-zero. If we can show this, then we can show a contradiction to UIC and the Myerson's lemma. Specifically, consider $b > a$ which both are sufficiently large and at which σ is continuous (Theorem 6.6 guarantees such a, b exist). By Lemma 6.9, under both scenarios, the user utility is 0 (assuming the bid, $\sigma(a)$ or $\sigma(b)$, is the same as the user's true value). By Lemma 6.5, for sufficiently large a, we have $\sigma(b) > \sigma(a)$. However, by Myerson's lemma, if the confirmation probabilities under a single $\sigma(a)$ and under a single $\sigma(b)$ are both non-zero, it cannot be that in both cases the user's utility is 0. Otherwise, the user with the true value $\sigma(b)$ can underbid $\sigma(a)$. Since the confirmation probability is non-zero and the payment is at most $\sigma(a)$, the user enjoys positive utility, which violates UIC.

We now show that the confirmation probability under a single bid $\sigma(a)$ for a sufficiently large a is non-zero assuming the miner honestly follows the inclusion rule. Suppose this is not true for the sake of a contradiction; that is, when the bid vector is just $\sigma(a)$ and the miner is honest, no one has

positive confirmation probability, so the social welfare is zero. Since the mechanism is non-trivial, there exists a scenario $\mathbf{b} = (b_1, \ldots, b_t)$ such that the total confirmation probability is non-zero, and let $x > 0$ be the confirmation probability of some bid b_i in $\mathbf{b}$. Consider $a > \sum_{j\in[n]} b_j/x$, and suppose there is only one user with the true value a. We will show that the strategy $\sigma(a)$ does not maximize social welfare which contradicts to the fact that σ is social-welfare-maximizing-strategy. Specifically, the user can post the bids $(b_1, \ldots, b_t)$ instead where position i corresponds to the user's primary bid and all others are fake bids. Note that under $\mathbf{b}$, the total user payment is at most $\sum_{j\in[n]} b_j$. Since the user's true value is greater than $\sum_{j\in[n]} b_j/x$ and its confirmation probability is x, the total user utility (and thus the social welfare) is positive under this new bidding strategy. □

Remark 6.11. Notice that Chung and Shi [CS23] showed that no TFM can simultaneously satisfy UIC and 1-SCP when the block size is finite. Their impossibility does not rely on MIC, while MIC is necessary for proving Theorem 6.10. Specifically, consider *the second-price auction without burning* defined as follows. Given the block size k, the miner includes the top up to k bids in the block. If the number of included bids is less than k, then all included bids are confirmed, and pay nothing. If the number of included bids is k, then the top $k-1$ bids are confirmed, while the k-th bid is unconfirmed (break ties arbitrarily). All confirmed bids pay the k-th bid, and all the payments go to the miner. This mechanism satisfies UIC since the confirmation and the payment satisfy the requirements of Myerson's lemma. The mechanism also satisfies OCA-proofness when σ is just bidding truthfully since the confirmed bids are always the ones with the highest true values, which maximizes the social welfare. However, it does not satisfy MIC, since the miner is incentivized to inject a fake bid to increase the k-th bid.

7 How to Circumvent the Impossibilities

In this section, we discuss possible ways to circumvent the impossibilities. Recall that in the definition of OCA-proofness, we require all users act independently in the globally optimal strategy σ, and σ should only output a single bid. If we remove either of the two requirements, we can have mechanisms that satisfies UIC, MIC and OCA-proofness, where we present in Sections 7.1 and 7.2. These mechanisms are somewhat contrived and not obviously relevant for real-world blockchain protocols; the primary point of these mechanisms is to "stress-test" the definitions, and these examples highlight the subtlety of the modeling and help us to have a better understanding of the trade-offs between various notions of collusion-resilience. In Section 7.3, we consider an even weaker definition, called *inclusion-rule-respecting*, which might be the weakest but still meaningful definition that captures the notion of "no way to steal from the protocol." Later in Section 8, we introduce a new model that captures the use of trusted hardware and formalize a Bayesian notion of incentive compatibility. We show that, in this model, it is possible to construct a TFM that satisfies Bayesian incentive compatibility.

7.1 Allowing the Globally Optimal Strategy to Coordinate

OCA-proofness requires that all users act independently in the globally optimal strategy σ. However, if we remove this restriction, then the following auction would simultaneously satisfy UIC, MIC, and OCA-proof (with the aforementioned relaxation).

Posted price auction with random selection and total burning. The mechanism is parametrized with some reserve price $r > 0$ and block size k. Any bid that is at least r is considered eligible.

Randomly select up to k bids to include in the block. All included bids are confirmed and they each pay r. All payment is burnt and the miner receives nothing. It is not hard to see that the above mechanism satisfies UIC and MIC. It also satisfies OCA-proofness with the aforementioned relaxation since the global coalition can just have the users with the top k values bid more than r, while everyone else bids 0.

Note that allowing all users to coordinate bids does not trivialize the definition, as this relaxed OCA-proofness definition still requires that the miner follow the TFM's inclusion rule honestly. For example, in the auction above, the miner has to select k bids uniformly at random among all eligible bids.

7.2 Allowing the Globally Optimal Strategy to Output Multiple Bids

We next show that when we allow the globally optimal strategy $\boldsymbol{\sigma}$ to output multiple bids, then we can actually construct a truthful mechanism that satisfies UIC, MIC, and OCA-proofness. Similar to Section 6.1, we try to signal to the mechanism when everyone is adopting the globally optimal strategy σ. In this mechanism, since σ can output multiple bids, users can signal through fake bids. We assume the miner and the mechanism can distinguish whether a set of bids come from the same user. In practice, this can be implemented by checking whether these bids are signed by the same public key. Specifically, consider the following mechanism parametrized by a reserve price $r > 1$.

- **Globally optimal strategy $\boldsymbol{\sigma}(v)$**: If the user's true value $v \geq r$, $\boldsymbol{\sigma}(v)$ outputs $(r, r/v)$ where $r/v \in (0, 1]$ is a uniquely decodable encoding of the user's true value v. Otherwise, if $v < r$, $\boldsymbol{\sigma}(v)$ simply outputs $\emptyset$, i.e., the user drops its bid.
- **Inclusion rule.**
  - **Case 1**: Check if it is the case that every user $i \in [n]$ submitted a message of the form $(r, r/v_i)$ where $r/v_i \in (0, 1]$ for all $i \in [n]$. If so, decode the true value vector $(v_1, \ldots, v_n)$. Choose the $\min(n, k)$ bids with the highest true values, and include their corresponding r bids in the block.
  - **Case 2**: Else, if the number of bids is even, henceforth denoted $2n$, check whether there are exactly n bids at r and n number of bids that lie in the range $(0, 1]$. If so, interpret the bid vector as $(r, r/v_1), \ldots, (r, r/v_n)$, as if each pair $(r, r/v_i)$ comes from the i-th user for $i \in [n]$. Now, apply the algorithm of Case 1.
  - **Case 3**: Else, if the number of bids is odd, henceforth denoted $2n+1$, check if there are n bids at r and n bids in the range of $(0, 1]$, and one last bid denoted r'. If $r' > r$, then choose r' to include in the block, and among the remaining n bids at r, randomly choose $\min(k-1, n)$ of them to include in the block; else if $r' \leq r$, randomly choose $\min(k, n)$ bids among all the r-bids to include in the block.
  - **Case 4**: In all other cases, randomly choose $\min(k, n)$ bids among all bids that are *strictly larger than* r to include in the block.
- **Confirmation, payment, and miner revenue rules**: All included bids that are at least r are confirmed. Each confirmed bid pays r, and the miner gets nothing.

Claim 7.1. *The above mechanism satisfies UIC, MIC, and a relaxed notion of OCA-proofness in which σ is allowed to output multiple bids.*

Proof. MIC is obvious since the miner always gets 0 revenue. For OCA-proofness, observe that if the global coalition all take the strategy $\boldsymbol{\sigma}$, then case 1 of the inclusion rule will be triggered, and thus the social welfare is maximized.

Proving UIC is the most complicated part. Fix any user i. Notice that each confirmed bid pays r, so if user i's true value is at most r, it cannot get positive utility regardless its bid. Thus, bidding truthfully is a dominant strategy. Henceforth, we assume user i's true value is strictly larger than r. There are two possibilities.

1. First, before user i submits its bid, the bid vector contains exactly n bids at r and n bids lie in the range $(0, 1]$ for some n. Then, bidding truthfully will lead to Case 3, where user i's bid is guaranteed to be included and confirmed, and thus maximizes the utility.

2. Second, before user i submits its bid, the bid vector does not contain exactly n bids at r and n bids lie in the range $(0, 1]$ for some n. Suppose there is no bid strictly larger than r in the bid vector before user i submits its bid. Then, if user i bids truthfully, its bid will be the only bid larger than r, and the inclusion rule goes to Case 4. Since $k \geq 1$, user i's bid is guaranteed to be included and confirmed, and thus maximizes the utility. On the other hand, suppose there are $t \geq 1$ bids strictly larger than r in the bid vector before user i submits its bid. Then, the inclusion can never go to Case 1 or Case 2. If user i bids truthfully, the inclusion rule must go to Case 4, and user i's bid is included with probability $\min(k, t+1)/(t+1)$. However, if the strategic bidding leads to Case 3, it must be $t = 1$. Then, user i's inclusion probability has dropped from $1/2$ to 0 if $k = 1$, or from 1 to at most 1 if $k \geq 2$. On the other hand, if the strategic bidding leads to Case 4, any untruthful bidding $> r$ leads to the same inclusion probability and the same payment, so the utility does not change, while any untruthful bidding $\leq r$ leads to the zero inclusion probability which leads to zero utility. Finally, notice that injecting fake bids never helps in Case 4. Thus, in all cases, the utility is better than bidding truthfully.

□

Recall that weakly symmetric mechanisms do not make use of the bidders' identities or other auxiliary information except for tie-breaking among equal bids. Because the mechanism above relies on the fact that the inclusion rule can distinguish whether a set of bids come from the same user, it may not be obvious that the mechanism satisfies weak symmetry. The next claim proves this formally.

Claim 7.2. *The above mechanism satisfies weak symmetry.*

Proof. To show weak symmetry, we give an equivalent description as follows. Given a bid vector $\mathbf{b}$, the sorting algorithm checks whether $\mathbf{b}$ satisfies the condition of Case 1. If so, the sorting algorithm places the multiple r bids corresponding to the true value vector $(v_1, \ldots, v_n)$ (defined in Case 1 above) from high to low, and sorts the rest in the descending order. Otherwise, if $\mathbf{b}$ does not satisfy the condition of Case 1, the sorting algorithm sorts $\mathbf{b}$ in the descending order and breaks tie arbitrarily.

In the original mechanism, the only case that depends on the extra information (identity, in this case) is Case 1. By applying the sorting algorithm described above, the inclusion rule can implement Case 1 by only depending on the amount of the bids and their relative position in the sorted bid vector.

□

7.3 Inclusion-Rule-Respecting

It is also natural to consider a further relaxation of OCA-proofness henceforth called "inclusion-rule-respecting". In comparison with OCA-proofness, inclusion-rule-respecting only requires that the globally optimal strategy σ not involve altering the inclusion rule, and all other constraints on σ are removed. Inclusion-rule-respecting is strictly weaker than the notions in Section 7.1 and Section 7.2. Therefore, the mechanisms mentioned in Section 7.1 and Section 7.2 simultaneously satisfy UIC, MIC, and inclusion-rule-respecting too.

The "inclusion-rule-respecting" notion has an intuitive interpretation: it discourages miners from altering the intended protocol implementation; in other words, all profiting strategies can be implemented as bidding strategies above the protocol layer. In this sense, "inclusion-rule-respecting" also captures the intuition of "no way to steal from the protocol". One can also view it as follows: global SCP captures "not stealing from the protocol" where the protocol is the union of the miner's inclusion rule as well as users' honest bidding strategies; and inclusion-rule-respecting captures the same notion but where the protocol is only the underlying blockchain protocol. OCA-proofness is somewhere in between.

8 Transaction Fee Mechanisms in the Trusted Hardware Model

In this section, we introduce a new model, the trusted-hardware model, which also helps us bypass the impossibility results. Trusted hardware, like Intel SGX, refers to a secure computation environment whose execution is enforced and whose internal state remains hidden from strategic players. In a transaction fee mechanism, trusted hardware can enforce the inclusion rule and ensure that strategic players cannot view or tamper with honest users' bids during block construction.

We first define the model formally in Section 8.1, then introduce the Bayesian variants of incentive compatibility properties in Section 8.2. In Section 8.3, we show that the second-price auction with reserve satisfies UIC, Bayesian MIC, and global SCP demonstrating that we can bypass the impossibility by adopting the Bayesian notion of incentive compatibility. Finally, in Section 8.4, we show that there remains a fundamental trade-off between global SCP and 1-SCP even if we only aim for Bayesian incentive compatibility.

8.1 The Trusted Hardware Model

Recall that in the plain model (Section 2.1), the miner observes all honest users' bids and may deviate from the inclusion rule. In contrast, the trusted-hardware model assumes that the inclusion rule is enforced within a trusted execution environment (e.g., Intel SGX), which users interact with directly via encrypted bids. In this setting, when designing a mechanism, we only need to specify the *allocation rule* which is a combination of the inclusion and confirmation rules. We say a TFM is *trivial* if the allocation probability of all transactions is always zero; otherwise, it is called *non-trivial*.

Below, we define the game that captures the using of trusted hardware in the transaction fee mechanism design. To formalize the fact that the bids are hidden from the miner, we assume the existence of an ideal trusted functionality T who the honest users interact with to submit their bids and execute the mechanism rules. When the strategic player is an individual user, the execution of a transaction fee mechanism in the trusted-hardware model is described by the following game:

1. Each honest user submits a bid to T. The strategic user can submit a non-negative number of bids to T.

2. The trusted party T collects all bids and outputs the outcomes of all players according to the allocation rule, the payment rule, and the miner revenue rule.

When the strategic player, denoted as $\mathcal{C}$, is the miner or a coalition of the miner and some users, the execution of a transaction fee mechanism in the trusted-hardware model is described by the following game:

1. Each honest user i submits its identity i to the miner, and its bid b_i to T.
2. Let $\mathcal{H}$ denote the set of honest identities received by the miner. The coalition $\mathcal{C}$ chooses a subset $\mathcal{H}' \subseteq \mathcal{H}$, and a vector of bids $\mathbf{b}_{\mathcal{C}}$ of arbitrary length. Then, $\mathcal{C}$ submits $\mathcal{H}'$ and $\mathbf{b}_{\mathcal{C}}$ to T.
3. The trusted party T outputs the outcomes of all players according to the allocation rule, payment rule, and the miner revenue rule by using $\{b_i\}_{i \in \mathcal{H}'} \cup \mathbf{b}_{\mathcal{C}}$ as the input.

In the games above, an honest user always takes on a single identity, and submits a bid according to the bidding rule. However, a strategic user can take a non-negative number of identities, and submit bids on behalf of these identities which possibly deviate from the bidding rule. An honest miner does not take on any identity or submit any bid, and it always forwards all users' identities to T. A strategic miner can take a non-negative number of identities and submit fake bids. Moreover, we allow the strategic miner to choose a subset $\mathcal{H}' \subseteq \mathcal{H}$ of users' identities. This captures the fact that the trusted hardware relies on the miner to relay the bids, so the miner can censor some honest users' bids by not forwarding them to the trusted hardware. A miner-user coalition can adopt a combination of the above strategies.

Remark 8.1. Shi, Chung, and Wu [SCW23] introduced the *MPC-assisted model*, which further restricts the miner's strategy space. In this model, the mechanism is executed by a multiparty computation (MPC) protocol run by multiple miners. The MPC-assisted model is appropriate when the strategic players do not control all the parties participating in the MPC. However, global SCP and OCA-proofness focus on the global coalition of all users and the miners. If such a global coalition corrupts all MPC participants, the protocol's security is compromised. Therefore, the MPC-assisted model is incompatible with global SCP and OCA-proofness.

In terms of strategy space, the trusted hardware model lies between the plain model and the MPC-assisted model. Both the trusted hardware and MPC-assisted models prevent the strategic miner from learning the honest users' bids. However, unlike in the MPC-assisted model, the trusted hardware model still allows the strategic miner to censor some bids.

8.2 Defining Bayesian Incentive Compatibility

The properties, UIC, MIC, and SCP, defined in Section 2.2 require that the honest behavior maximizes the strategic players' utilities *for any* bid vector submitted by the honest players. In contrast, Bayesian incentive compatibility assumes the honest players' values are drawn from a known distribution, and it only requires that the honest behavior maximizes the strategic players' expected utilities *over the distribution of the honest players' values* assuming the honest players follow the bidding rule. This notion is particularly justified when the strategic players have to submit their bids without having seen other honest users' bids.

Let $\mathcal{D}$ denote a distribution representing the true values of honest users, which is known to the strategic players. Let β denote the bidding rule, and we define $\beta(\mathcal{D})$ as the distribution of bids when a user follows the bidding rule β and its true value is drawn from $\mathcal{D}$. The Bayesian incentive compatibility properties are defined as follows.

Definition 8.2 (Bayesian user incentive compatible)**.** We say a TFM satisfies *Bayesian user incentive compatible (Bayesian UIC)* iff the following holds: for any number n of users, for any user i, for any true value v_i of user i, and for any bid vector $\mathbf{b}_i$ submitted by user i, it holds that

$$\mathop{\mathbb{E}}_{\mathbf{b}_{-\mathcal{C}} \sim \beta(\mathcal{D})^{n-1}} [\mathsf{util}_i(\mathbf{b}_{-\mathcal{C}}, v_i)] \geq \mathop{\mathbb{E}}_{\mathbf{b}_{-\mathcal{C}} \sim \beta(\mathcal{D})^{n-1}} [\mathsf{util}_i(\mathbf{b}_{-\mathcal{C}}, \mathbf{b}_i)],$$

where $\mathsf{util}_i(\mathbf{b})$ denotes user i's expected utility (taken over the random coins of the TFM) assuming the miner implements the TFM honestly.

Definition 8.3 (Bayesian miner incentive compatible)**.** We say a TFM satisfies *Bayesian miner incentive compatible (Bayesian MIC)* iff the following holds: for any set $\mathcal{H}$ of users with arbitrary size n, for any subset $\mathcal{H}' \subseteq \mathcal{H}$, and for any bid vector $\mathbf{b}_{\mathcal{M}}$ submitted by the miner, it holds that

$$\mathop{\mathbb{E}}_{\mathbf{b}_{\mathcal{H}} \sim \beta(\mathcal{D})^{n}} [\mathsf{util}_{\mathcal{M}}(\mathbf{b}_{\mathcal{H}})] \geq \mathop{\mathbb{E}}_{\mathbf{b}_{\mathcal{H}} \sim \beta(\mathcal{D})^{n}} [\mathsf{util}_{\mathcal{M}}(\mathbf{b}_{\mathcal{H}'}, \mathbf{b}_{\mathcal{M}})],$$

where $\mathbf{b}_{\mathcal{H}'}$ denotes the subvector of $\mathbf{b}_{\mathcal{H}}$ corresponding to the users in $\mathcal{H}'$, and $\mathsf{util}_{\mathcal{M}}(\mathbf{b})$ denotes the miner's expected utility (taken over the random coins of the TFM) given the input to the allocation rule is $\mathbf{b}$.

Definition 8.4 (Bayesian side-contract-proof)**.** We say a TFM satisfies *Bayesian c-side-contract-proof (Bayesian c-SCP)* iff the following holds: for any miner-user coalition $\mathcal{C}$ consisting of the miner and at most c users, for any true values $\mathbf{v}_{\mathcal{C}}$ of the users in $\mathcal{C}$, for any bid vector $\mathbf{b}_{\mathcal{C}}$ submitted by $\mathcal{C}$, for any set $\mathcal{H}$ of honest users with arbitrary size n, and for any subset $\mathcal{H}' \subseteq \mathcal{H}$, it holds that

$$\mathop{\mathbb{E}}_{\mathbf{b}_{\mathcal{H}} \sim \beta(\mathcal{D})^{n}} [\mathsf{util}_{\mathcal{C}}(\mathbf{b}_{\mathcal{H}}, \mathbf{v}_{\mathcal{C}})] \geq \mathop{\mathbb{E}}_{\mathbf{b}_{\mathcal{H}} \sim \beta(\mathcal{D})^{n}} [\mathsf{util}_{\mathcal{C}}(\mathbf{b}_{\mathcal{H}'}, \mathbf{b}_{\mathcal{C}})],$$

where $\mathbf{b}_{\mathcal{H}'}$ denotes the subvector of $\mathbf{b}_{\mathcal{H}}$ corresponding to the users in $\mathcal{H}'$, and $\mathsf{util}_{\mathcal{C}}(\mathbf{b})$ denotes the coalition's expected utility (taken over the random coins of the TFM) given the input to the allocation rule is $\mathbf{b}$.

Because the strategic players considered in global SCP and OCA-proofness are the coalition of all users and the miner, and there are no honest users outside the coalition, Bayesian notions of global SCP and OCA-proofness are equivalent to their *ex post* counterparts, i.e., Definition 2.5 and Definition 2.6.

8.3 Second-Price Auction Achieves UIC + Bayesian MIC + Global SCP

In this section, we assume each user's true value is independently drawn from a distribution $\mathcal{D}$, and we assume the probability density function of $\mathcal{D}$ exists. Let F and f denote the cumulative distribution function and probability density function of $\mathcal{D}$, respectively. We define the *virtual value* function as $\phi(v) = v - \frac{1-F(v)}{f(v)}$. We assume $\mathcal{D}$ is *regular*, i.e., $\phi(v)$ is strictly increasing and f is strictly positive over the support. We set the reserve price to be $r = \phi^{-1}(0)$. Note that for any bounded distribution over $[0, v_{\max}]$, we have $\phi(v_{\max}) = v_{\max} \geq 0$, so $\phi^{-1}(0)$ exists.

Second-price auction with reserve

Parameters: the reserve r and the block size k.
Inputs: a bid vector $(b_1, \ldots, b_N)$ where $b_1 \geq \cdots \geq b_N$ (break tie arbitrarily).

- **Bidding rule**: Each user bids its true value.
- **Allocation rule and payment rule**: All bids among $b_1, \ldots, b_k$ that are at least r are

confirmed. Each confirmed bid pays $\max(r, b_{k+1})$.

- **Miner revenue rules**: For each confirmed bid, the reserve is burnt, and the miner gets the rest.

Theorem 8.5. *The second-price auction with reserve above satisfies the following properties:*

- *It satisfies UIC, and global SCP for users' values regardless of distribution $\mathcal{D}$.*
- *If each buyer's true value is i.i.d. sampled from a bounded regular distribution $\mathcal{D}$, it additionally satisfies Bayesian MIC.*

Proof. We start with the proof of UIC and global SCP. The proof of Bayesian MIC is more involved, so we will postpone it to the end.

UIC. Since the allocation and payment rules align with the conditions of Myerson's lemma, a user cannot increase their utility by bidding untruthfully. Additionally, injecting fake bids cannot reduce the payment required. Thus, the mechanism satisfies UIC. This property holds even if users observe others' bids before submitting their own.

Global SCP. Each confirmed bid burns an amount r, so confirming a transaction with a value $< r$ reduces total social welfare. On the other hand, confirming a transaction with value $v \geq r$ increases social welfare by $v - r$, regardless of the exact payment. Because the mechanism always confirms the highest bids that are at least r, everyone submits the true value maximizes the social welfare. Thus, it satisfies global SCP.

Bayesian MIC. In the second-price auction, censoring honest users' bids never increases the payment from confirmed users nor the number of confirmed users. Hence, we assume the miner always forwards all honest users' bids to the trusted hardware, while it may inject fake bids.

Let Π denote the second-price auction with reserve r described above. Define Π' as an identical mechanism, except that no payment is burned: the miner receives the full payment. We will show the following:

1. Π' satisfies Bayesian MIC.[10]
2. If the miner can profit from deviating in Π, it can also profit from deviating in Π'.

This implies that Π must also satisfy Bayesian MIC.

We first show that Π' satisfies Bayesian MIC. Let d be the payment from a confirmed user in an honest execution, so the miner receives d per confirmed bid. Now, suppose the miner deviates from the inclusion rule or injects fake bids so that the payment becomes d'. Since only confirmed bids generate revenue, the deviation is profitable only if

$$d' \cdot \Pr[v \geq d' \mid v \geq d] > d, \tag{1}$$

or equivalently, $d' \cdot (1 - F(d')) > d \cdot (1 - F(d))$.

Since Equation (1) only holds if $d' > d$, we assume $d' > d$ henceforth. We define $h(v) = v \cdot (1 - F(v))$, and we have $h'(v) = -\phi(v) \cdot f(v)$. Because $\phi(v)$ is strictly increasing, $h(v)$ is non-increasing for $v \geq \phi^{-1}(0)$. Since $d' > d \geq r = \phi^{-1}(0)$, Equation (1) cannot hold. Thus, fake bid injection does not increase the miner's utility in Π'.

[10] Although Π' satisfies Bayesian MIC, it does not satisfy global SCP; see Theorem 8.6.

Now we are ready to show that Π also satisfies Bayesian MIC. Given any bid vector $\mathbf{b}$ from honest users, let P denote the total payment collected from the confirmed users, and let t be the number of confirmed bids in an honest execution of Π. In this case, the miner's utility is $P - t \cdot r$. On the other hand, let P' denote the total payment collected from the honest confirmed users (not including the miner's payment), and let t' be the number of confirmed bids (including fake bids) if the miner injects some fake bids $\mathbf{f}$ in Π. In this case, the miner's utility is $P' - t' \cdot r$. Note that t, t', P, P' are the functions of $\mathbf{b}$ and $\mathbf{f}$.

For the sake of contradiction, suppose Π does not satisfy Bayesian MIC. That is, there exists a number n of users such that the miner's expected utility increases after injecting fake bids $\mathbf{f}$ when n users' true values are drawn from $\mathcal{D}$. Thus, we have $\mathbb{E}_{\mathbf{b}\sim\mathcal{D}^n}[P' - t' \cdot r] > \mathbb{E}_{\mathbf{b}\sim\mathcal{D}^n}[P - t \cdot r]$. After inject fake bids, the number of confirmed bids (including fake bids) never decreases, so we have $t' \geq t$. Thus, we have $\mathbb{E}_{\mathbf{b}\sim\mathcal{D}^n}[P - t \cdot r] \geq \mathbb{E}_{\mathbf{b}\sim\mathcal{D}^n}[P - t' \cdot r]$. Combining the two inequalities, we have

$$\mathop{\mathbb{E}}_{\mathbf{b}\sim\mathcal{D}^n}\left[P'\right] > \mathop{\mathbb{E}}_{\mathbf{b}\sim\mathcal{D}^n}\left[P\right].$$

Because the allocation rule and the payment rule of Π and Π' are identical, for any $\mathbf{b}$ and $\mathbf{f}$, the number of confirmed users and the corresponding payments are identical in Π and Π'. Consequently, the miner can profit by injecting fake bids $\mathbf{f}$ in Π', and it contradicts the fact that Π' satisfies Bayesian MIC. □

Remark 8.6 (On the Necessity of Burning the Reserve)**.** Burning the reserve is essential for ensuring global SCP. If the reserve is not burned and the miner receives the full payment, then in situations where fewer than k users have values $\geq r$, low-value users are incentivized to overbid to occupy block space. Without burning, the coalition's utility increases with these overbids. Conversely, if all payments are burned, the global coalition is incentivized to reduce payments, e.g., by having all but the top k bidders bid 0 to keep the price at r.

Remark 8.7 (First-price auction satisfies Bayesian UIC, MIC, and global SCP in the permissioned setting)**.** We have seen that the second-price auction can bypass the impossibility result by appealing to the Bayesian notion of MIC. Interestingly, the first-price auction can also bypass the impossibility—this time via the Bayesian notion of UIC—provided we are in a *permissioned setting*, where the number of users n is fixed *a priori* and is common knowledge to all players and the mechanism.

Under the assumption that other users' bids are i.i.d. draws from a known distribution $\mathcal{D}$, a bidder with true value v in a first-price auction has a Bayesian optimal strategy to bid $\gamma(n) \cdot v$, where $\gamma(n) \in [0, 1]$ depends on n.[11] When n is known *a priori*, this equilibrium bidding rule ensures that the first-price auction satisfies Bayesian UIC. The first-price auction also satisfies MIC, as shown in [Rou20], and it satisfies global SCP since it always selects the bidders with the highest values.

However, Bayesian UIC here crucially relies on the permissioned setting. In designing a TFM for blockchains, we require a mechanism that works for any number n of users. In such a setting, no single dominant bidding strategy exists that works for all possible n.

Remark 8.8 (Smaller Reserve in the Permissioned Setting)**.** If the number n of users is known in advance (e.g., in a permissioned setting), the reserve can be set below $\phi^{-1}(0)$. This reduces the burned payment and increases both miner revenue and social welfare.

Formally, given n, reserve r, and a fake bid vector $\mathbf{f}$ injected by the miner, let:

[11] For an educational derivation, see, e.g., https://cs.brown.edu/courses/cs1951k/lectures/2020/first_price_auctions.pdf

- $P(r, n, \mathbf{f})$ be the expected total payment from honest users (excluding fake bids),
- $B(r, n, \mathbf{f})$ be the expected total amount burned (including fake bids).

Bayesian MIC holds if

$$P(r, n, \emptyset) - B(r, n, \emptyset) \geq P(r, n, \mathbf{f}) - B(r, n, \mathbf{f}), \tag{2}$$

where the LHS is the miner's expected utility under honest execution, and the RHS under fake bids $\mathbf{f}$.

For example, if values are i.i.d. drawn from the uniform distribution over $[0, 1]$, then $\phi(v) = 2v-1$ and $\phi^{-1}(0) = 1/2$. The smallest r satisfying Equation (2) for various n is as follows:

n	1	2	3	4	5	10	100	1000
r	0.250	0.305	0.335	0.355	0.370	0.410	0.480	0.497

As n grows, $r \to \phi^{-1}(0) = 1/2$. Intuitively, for large n, the block is almost surely full, so we have $B(r, n, \mathbf{f}) \approx B(r, n, \emptyset)$, and Equation (2) reduces to

$$P(r, n, \emptyset) \geq P(r, n, \mathbf{f}),$$

matching the condition in classical optimal mechanism design.

8.4 Impossibility of Bayesian UIC + Bayesian MIC + Bayesian 1-SCP + Global SCP

In this section, we will show that for any mechanism satisfying Bayesian UIC, Bayesian MIC, Bayesian 1-SCP and global SCP, the social welfare must be zero. Notice that if users' true values are sampled from a bounded distribution over $[0, M]$, we can run the posted-price auction with a price M: only bids at least M are confirmed, each pays M, and all payments are burned. Although some transactions are confirmed, the mechanism is still undesired because everyone's utility is zero. Our impossibility result in this section demonstrates that it is a consequence of simultaneously achieving all four incentive compatibility properties.

We begin by recalling the following lemma from [SCW23].

Lemma 8.9 (Theorem C.5 in [SCW23])**.** *Suppose each user's true value is i.i.d. sampled from a distribution $\mathcal{D}$. For any (possibly randomized) TFM satisfying Bayesian UIC, Bayesian MIC, and Bayesian 1-SCP, and for any number n of users, it must be*

$$\mathop{\mathbb{E}}_{\mathbf{b} \sim \mathcal{D}^n}[\mu(\mathbf{b})] = 0.$$

Lemma 8.10. *Suppose each user's true value is i.i.d. sampled from a distribution $\mathcal{D}$. Consider any (possibly randomized) TFM satisfying Bayesian UIC and global SCP. Then, for any $n \geq 1$, for any user i, and for any b_i, b_i', it holds that*

$$\mathop{\mathbb{E}}_{\mathbf{b}_{-i} \sim \mathcal{D}^n}[\mathsf{SW}_{-i}(\mathbf{b}_{-i}, b_i)] = \mathop{\mathbb{E}}_{\mathbf{b}_{-i} \sim \mathcal{D}^n}[\mathsf{SW}_{-i}(\mathbf{b}_{-i}, b_i')].$$

Proof. The proof is similar to the proof of Lemma B.6, while the mechanism only satisfies Bayesian UIC instead of ex post UIC. We define

$$\overline{x}_i(\cdot) = \mathop{\mathbb{E}}_{\mathbf{b}_{-i} \sim \mathcal{D}^n}[x_i(\mathbf{b}_{-i}, \cdot)], \quad \overline{p}_i(\cdot) = \mathop{\mathbb{E}}_{\mathbf{b}_{-i} \sim \mathcal{D}^n}[p_i(\mathbf{b}_{-i}, \cdot)].$$

By Bayesian UIC, for any user i, when user i's true value is b_i, it must be

$$\begin{aligned}\mathop{\mathbb{E}}_{\mathbf{b}_{-i}\sim\mathcal{D}^n}\left[\mathsf{util}_i(\mathbf{b}_{-i}, b_i) - \mathsf{util}_i(\mathbf{b}_{-i}, b_i')\right] &= [b_i \cdot \overline{x}_i(b_i) - \overline{p}_i(b_i)] - [b_i \cdot \overline{x}_i(b_i') - \overline{p}_i(b_i')] \\ &\le [b_i' - b_i][\overline{x}_i(b_i') - \overline{x}_i(b_i)].\end{aligned}$$

By global SCP, for any bid vector $\mathbf{b}_{-i}$, when user i's true value is b_i, it must be

$$\mathsf{SW}_{-i}(\mathbf{b}_{-i}, b_i') - \mathsf{SW}_{-i}(\mathbf{b}_{-i}, b_i) \le \mathsf{util}_i(\mathbf{b}_{-i}, b_i) - \mathsf{util}_i(\mathbf{b}_{-i}, b_i').$$

Thus, we have

$$\mathop{\mathbb{E}}_{\mathbf{b}_{-i}\sim\mathcal{D}^n}\left[\mathsf{SW}_{-i}(\mathbf{b}_{-i}, b_i') - \mathsf{SW}_{-i}(\mathbf{b}_{-i}, b_i)\right] \le [b_i' - b_i][\overline{x}_i(b_i') - \overline{x}_i(b_i)].$$

Now, for any b_i and b_i', we can divide the interval $[b_i, b_i']$ into L equally sized segments $b_i^{(0)}, \ldots, b_i^{(L)}$ where $b_i^{(0)} = b_i$ and $b_i^{(L)} = b_i'$. Then, we have

$$\begin{aligned}\mathop{\mathbb{E}}_{\mathbf{b}_{-i}\sim\mathcal{D}^n}\left[\mathsf{SW}_{-i}(\mathbf{b}_{-i}, b_i') - \mathsf{SW}_{-i}(\mathbf{b}_{-i}, b_i)\right] &= \sum_{j=0}^{L-1} \mathop{\mathbb{E}}_{\mathbf{b}_{-i}\sim\mathcal{D}^n}\left[\mathsf{SW}_{-i}(\mathbf{b}_{-i}, b_i^{(j+1)}) - \mathsf{SW}_{-i}(\mathbf{b}_{-i}, b_i^{(j)})\right] \\ &\le \sum_{j=0}^{L-1} (b_j^{(j+1)} - b_j^{(j)}) \cdot \left[\overline{x}_i(b_i^{(j+1)}) - \overline{x}_i(b_i^{(j)})\right] \\ &= \frac{b_i' - b_i}{L} \cdot \left[\overline{x}_i(b_i') - \overline{x}_i(b_i)\right].\end{aligned}$$

By taking the limit for $L \to \infty$, we have $\mathop{\mathbb{E}}_{\mathbf{b}_{-i}\sim\mathcal{D}^n}[\mathsf{SW}_{-i}(\mathbf{b}_{-i}, b_i)] \le \mathop{\mathbb{E}}_{\mathbf{b}_{-i}\sim\mathcal{D}^n}[\mathsf{SW}_{-i}(\mathbf{b}_{-i}, b_i')]$. Because the argument above holds for any b_i and b_i', it must be $\mathop{\mathbb{E}}_{\mathbf{b}_{-i}\sim\mathcal{D}^n}[\mathsf{SW}_{-i}(\mathbf{b}_{-i}, b_i)] = \mathop{\mathbb{E}}_{\mathbf{b}_{-i}\sim\mathcal{D}^n}[\mathsf{SW}_{-i}(\mathbf{b}_{-i}, b_i')]$. □

Lemma 8.11. *Suppose each user's true value is i.i.d. sampled from a distribution $\mathcal{D}$. Consider any (possibly randomized) TFM satisfying Bayesian UIC and global SCP. Then, for any $n \ge 1$, for any user i, and for any b_i, it holds that*

$$\mathop{\mathbb{E}}_{\mathbf{b}_{-i}\sim\mathcal{D}^n}[\mathsf{SW}_{-i}(\mathbf{b}_{-i})] = \mathop{\mathbb{E}}_{\mathbf{b}_{-i}\sim\mathcal{D}^n}[\mathsf{SW}_{-i}(\mathbf{b}_{-i}, b_i)].$$

Proof. First, we show that $\mathop{\mathbb{E}}_{\mathbf{b}_{-i}\sim\mathcal{D}^n}[\mathsf{SW}_{-i}(\mathbf{b}_{-i})] = \mathop{\mathbb{E}}_{\mathbf{b}_{-i}\sim\mathcal{D}^n}[\mathsf{SW}_{-i}(\mathbf{b}_{-i}, 0_i)]$. To see this, if $\mathop{\mathbb{E}}_{\mathbf{b}_{-i}\sim\mathcal{D}^n}[\mathsf{SW}_{-i}(\mathbf{b}_{-i})] < \mathop{\mathbb{E}}_{\mathbf{b}_{-i}\sim\mathcal{D}^n}[\mathsf{SW}_{-i}(\mathbf{b}_{-i}, 0_i)]$, then consider a world consists of n users. The global coalition can inject a fake bid 0_i, and the social welfare increases. On the other hand, if $\mathop{\mathbb{E}}_{\mathbf{b}_{-i}\sim\mathcal{D}^n}[\mathsf{SW}_{-i}(\mathbf{b}_{-i})] > \mathop{\mathbb{E}}_{\mathbf{b}_{-i}\sim\mathcal{D}^n}[\mathsf{SW}_{-i}(\mathbf{b}_{-i}, 0_i)]$, then consider a world consists of $n+1$ users where user i's true value is zero. Because user i's utility is zero under an honest execution, we have $\mathop{\mathbb{E}}_{\mathbf{b}_{-i}\sim\mathcal{D}^n}[\mathsf{SW}_{-i}(\mathbf{b}_{-i}, 0_i)] = \mathop{\mathbb{E}}_{\mathbf{b}_{-i}\sim\mathcal{D}^n}[\mathsf{SW}(\mathbf{b}_{-i}, 0_i)]$. If user i drops out, the social welfare becomes $\mathop{\mathbb{E}}_{\mathbf{b}_{-i}\sim\mathcal{D}^n}[\mathsf{SW}_{-i}(\mathbf{b}_{-i})]$, which is larger than the honest execution. Thus, it must be $\mathop{\mathbb{E}}_{\mathbf{b}_{-i}\sim\mathcal{D}^n}[\mathsf{SW}_{-i}(\mathbf{b}_{-i})] = \mathop{\mathbb{E}}_{\mathbf{b}_{-i}\sim\mathcal{D}^n}[\mathsf{SW}_{-i}(\mathbf{b}_{-i}, 0_i)]$. The lemma directly follows from the argument above and Theorem 8.10. □

Lemma 8.12. *Suppose each user's true value is i.i.d. sampled from a distribution $\mathcal{D}$, and let $u = \mathop{\mathbb{E}}_{b\sim\mathcal{D}^n}[\mathsf{util}(b)]$ denote the expected utility of the lone user if only one user submits a bid drawn from $\mathcal{D}$. Then, for any possibly randomized TFM that simultaneously satisfy Bayesian UIC, Bayesian MIC, Bayesian 1-SCP, and global SCP, and for any n, it must be*

$$\mathop{\mathbb{E}}_{\mathbf{b}\sim\mathcal{D}^n}[\mathsf{SW}(\mathbf{b})] = n \cdot u.$$

Proof. We show the lemma by induction on the number of users n. By Theorem 8.9, the miner revenue must be zero. Thus, $\mathop{\mathbb{E}}_{b\sim\mathcal{D}}[\mathsf{SW}(b)] = \mathop{\mathbb{E}}_{b\sim\mathcal{D}}[\mathsf{util}(b)] = u$, and the base case $n = 1$ holds.

Next, we show the induction step. Suppose the lemma holds for k users, i.e., $\mathop{\mathbb{E}}_{\mathbf{b}\sim\mathcal{D}^k}[\mathsf{SW}(\mathbf{b})] = k\cdot u$ for an integer k. We first show that each user's expected utility is the same regardless of the identity. By weak symmetry, for any n, the social welfare $\mathop{\mathbb{E}}_{\mathbf{b}\sim\mathcal{D}^n}[\mathsf{SW}(\mathbf{b})]$ is independent of the user identities. By Theorem 8.11, for any user identity i, we have $\mathop{\mathbb{E}}_{\mathbf{b}\sim\mathcal{D}^n}[\mathsf{SW}(\mathbf{b})] = \mathop{\mathbb{E}}_{\mathbf{b}\sim\mathcal{D}^{n+1}}[\mathsf{SW}_{-i}(\mathbf{b})]$. For any bid vector $\mathbf{b}$, we have $\mathsf{SW}(\mathbf{b}) = \mathsf{SW}_{-i}(\mathbf{b}) + \mathsf{util}_i(\mathbf{b})$. Thus, we have $\mathop{\mathbb{E}}_{\mathbf{b}\sim\mathcal{D}^{n+1}}[\mathsf{util}_i(\mathbf{b})] = \mathop{\mathbb{E}}_{\mathbf{b}\sim\mathcal{D}^{n+1}}[\mathsf{SW}(\mathbf{b})] - \mathop{\mathbb{E}}_{\mathbf{b}\sim\mathcal{D}^{n}}[\mathsf{SW}(\mathbf{b})]$, which implies that each user's expected utility is the same regardless of the identity for any n.

By assumption, we have $\mathop{\mathbb{E}}_{\mathbf{b}\sim\mathcal{D}^k}[\mathsf{SW}(\mathbf{b})] = k \cdot u$. Because the miner revenue is always zero and $\mathop{\mathbb{E}}_{\mathbf{b}\sim\mathcal{D}^k}[\mathsf{SW}(\mathbf{b})] = \mathop{\mathbb{E}}_{\mathbf{b}\sim\mathcal{D}^{k+1}}[\mathsf{SW}_{-i}(\mathbf{b})]$, we have $\mathop{\mathbb{E}}_{\mathbf{b}\sim\mathcal{D}^{k+1}}[\mathsf{util}_i(\mathbf{b})] = u$ for any user i. Finally, $\mathop{\mathbb{E}}_{\mathbf{b}\sim\mathcal{D}^{k+1}}[\mathsf{util}_i(\mathbf{b})] = \mathop{\mathbb{E}}_{\mathbf{b}\sim\mathcal{D}^{k+1}}[\mathsf{util}_i(\mathbf{b})] + \mathop{\mathbb{E}}_{\mathbf{b}\sim\mathcal{D}^{k}}[\mathsf{SW}(\mathbf{b})] = k \cdot u + u = (k+1)\cdot u$. Thus, the lemma holds for $k+1$ users, and we conclude the proof. □

Lemma 8.13. *Let $\mathcal{D}$ be any non-negative distribution with bounded mean. Given any natural numbers n, let $X_1, \ldots, X_n$ be i.i.d. random variables drawn from $\mathcal{D}$, and let $Y_n = \max(X_1, \ldots, X_n)$. Then, it holds that*

$$\lim_{n\to\infty} \mathbb{E}[Y_n]/n = 0.$$

Proof. Given any real number $t > 0$, for all $i \in [n]$, let A_i^t and B_i^t be the random variables defined as follows:

$$A_i^t = \begin{cases} X_i, & \text{if } X_i \le t, \\ 0, & \text{otherwise;} \end{cases} \text{ and } B_i^t = \begin{cases} X_i, & \text{if } X_i > t, \\ 0, & \text{otherwise.} \end{cases}$$

By definition, we have $X_i = A_i^t + B_i^t$. Thus, we have

$$Y_n = \max(X_1, \ldots, X_n) \le \max(A_1^t, \ldots, A_n^t) + \max(B_1^t, \ldots, B_n^t) \le t + \max(B_1^t, \ldots, B_n^t).$$

Next, fix any $\epsilon > 0$. Because the mean of $\mathcal{D}$ is bounded, there exists a threshold t such that $\mathbb{E}[B_1^t] \le \epsilon$. Because $\max(B_1^t, \ldots, B_n^t) \le \sum_{i=1}^n B_i^t$, we have

$$\mathbb{E}[\max(B_1^t, \ldots, B_n^t)] \le \sum_{i=1}^{n} \mathbb{E}[B_i^t] = n \cdot \mathbb{E}[B_1^t] \le n \cdot \epsilon.$$

Combining the arguments above, we have

$$\mathbb{E}[Y_n] \le t + \mathbb{E}[\max(B_1^t, \ldots, B_n^t)] \le t + n \cdot \epsilon.$$

Thus, we have

$$\lim_{n\to\infty} \frac{\mathbb{E}[Y_n]}{n} \le \epsilon.$$

Because the argument holds for any $\epsilon > 0$, we conclude that $\lim_{n\to\infty} \mathbb{E}[Y_n]/n = 0$. □

Theorem 8.14. *Let $\mathcal{D}$ be any non-negative distribution with bounded mean. Suppose each user's true value is i.i.d. sampled from $\mathcal{D}$, and suppose the block size is finite. Then, for any possibly randomized TFM that simultaneously satisfy Bayesian UIC, Bayesian MIC, Bayesian 1-SCP, and global SCP, the expected social welfare must be zero.*

Proof. Let k be the block size, and let m be the mean of $\mathcal{D}$. For the sake of contradiction, suppose there exists a number n of users such that the expected social welfare $U \coloneqq \mathop{\mathbb{E}}_{\mathbf{b}\sim\mathcal{D}^n}[\mathsf{SW}(\mathbf{b})]$ is positive. By Theorem 8.12, we have $u \coloneqq \mathop{\mathbb{E}}_{b\sim\mathcal{D}^n}[\mathsf{util}(b)] = U/n > 0$.

Theorem 8.13 implies that for any $\epsilon > 0$, and for sufficiently large N, when there are N users whose values are i.i.d. sampled from $\mathcal{D}$, the expectation of the largest user's value is strictly smaller than $N\epsilon$. Because only at most k users can be confirmed, the expected social welfare is strictly smaller than $Nk\epsilon$. Let $\epsilon < u/k$. By Theorem 8.12, when there are N users, the expected social welfare is $Nu > Nk\epsilon$, which leads to a contradiction. □

9 Static Revelation Principle for Transaction Fee Mechanisms

Informally speaking, the *revelation principle* says that any mechanism can be simulated by an equivalent truthful mechanism. Whenever the revelation principle holds, without loss of generality, we may assume that the users' honest bidding strategy is simply truth-telling. We focus here on a revelation principle for "static" mechanisms in which users interact with a mechanism simply by submitting a bid.[12]

TFMs differ from traditional auctions in that they additionally need to satisfy MIC and collusion-resilience, and also the separation of the inclusion rule (executed by the miner) and the confirmation/payment rules (executed by the blockchain). Therefore, revelation principles are not "automatic." In this section, we prove that for any fixed c, the revelation principle holds for TFMs that must satisfy UIC, MIC, and c-SCP. As a direct corollary, the revelation principle holds for TFMs that satisfy UIC, MIC, and global SCP. The main subtlety in the proof is that when we bake the user's non-truth-telling bidding rule β into the mechanism itself, not only do we need to have the miner's inclusion rule execute β, the blockchain's confirmation/payment rules must also double-check the correct enforcement of β again. Otherwise, a strategic miner may not honestly enforce β.

Notice that we focus on the mechanisms that satisfy c-SCP instead of OCA-proofness in this section. c-SCP requires that following the honest bidding rule and the honest inclusion rule is a dominant strategy for a coalition consisting of the miner and at most c users. For a non-truthful mechanism with the bidding rule β, β is used by users to satisfy both the UIC and the c-SCP conditions. In Section 9.1, we assume the honest bidding rule β only outputs a single bid. Later in Section 9.2, we generalize the proof so that β can output any non-negative number of bids.

9.1 Static Revelation Principle: Bidding Rules That Output Single Bid

The outcome of a TFM is defined as a tuple $(\mathbf{x}, \mathbf{p}, \mu)$ where $\mathbf{x} \in \{0,1\}^t$ is a bit-vector indicating whether each bid is confirmed or not, $\mathbf{p} \in \mathbb{R}^t_{\geq 0}$ is the vector of payments for all bids, and $\mu \in \mathbb{R}_{\geq 0}$ is the miner's revenue.

[12] More generally, typical revelation principles in auction theory transform non-direct mechanisms (in which the action space may be different from the valuation space, for example due to multiple rounds of user-mechanism interaction) into a direct mechanisms.

Theorem 9.1 (Static revelation principle)**.** *Let c be any natural number. Suppose Π is a non-truthful TFM for block size k that is UIC, MIC, and c-SCP with an individually rational bidding rule β. If β always outputs a single bid, then there exists a truthful TFM Π' for block size k that is UIC, MIC, and c-SCP such that 1) the honest bidding rule is truth-telling, and 2) given any vector of true values, the outcome under an honest execution of Π is identically distributed as the outcome under an honest execution of Π'.*

Proof. Suppose we are given a TFM $\Pi = (\beta, \mathbf{I}, \mathbf{C}, \mathbf{P}, \mu)$ where $\beta : \mathbb{R}_{\geq 0} \to \mathbb{R}$ denotes the user's honest bidding rule, $\mathbf{I}$ denotes the inclusion rule, $\mathbf{C}$ denotes the confirmation rule, $\mathbf{P}$ denotes the payment rule, and μ denotes the miner revenue rule. Given a vector $\mathbf{b} = (b_1, \ldots, b_t)$, let $\beta(\mathbf{b})$ denote the element-wise application of β; that is, $\beta(\mathbf{b}) = (\beta(b_1), \ldots, \beta(b_t))$. We construct a truthful mechanism $\Pi' = (\mathbf{I}', \mathbf{C}', \mathbf{P}', \mu')$ as follows:

- Inclusion rule $\mathbf{I}'$: Given a bid vector $\mathbf{b}$, the miner selects the bids by using $\mathbf{I}$ as if the bid vector is $\beta(\mathbf{b})$. That is, the miner selects the bids from $\mathbf{b}$ if the corresponding bids in $\beta(\mathbf{b})$ are selected by $\mathbf{I}$.
- Confirmation rule $\mathbf{C}'$, payment rule $\mathbf{P}'$ and miner revenue rule μ': Given a created block $\mathbf{B}$, the mechanism applies $\mathbf{C}$ and $\mathbf{P}$ to $\beta(\mathbf{B})$. Then, a bid b in the original block $\mathbf{B}$ is confirmed if and only if its corresponding bid $\beta(b)$ is confirmed when applying $\mathbf{C}$ to $\beta(\mathbf{B})$, and its payment is the same as its corresponding bid when applying $\mathbf{P}$ to $\beta(\mathbf{B})$. The miner revenue is $\mu(\beta(\mathbf{B}))$; that is, applying μ as if the block is $\beta(\mathbf{B})$.

Notice that β is individually rational under Π, so the payment corresponding to the bid $\beta(v)$ never exceeds v. Thus, under the induced mechanism Π', the payment never exceeds the bid, so the individual rationality of the payment rule is respected.

Next, by directly checking the syntax, the user's confirmation probability and expected payment when submitting a bid b under Π' is identically distributed to submitting a bid $\beta(b)$ under Π assuming the miner follows the inclusion rule. Similarly, the miner's revenue when creating a block $\mathbf{B}$ under $\Pi'$is identically distributed to creating a block $\beta(\mathbf{B})$ under Π. Thus, the outcome under an honest execution of Π is identically distributed as the outcome under an honest execution of Π'.

Finally, we show that the induced mechanism Π' satisfies UIC, MIC, and global SCP. Since submitting a bid b and creating a block $\mathbf{B}$ under Π' is equivalent to submitting a bid $\beta(b)$ and creating a block $\beta(\mathbf{B})$ under Π, respectively, the fact that Π' satisfies UIC and global SCP directly follows from the fact that Π satisfies UIC and global SCP under the user's honest bidding rule β.

For proving Π' satisfies MIC, suppose Π' is not MIC for the sake of contradiction. In this case, there exists a bid vector $\mathbf{b}$ and a block $\tilde{\mathbf{B}}$ such that if the miner creates the block $\tilde{\mathbf{B}}$ instead of $\mathbf{I}'(\mathbf{b})$, the miner's utility increases. Then, consider another miner under the mechanism Π, and imagine that the bid vector is $\beta(\mathbf{b})$. Notice that $\mu(\beta(\tilde{\mathbf{B}})) = \mu'(\tilde{\mathbf{B}})$ and $\mu(\mathbf{I}(\beta(\mathbf{b}))) = \mu'(\mathbf{I}'(\mathbf{b}))$. Thus, if the miner creates the block $\beta(\tilde{\mathbf{B}})$ instead of $\mathbf{I}(\beta(\mathbf{b}))$, the miner's utility increases. It violates the fact that Π is MIC. Thus, Π' must satisfy MIC.

□

9.2 Static Revelation Principle: Allowing Bidding Rules that Output Multiple Bids

In this section, we extend Theorem 9.1 to allow the bidding rule to output multiple bids. The basic idea for proving the revelation principle is still the same — trying to bake the non-truth-telling bidding rule β into the mechanism itself. However, when β may output multiple bids, a user i will submit a vector $\mathbf{b}_i$ of bids instead of a single real number, and the inclusion rule may ask the miner

to only select a subset of the bids from $\mathbf{b}_i$. Therefore, the new challenge in this case is how to make the blockchain's confirmation/payment rules also double-check the correct enforcement of β again, given that the blockchain can only see a subset of the bids from $\mathbf{b}_i$.

To handle this difficulty, we need to relax the syntax of the inclusion rule. Earlier in our definitions (Section 2.1), we require that the inclusion rule outputs a subset of the input bid vector. In this section, we slightly relax this syntax requirement, and allow the honest inclusion rule to append some metadata for each bid.[13] The purpose of the metadata in the honest execution is to encode the auxiliary information for the blockchain's confirmation/payment rules. This relaxation in the syntax is non-essential for the impossibility results in this paper, since all the proofs for our impossibility results (Sections 4, 5 and 6.2) still hold even if the honest inclusion rule can append the metadata. Now, we are ready to present the revelation principle.

Theorem 9.2. *Let c be any natural number. Given any non-truthful TFM Π for block size k that is UIC, MIC, and c-SCP with an individually rational bidding rule, there exists a truthful TFM Π' for block size k that is UIC, MIC, and c-SCP such that 1) the honest bidding rule is truth-telling, and 2) given any vector of true values, the outcome under an honest execution of Π is identically distributed as the outcome under an honest execution of Π'.*

Proof of Theorem 9.2. The rest of this section is dedicated to proving Theorem 9.2. Suppose we are given a TFM $\Pi = (\beta, \mathbf{I}, \mathbf{C}, \mathbf{P}, \mu)$ where $\beta : \mathbb{R}_{\geq 0} \to \mathbb{R}^*$ denotes the user's honest bidding strategy, $\mathbf{I}$ denotes the inclusion rule, $\mathbf{C}$ denotes the confirmation rule, $\mathbf{P}$ denotes the payment rule, and μ denotes the miner revenue rule. We construct a truthful mechanism $\Pi' = (\mathbf{I}', \mathbf{C}', \mathbf{P}', \mu')$ as follows:

- Inclusion rule $\mathbf{I}'$: given a bid vector $\mathbf{b} = (b_1, \ldots, b_t)$, let $\beta(\mathbf{b}) = (\beta(b_1), \ldots, \beta(b_t))$. Create a block $\mathbf{B}'$ that includes the following:
  - for each $j \in [t]$, if some bid in the vector $\beta(b_j)$ is selected by the original inclusion rule $\mathbf{I}(\beta(\mathbf{b}))$, then include b_j in $\mathbf{B}'$;
  - for each included bid b_j, attach the following extra annotation in $\mathbf{B}'$:

    $$\mathsf{info}(b_j) = \text{bids in } \beta(b_j) \text{ selected by } \mathbf{I}(\beta(\mathbf{b}))$$

  A *valid block* $\mathbf{B}$ is one in which the annotation for each $b_j \in \mathbf{B}$ is a subset of the vector $\beta(b_j)$.
- The confirmation rule $\mathbf{C}'$, the payment rule $\mathbf{P}'$, and the miner revenue rule μ': If the input block is not valid, then no one is confirmed, and the miner revenue is zero. Else, parse the block as $\{b'_j, \mathbf{s}'_j = \mathsf{info}(b'_j)\}_j$, form an imaginary block $\{\mathbf{s}'_j\}_j$, and run the original confirmation rule $\mathbf{C}$, payment rule $\mathbf{P}$ and miner revenue rule μ on $\{\mathbf{s}'_j\}_j$. A bid b'_j in the input block is considered confirmed if the primary bid in $\beta(b'_j)$ is contained in $\mathbf{s}'_j$ and is confirmed by $\mathbf{C}(\{\mathbf{s}'_j\}_j)$, and its payment is the sum of all payments of $\mathbf{s}'_j$ in $\mathbf{P}(\{\mathbf{s}'_j\}_j)$. The miner is paid $\mu(\{\mathbf{s}'_j\}_j)$.

More intuitively, in the truthful mechanism Π', the miner simulates β by applying β to all bids it has received, resulting in $\beta(\mathbf{b})$. It then simulates the original inclusion rule of the non-truthful mechanism, and the result is $\mathbf{I}(\beta(\mathbf{b}))$, indicating which of $\beta(\mathbf{b})$ would have been included by the original $\mathbf{I}$. The miner includes in the block each bid b_j such that at least one of $\beta(b_j)$ would have been included by $\mathbf{I}(\beta(\mathbf{b}))$, and for each included bid b_j, it attaches the extra information which

[13] We assume the metadata do not occupy the block space. See Theorem 9.5 for more discussion.

bids among $\beta(b_j)$ would have been included by $\mathbf{I}(\beta(\mathbf{b}))$. This way, the blockchain's confirmation, payment, and miner-revenue rules can double-check the correct enforcement of β again, before simulating the effect of the original confirmation, payment, and miner revenue rules of the non-truthful mechanism — formally, the checking is accomplished through a block validity check in our reduction.

Lemma 9.3. *Assume the miner follows the mechanism honestly. Then, the outcome when all users follow the bidding rule β under the mechanism Π is identically distributed as the outcome when all users bid truthfully under the mechanism Π'.*

Proof. Let $(v_1, \ldots, v_t)$ be the true values of all users. Under mechanism Π, if all users follow the bidding rule β, the created block will be $\mathbf{B} = \mathbf{I}(\beta(v_1), \ldots, \beta(v_t))$. For any user i, since the fake bids have no intrinsic values, user i's bid is considered as confirmed if and only if its primary bid in $\beta(v)$ is confirmed. User i's payment is the sum of the payments of all its bids $\beta(v)$. The miner revenue in this case is $\mu(\mathbf{B})$.

Under mechanism Π', if all users bid truthfully, the created block will be $\mathbf{B}'$ parsed as $\{b'_j, \mathsf{info}(b'_j)\}_j$. Notice that $\{\mathsf{info}(b'_j)\}_j$ is identically distributed as $\mathbf{B}$. The bid b'_j under mechanism Π' is considered confirmed if the primary bid in $\beta(b'_j)$ is contained in $\mathsf{info}(b'_j)$ and is confirmed by $\mathbf{C}(\{\mathsf{info}(b'_j)\}_j)$, and its payment is the sum of all payments of $\mathbf{s}'_j$ in $\mathbf{P}(\{\mathsf{info}(b'_j)\}_j)$. The miner revenue in this case is $\mu(\{\mathsf{info}(b'_j)\}_j)$.

Because $\{\mathsf{info}(b'_j)\}_j$ is identically distributed as $\mathbf{B}$, the outcomes of the honest execution of two mechanisms are also identically distributed by checking all the syntax above.

□

Lemma 9.4. *The following statements hold.*

- *If Π is UIC, then Π' is UIC.*
- *If Π is MIC, then Π' is MIC.*
- *For all c, if Π is c-SCP, then Π' is c-SCP.*

Proof. We will prove three properties individually.

Π' is UIC. For the sake of contradiction, suppose Π' is not UIC. That is, under mechanism Π', there exists a user i with true value v and a bid vector $\mathbf{b}_{-i}$ such that user i's utility increases if it submits a vector $\mathbf{f}$ (possibly includes some fake bids) instead of v when other users' bids are $\mathbf{b}_{-i}$. Denote the above as Scenario 1. Then, consider another Scenario 2, where the mechanism is Π, and other users' bids are $\beta(\mathbf{b}_{-i})$. The strategic user i can submit a vector $\beta(\mathbf{f})$ instead of $\beta(v)$.

By Theorem 9.3, the confirmation probability and the payment of the bids $\mathbf{f}$ (v, resp.) in Scenario 1 are identically distributed as the confirmation probability and the payment of the bids $\beta(\mathbf{f})$ ($\beta(v)$, resp.) in Scenario 2. Thus, the expected utility of user i if it submits $\mathbf{f}$ (v, resp.) in Scenario 1 is the same as the expected utility of user i if it submits $\beta(\mathbf{f})$ ($\beta(v)$, resp.) in Scenario 2. Thus, in Scenario 2, user i's utility increases if it $\beta(\mathbf{f})$ instead of $\beta(v)$. It violates UIC under Π, so Π' must satisfy UIC.

Π' is MIC. For the sake of contradiction, suppose Π' is not MIC. That is, there exists a bid vector $\mathbf{b}$ and a block $\tilde{\mathbf{B}} = \{b'_j, \mathbf{s}'_j = \mathsf{info}(b'_j)\}_j$ such that if the miner creates the block $\tilde{\mathbf{B}}$ instead of $\mathbf{I}'(\mathbf{b})$, the miner's utility increases. Denote the above as Scenario 1. Then, consider another Scenario 2, where the mechanism is Π, and the incoming bid vector is $\beta(\mathbf{b})$. The strategic miner can create a block $\{\mathbf{s}'_j\}_j$ instead of $\mathbf{I}(\beta(\mathbf{b}))$.

Recall the miner's utility is the revenue minuses the cost of injecting the fake bids. The miner revenue of the block $\tilde{\mathbf{B}}$ ($\mathbf{I}'(\mathbf{b})$, resp.) in Scenario 1 is the same as the miner revenue of the block $\{\mathbf{s}'_j\}_j$ ($\mathbf{I}(\beta(\mathbf{b}))$, resp.) in Scenario 2. If $\{\mathbf{s}'_j\}_j$ contains some fake bids, the expected payment of those fake bids in Scenario 2 is also the same as the expected payment of the fake bids when the created block is $\tilde{\mathbf{B}}$ in Scenario 1. Therefore, in Scenario 2, if the miner creates $\{\mathbf{s}'_j\}_j$ instead of $\mathbf{I}(\beta(\mathbf{b}))$, the miner's utility increases. It violates MIC under Π, so Π' must satisfy MIC.

Π' is c-SCP for all c. For the sake of contradiction, suppose Π' is not c-SCP for some c. That is, there exist a coalition C formed by the miner and at most c users, a bid vector $\mathbf{b}_{-C}$, and a block $\tilde{\mathbf{B}}$ such that when all non-colluding users' bids are $\mathbf{b}_{-C}$, C's utility increases if the created block is $\tilde{\mathbf{B}}$ instead of $\mathbf{I}'(\mathbf{b}_{-C}, \mathbf{b}_C)$, where $\mathbf{b}_C$ is the true values of all users in C. Denote the above as Scenario 1. Then, consider another Scenario 2, which is the same as scenario 1 except that the mechanism is Π, and the bid vector of non-colluding user is $\beta(\mathbf{b}_{-C})$. Parse $\tilde{\mathbf{B}} = \{b'_j, \mathbf{s}'_j = \mathsf{info}(b'_j)\}_j$. The coalition C can create a block $\{\mathbf{s}'_j\}_j$ instead of $\mathbf{I}(\beta(\mathbf{b}_{-C}, \mathbf{b}_C))$.

Following the same argument as UIC and MIC, we conclude that the joint utility of C of the block $\tilde{\mathbf{B}}$ ($\mathbf{I}'(\mathbf{b}_{-C}, \mathbf{b}_C)$, resp.) in Scenario 1 is the same as the joint utility of C of the block $\{\mathbf{s}'_j\}_j$ ($\mathbf{I}(\beta(\mathbf{b}_{-C}, \mathbf{b}_C))$, resp.) in Scenario 2. Thus, in Scenario 2, if the coalition creates $\{\mathbf{s}'_j\}_j$ instead of $\mathbf{I}(\beta(\mathbf{b}_{-C}, \mathbf{b}_C))$, the joint utility of C increases. It violates c-SCP under Π, so Π' must satisfy c-SCP. □

The proof of Theorem 9.2 directly follows from Theorem 9.3 and Theorem 9.4.

Remark 9.5. We assume the metadata appended by the inclusion rule do not occupy the block space. If the metadata occupy the block space, it is not hard to see that we can encode a valid block using a vector of length at most $3k$. For example, we can divide the block into three parts: the first part encodes users' original bids b_j; the second part encodes $\beta(b_j)$; and the third part encodes the mapping between the first part and the second part.

References

[ACH11] Gilad Asharov, Ran Canetti, and Carmit Hazay. Towards a game theoretic view of secure computation. In *Eurocrypt*, 2011.

[ADGH06] Ittai Abraham, Danny Dolev, Rica Gonen, and Joseph Halpern. Distributed computing meets game theory: Robust mechanisms for rational secret sharing and multiparty computation. In *PODC*, 2006.

[AL11] Gilad Asharov and Yehuda Lindell. Utility dependence in correct and fair rational secret sharing. *Journal of Cryptology*, 24(1), 2011.

[BCD+] Vitalik Buterin, Eric Conner, Rick Dudley, Matthew Slipper, and Ian Norden. Ethereum improvement proposal 1559: Fee market change for eth 1.0 chain. https://github.com/ethereum/EIPs/blob/master/EIPS/eip-1559.md.

[BEOS19] Soumya Basu, David A. Easley, Maureen O'Hara, and Emin Gün Sirer. Towards a functional fee market for cryptocurrencies. *CoRR*, abs/1901.06830, 2019.

[BGR23] Maryam Bahrani, Pranav Garimidi, and Tim Roughgarden. Transaction fee mechanism design with active block producers. *arXiv preprint arXiv:2307.01686*, 2023.

[CCWS21] Kai-Min Chung, T-H. Hubert Chan, Ting Wen, and Elaine Shi. Game-theoretic fairness meets multi-party protocols: The case of leader election. In *CRYPTO*. Springer-Verlag, 2021.

[CGL+18] Kai-Min Chung, Yue Guo, Wei-Kai Lin, Rafael Pass, and Elaine Shi. Game theoretic notions of fairness in multi-party coin toss. In *TCC*, volume 11239, pages 563–596, 2018.

[CMW23] Davide Crapis, Ciamac C Moallemi, and Shouqiao Wang. Optimal dynamic fees for blockchain resources. *arXiv preprint arXiv:2309.12735*, 2023.

[CS23] Hao Chung and Elaine Shi. Foundations of transaction fee mechanism design. In *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 3856–3899. SIAM, 2023.

[DR07] Yevgeniy Dodis and Tal Rabin. Cryptography and game theory. In *AGT*, 2007.

[EFW22] Meryem Essaidi, Matheus V. X. Ferreira, and S. Matthew Weinberg. Credible, strategyproof, optimal, and bounded expected-round single-item auctions for all distributions. In Mark Braverman, editor, *13th Innovations in Theoretical Computer Science Conference, ITCS 2022, January 31 - February 3, 2022, Berkeley, CA, USA*, volume 215 of *LIPIcs*, pages 66:1–66:19, 2022.

[FMPS21] Matheus V. X. Ferreira, Daniel J. Moroz, David C. Parkes, and Mitchell Stern. Dynamic posted-price mechanisms for the blockchain transaction-fee market. *CoRR*, abs/2103.14144, 2021.

[FW20] Matheus V. X. Ferreira and S. Matthew Weinberg. Credible, truthful, and two-round (optimal) auctions via cryptographic commitments. In Péter Biró, Jason D. Hartline, Michael Ostrovsky, and Ariel D. Procaccia, editors, *EC '20: The 21st ACM Conference on Economics and Computation, Virtual Event, Hungary, July 13-17, 2020*, pages 683–712. ACM, 2020.

[GKM+13] Juan A. Garay, Jonathan Katz, Ueli Maurer, Björn Tackmann, and Vassilis Zikas. Rational protocol design: Cryptography against incentive-driven adversaries. In *FOCS*, 2013.

[GKTZ15] Juan Garay, Jonathan Katz, Björn Tackmann, and Vassilis Zikas. How fair is your protocol? a utility-based approach to protocol optimality. In *PODC*, 2015.

[GLR10] Ronen Gradwohl, Noam Livne, and Alon Rosen. Sequential rationality in cryptographic protocols. In *FOCS*, 2010.

[GTZ15] Juan A. Garay, Björn Tackmann, and Vassilis Zikas. Fair distributed computation of reactive functions. In *DISC*, volume 9363, pages 497–512, 2015.

[GY22] Yotam Gafni and Aviv Yaish. Greedy transaction fee mechanisms for (non-) myopic miners. *arXiv preprint arXiv:2210.07793*, 2022.

[GY24] Yotam Gafni and Aviv Yaish. Barriers to collusion-resistant transaction fee mechanisms. *arXiv preprint arXiv:2402.08564*, 2024.

[Har] Jason Hartline. Lectures on optimal mechanism design. http://users.eecs.northwestern.edu/~hartline/omd.pdf.

[HT04] Joseph Halpern and Vanessa Teague. Rational secret sharing and multiparty computation. In *STOC*, 2004.

[Kat08] Jonathan Katz. Bridging game theory and cryptography: Recent results and future directions. In *Theory of Cryptography Conference*, pages 251–272. Springer, 2008.

[KKLP23] Aggelos Kiayias, Elias Koutsoupias, Philip Lazos, and Giorgos Panagiotakos. Tiered mechanisms for blockchain transaction fees. *arXiv preprint arXiv:2304.06014*, 2023.

[KMSW22] Ilan Komargodski, Shin'ichiro Matsuo, Elaine Shi, and Ke Wu. log*-round game-theoretically-fair leader election. In *CRYPTO*, 2022.

[KN08] Gillat Kol and Moni Naor. Cryptography and game theory: Designing protocols for exchanging information. In *TCC*, 2008.

[LRMP23] Stefanos Leonardos, Daniël Reijsbergen, Barnabé Monnot, and Georgios Piliouras. Optimality despite chaos in fee markets. In *International Conference on Financial Cryptography and Data Security*, pages 346–362. Springer, 2023.

[LSZ19] Ron Lavi, Or Sattath, and Aviv Zohar. Redesigning bitcoin's fee market. In *The World Wide Web Conference, WWW 2019*, pages 2950–2956, 2019.

[Mye81] Roger B. Myerson. Optimal auction design. *Math. Oper. Res.*, 6(1), 1981.

[Ndi23] Abdoulaye Ndiaye. Blockchain price vs. quantity controls. *Quantity Controls (July 27, 2023)*, 2023.

[OPRV09] Shien Jin Ong, David C. Parkes, Alon Rosen, and Salil P. Vadhan. Fairness with an honest minority and a rational majority. In *TCC*, 2009.

[PS17] Rafael Pass and Elaine Shi. Fruitchains: A fair blockchain. In *PODC*, 2017.

[Rou20] Tim Roughgarden. Transaction fee mechanism design for the Ethereum blockchain: An economic analysis of EIP-1559. Manuscript, https://timroughgarden.org/papers/eip1559.pdf, 2020.

[Rou21] Tim Roughgarden. Transaction fee mechanism design. In *EC*, 2021.

[SCW23] Elaine Shi, Hao Chung, and Ke Wu. What can cryptography do for decentralized mechanism design? In *ITCS*, volume 251 of *LIPIcs*, pages 97:1–97:22. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2023.

[TY23] Wenpin Tang and David D Yao. Transaction fee mechanism for proof-of-stake protocol. *arXiv preprint arXiv:2308.13881*, 2023.

[WAS22] Ke Wu, Gilad Asharov, and Elaine Shi. A complete characterization of game-theoretically fair, multi-party coin toss. In *Eurocrypt*, 2022.

[WSC24] Ke Wu, Elaine Shi, and Hao Chung. Maximizing Miner Revenue in Transaction Fee Mechanism Design. In Venkatesan Guruswami, editor, *15th Innovations in Theoretical Computer Science Conference (ITCS 2024)*, 2024.

[XFP23] Matheus Venturyne Xavier Ferreira and David C Parkes. Credible decentralized exchange design via verifiable sequencing rules. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, pages 723–736, 2023.

[Yao] Andrew Chi-Chih Yao. An Incentive Analysis of Some Bitcoin Fee Designs (Invited Talk). In *ICALP 2020*.

[ZCZ22] Zishuo Zhao, Xi Chen, and Yuan Zhou. Bayesian-nash-incentive-compatible mechanism for blockchain transaction fee allocation. https://arxiv.org/abs/2209.13099, 2022.

A Comparison of Collusion-Resilience Notions

In this work, we discuss three different definitions of collusion-resilience: c-SCP, global SCP, and OCA-proofness. Both global SCP and OCA-proofness focus on the global coalition consisting of the miner and all users, while c-SCP considers a coalition formed by the miner and at most c users. Both c-SCP and global SCP require following the honest bidding rule is a dominant strategy, so if a TFM satisfies c-SCP for all c, it must also satisfy global SCP. On the other hand, an OCA-proof TFM allows each user to bid according to an individually rational bidding strategy $\sigma : \mathbb{R}_{\geq 0} \to \mathbb{R}$ as long as σ does not depend on others' bids. Hence, for any TFM whose bidding rule only outputs a single real, the TFM satisfies global SCP implies it also satisfies OCA-proofness by simply taking the honest bidding rule as the individually rational bidding strategy σ. However, in Section 2.1, the honest bidding rule can output multiple bids in general. In this case, global SCP may not imply OCA-proofness. Consider the following somewhat contrived mechanism.

Even Auction

Parameters: the block size is infinite.

Mechanism:

- **Bidding rule.** Each user submits two bids $(v, 0)$, where v is the true value.
- **Inclusion rule.** Include all pending bids.
- **Confirmation rule.** If the number of the included bids is odd, no one is confirmed. Otherwise, if the number of the included bids is even, all included bids are confirmed.
- **Payment rule and miner revenue rule.** Everyone pays nothing, and the miner revenue is zero.

Lemma A.1. *The even auction above satisfies global SCP but not OCA-proofness.*

Proof. If everyone follows the bidding rule, the number of the included bids must be even, so everyone's bid is confirmed. Thus, the social welfare is maximized, and the mechanism satisfies global SCP. However, when there are an odd number of users, following any bidding strategy $\sigma : \mathbb{R}_{\geq 0} \to \mathbb{R}$ will lead to an odd number of bids, and no one is confirmed. Thus, it does not satisfy OCA-proofness. □

Notice that whether the individually rational bidding strategy σ outputs a single bid or multiple bids is important. If we require σ to only output a single real as we defined in Definition 2.6, there is no non-trivial TFM that can satisfy UIC, MIC, and OCA-proofness as we proved in Theorem 6.10. However, if we allow σ to output multiple bids, we obtain a feasibility in Section 7.2.

When we analyze whether a TFM satisfies c-SCP for some c, 1-SCP is the weakest requirement since c'-SCP always implies c-SCP for any $c' > c$ by definition. Interestingly, even so, 1-SCP is incomparable with global SCP and OCA-proofness. The relations between 1-SCP, global SCP, and OCA-proofness is depicted in Figure 1.

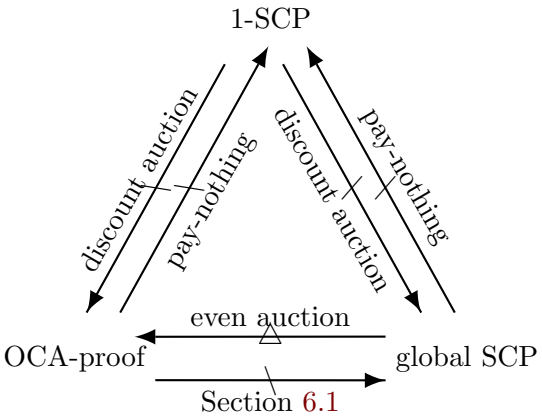


Figure 1: Relationship between collusion-resilience notions.

We explain Figure 1 in more detail below:

- global SCP $\overset{\triangle}{\Longrightarrow}$ OCA-proofness: As we discussed above, if the honest bidding rule only outputs a single real, global SCP implies OCA-proofness; if the honest bidding rule outputs multiple reals, Theorem A.1 states that global SCP does not imply OCA-proofness.
- OCA-proofness $\not\Longrightarrow$ global SCP: In Section 6.1, we give a non-truthful TFM that satisfies UIC, MIC and OCA-proofness. However, in Theorem 5.6, we show that no non-trivial TFM can satisfy UIC, MIC and global SCP no matter it is truthful or non-truthful. Combining these results, we conclude that OCA-proofness does not imply global SCP.
- OCA-proofness $\not\Longrightarrow$ 1-SCP, global SCP $\not\Longrightarrow$ 1-SCP: Conceptually, since a strategic miner can unilaterally decide which transactions are included, the miner can favor the colluding user's bid even if it harms the social welfare. This is formally proven in Theorem A.2.
- 1-SCP $\not\Longrightarrow$ OCA-proofness, 1-SCP $\not\Longrightarrow$ global SCP: The counterexample is given in Theorem A.3. Intuitively, the attack relies on the fact that many colluding users "amortize" the cost of deviation.

Below, we introduce the (somewhat contrived) auctions for constructing the counterexamples.

Pay-Nothing Auction

Parameters: the block size k.

Mechanism:

- **Bidding rule.** Each user submits the true value.

- **Inclusion rule.** Choose the highest up to k bids.
- **Confirmation rule, payment rule and miner revenue rule.** All included bids are confirmed, everyone pays nothing, and the miner revenue is zero.

Lemma A.2. *The pay-nothing auction above satisfies global SCP and OCA-proofness. However, it does not satisfy* 1*-SCP.*

Proof. When everyone follows the mechanism, the users with the highest values will be confirmed, so the social welfare is maximized. Thus, the mechanism satisfies global SCP and OCA-proofness. However, if there are $k+1$ users with the true values $v_1 > \cdots > v_{k+1} > 0$. When everyone follows the mechanism, the $(k+1)$-th user's bid will not be included in the block, so the user's utility is zero. The platform can collude with the $(k+1)$-th user, and include its bid. Since the payment is always zero, the coalition now has positive utility, which violates 1-SCP. □

Discount Auction

Parameters: the fixed price r, and the block size is infinite.

Mechanism:

- **Bidding rule.** Each user submits the true value.
- **Inclusion rule.** Include all pending bids.
- **Confirmation rule.** Let S be the set of all included bids, and define f as

$$f(t) = \begin{cases} r, & \text{if } t \le 10; \\ r/2, & \text{otherwise.} \end{cases}$$

  Let t^* be the largest integer such that there are t^* bids in S larger than $f(t^*)$. All bids larger than $f(t^*)$ are confirmed.
- **Payment rule.** Each confirmed bid pays $f(t^*)$. Unconfirmed bids pay nothing.
- **Miner revenue rule.** The miner revenue is zero, and all payments are burnt.

Lemma A.3. *The discount auction above satisfies* 1*-SCP. However, it does not satisfy global SCP or OCA-proofness.*

Proof. We first show that the mechanism satisfies 1-SCP. Consider any coalition C formed by the miner and user i. Because the miner revenue is always zero, C's joint utility can only come from user i. Fix any bid vector $\mathbf{b}_{-i}$ from other users, user i's confirmation and the payment satisfy the requirements in Myerson's lemma. Thus, overbidding and underbidding cannot increase user i's utility. Next, we show that injecting fake bids does not help either. To increase user i's utility, the only way is to lower the payment. However, regardless of the bids, the payment for each bid is always in $[r/2, r]$, so injecting fake bids can only lower the payment by at most $r/2$. On the other hand, all confirmed bids must pay at least $r/2$, which means user i has to pay at least $r/2$ for each fake bid. Thus, injecting fake bids do not increase user i's utility, so we conclude that the mechanism satisfies 1-SCP.

Now, we show that the mechanism does not satisfy OCA-proofness. Imagine a scenario where there are 10 users each with the true value r. Without injecting any fake bid, the lowest possible payment for each user is $f(10) = r$, so the social welfare is at most 0. However, if the global coalition

inject one fake bid r, everyone's payment will further lower to $f(11) = r/2$. In this case, the social welfare becomes $10r - 11r/2 = 9r/2 > 0$. Thus, the mechanism does not satisfy OCA-proofness.

The mechanism does not satisfy global SCP either since the bidding rule only outputs a single real. If the mechanism satisfied global SCP, the global coalition could pick the suggested bidding rule as the individually rational bidding strategy σ to maximize the social welfare. Given that the mechanism does not satisfy OCA-proofness, it cannot satisfy global SCP either. □

B Alternative Proof of Theorem 5.6

The proof in Section 5 relies on the fact that the miner can deviate from the inclusion rule and choose an arbitrary subset of transactions to include in the block. In particular, Lemma 5.2 assumes that the global coalition can internally simulate the inclusion rule and strategically choose which transactions are included. In this section, we present an alternative proof of Theorem 5.6 which does not rely on deviations from the inclusion rule. As a result, this alternative proof rules out the possibility of achieving UIC, MIC, and global SCP even when the inclusion rule is strictly enforced, for example using trusted hardware.

The two proofs, however, are incomparable. The alternative proof in this section relies on the value domain being unbounded, since it analyzes the behavior of payments and burning as user values tend to infinity (see the proof intuition below). In contrast, the proof in Section 5 still holds even when the value domain is bounded, say $[0, M]$, assuming the confirmation probability under a single bid $a < M$ is non-zero. This assumption is mild. If it does not hold, i.e., if a single bid a has positive confirmation probability only when $a = M$, then its payment must be M by Myerson's lemma. Consequently, the single user's utility is always zero. Moreover, when the value domain is unbounded, we can always find a sufficiently large value M such that confirmation probability is positive as shown in Section 5. Because each proof has its advantages, we include both in this paper.

Proof intuition. Consider first a lonely world with only a single user submitting a true value v. By Myerson's lemma, the user's payment is given by $p(v) = v \cdot x(v) - \int_0^v x(t)\,dt$. We show that $\lim_{v\to\infty} p(v)/v = 0$ (Lemma B.4), i.e., when the bid v is sufficiently large, the user's payment becomes negligible compared to v. Since the miner's revenue is upper bounded by the total payment, the miner's revenue is also negligible in the lonely world.

Now, consider a crowded world where n users each submit the same bid v. We show that for any n, if the mechanism satisfies global SCP, then $\lim_{v\to\infty} Q_n(v)/v = 0$, where $Q_n(v)$ denotes the total amount of burnt coin (Lemma B.3 and Lemma B.4). Conceptually, burnt coins represent the "payment" of the global coalition, so the proof parallels the analysis of a single user's payment.

Next, we show that for any n, the miner revenue $\mu_n(v)$ in the crowded world must satisfy $\lim_{v\to\infty} \mu_n(v)/v = 0$ (Lemma B.5). If this is not true, i.e., $\mu_n(v)$ is large with respect to v, then there must be a user i whose payment is at least $\mu_n(v)/n$, which is also large with respect to v. However, as we have seen, the miner revenue is negligible in the lonely world. Thus, if there is only one user i submitting v, the miner can inject $n-1$ fake bids, inflating user i's payment to $\mu_n(v)/n$. Because the amount of burnt coins is negligible, the miner can extract $\mu_n(v)/n$ for user i, which is much larger than the honest case.

Finally, by global SCP, we show that adding an additional user to the system does not affect the joint utility of the original players (Lemma B.7). As we have shown, the miner's revenue is negligible, so Lemma B.7 implies that each original user's utility almost remains unchanged when an additional user is added. Thus, starting from the lonely world, where the only user has positive

utility u_v, the total social welfare in the crowded world becomes nu_v, growing linearly with n. However, because the block size is finite, this is impossible.

Remark B.1 (On the permissionless setting)**.** In this alternative proof, it is essential to consider the *permissionless setting*, where the number of users is not fixed *a priori* and strategic players may inject an arbitrary number of fake bids. This assumption is critical because the proof does not rely on any deviation from the inclusion rule. Hence, it remains valid even if the miner is forced to follow the inclusion rule exactly.

If, in contrast, we further assume that the number of users is fixed *a priori*, the miner cannot inject any fake bids, and its strategy space becomes empty. In such a restricted setting, MIC would hold trivially. However, the second-price auction satisfies both UIC and global SCP, which implies that the impossibility does not hold.

The rest of this section is devoted to the formal proof. For convenience, we will use the following notation. Consider any number n of users each with the true value v, and let $\mathbf{v}_n = (\underbrace{v, \ldots, v}_{n})$. Let $X_n(v) = \sum_i x_i(\mathbf{v}_n)$ denote the sum of all users' allocation probability, $P_n(v) = \sum_i p_i(\mathbf{v}_n)$ denote the total payment, $\mu_n(v) = \mu(\mathbf{v}_n)$ denote the miner revenue, and $Q_n(v) = P_n(v) - \mu_n(v)$ denote the total burnt coin.

Lemma B.2 (Technical lemma implied by the proof of Myerson's Lemma [Mye81, Har])**.** *Let $f(z)$ be a non-decreasing function. Suppose that $y \cdot (f(z) - f(y)) \le g(z) - g(y) \le z \cdot (f(z) - f(y))$ for any $z \ge y \ge 0$, and moreover, $g(0) = 0$. Then, it must be that*

$$g(z) = z \cdot f(z) - \int_0^z f(t)dt.$$

Lemma B.3. *For any n, if the mechanism satisfies global SCP, then $X_n(\cdot)$ must be non-decreasing, and the function $Q_n(\cdot)$ satisfies*

$$Q_n(v) = v \cdot X_n(v) - \int_0^v X_n(t)dt.$$

Proof. If all users bid b instead of their true value v, the social welfare is $v \cdot X_n(b) - Q_n(b)$. Consider any $0 \le y \le z$, by global SCP, it must be $y \cdot X_n(y) - Q_n(y) \ge y \cdot X_n(z) - Q_n(z)$, and $z \cdot X_n(z) - Q_n(z) \ge z \cdot X_n(y) - Q_n(y)$. Thus, we obtain the following "burning sandwich" inequality:

$$y \cdot [X_n(z) - X_n(y)] \le Q_n(z) - Q_n(y) \le z \cdot [X_n(z) - X_n(y)],$$

which implies that $X_n(\cdot)$ must be non-decreasing. Because $p_i(0) = 0$ for any user i, we have $Q_n(0) = 0$. By Lemma B.2, we have

$$Q_n(v) = v \cdot X_n(v) - \int_0^v X_n(t)dt.$$

□

Lemma B.4. *Consider any natural number k and any non-decreasing function $f(z) : \mathbb{R}_{\ge 0} \to [0, k]$. We define function $g(\cdot)$ as*

$$g(v) = v \cdot f(v) - \int_0^v f(t)dt.$$

Then, it holds that $\lim_{v \to \infty} g(v)/v = 0$.

Proof. It is equivalent to show that

$$\lim_{v\to\infty}\left[f(v)-\frac{1}{v}\int_0^v f(t)dt\right]=0.$$

It is well known that a bounded monotonic function must have a finite limit at infinity. Thus, we have $\lim_{v\to\infty} f(v)=L$ for some $L\in[0,k]$.

Next, we will show that $\lim_{v\to\infty}\frac{1}{v}\int_0^v f(t)dt=L$. In other words, we need to show that for any $\delta>0$, we can find a real number v^* such that $\frac{1}{v}\int_0^v f(t)dt\in[L-\delta,L+\delta]$ for all $v>v^*$. Because $\lim_{v\to\infty} f(v)=L$, there exists a real number v_0 such that $f(v)\in[L-\delta,L+\delta]$ for all $v>v_0$, so we have

$$(L-\delta)\cdot(v-v_0)\le\int_{v_0}^v f(t)dt\le(L+\delta)\cdot(v-v_0).$$

Because $\lim_{v\to\infty}(v-v_0)/v=1$, we have

$$\lim_{v\to\infty}\frac{1}{v}\int_{v_0}^v f(t)dt\in[L-\delta,L+\delta].$$

We can split the integral as

$$\frac{1}{v}\int_0^v f(t)dt=\frac{1}{v}\int_0^{v_0} f(t)dt+\frac{1}{v}\int_{v_0}^v f(t)dt.$$

The term $\int_0^{v_0} f(t)dt$ is a constant independent of v, so we have $\lim_{v\to\infty}\frac{1}{v}\int_0^{v_0} f(t)dt=0$. Combining the argument above, we have

$$L-\delta\le\lim_{v\to\infty}\frac{1}{v}\int_0^v f(t)dt=\lim_{v\to\infty}\frac{1}{v}\int_0^{v_0} f(t)dt+\lim_{v\to\infty}\frac{1}{v}\int_{v_0}^v f(t)dt\le L+\delta.$$

Because the argument above holds for any $\delta>0$, we have $\lim_{v\to\infty}\frac{1}{v}\int_0^v f(t)dt=L$. Therefore, we have

$$\lim_{v\to\infty}\left[f(v)-\frac{1}{v}\int_0^v f(t)dt\right]=0,$$

so we conclude the proof. □

Lemma B.5. *Suppose the mechanism satisfies UIC, MIC, and global SCP. Then, for any n, it must be* $\lim_{v\to\infty}\mu_n(v)/v=0$.

Proof. Consider any case where there are n users all submitting v. Because the total payment is $P_n(v)$, there must exist a user whose payment is at least $P_n(v)/n$.

Now, imagine that there is only one user with identity i submitting v, so the miner's utility is $\mu_1(v)$ in the honest execution. The miner randomly chooses $n-1$ fake identities $\mathcal{I}$, and checks whether i's expected payment $p_i(\mathbf{v}_n)$ is at least $P_n(v)/n$ given n identities $\mathcal{I}\cup\{i\}$ all submitting v. If not, the miner plays honestly without injecting any fake bids. Otherwise, the miner injects $n-1$ fake bids v on behalf of $\mathcal{I}$. In this case, i's payment becomes $p_i(\mathbf{v}_n)$, and the miner's utility is $p_i(\mathbf{v}_n)-Q_n(v)$. By MIC, it must be

$$\mu_1(v)\ge p_i(\mathbf{v}_n)-Q_n(v)\ge\frac{P_n(v)}{n}-Q_n(v).\tag{3}$$

Because the honest miner revenue is upper bounded by the payment, we have $\mu_1(v)\le P_1(v)$ and $\mu_n(v)\le P_n(v)$. Together with Equation (3), we have

$$\mu_n(v)\le n\cdot[P_1(v)+Q_n(v)].\tag{4}$$

Next, by UIC and Myerson's lemma, the only user i's payment is subject to $P_1(v) = v \cdot x(v) - \int_0^v x(v)dt$. Moreover, by Lemma B.3, we have $Q_n(v) = v \cdot X_n(v) - \int_0^v X_n(t)dt$. Thus, by Lemma B.4, we have $\lim_{v\to\infty} P_1(v)/v = 0$ and $\lim_{v\to\infty} Q_n(v)/v = 0$. Because the miner revenue must be non-negative, fix any n, Equation (4) implies $\lim_{v\to\infty} \mu_n(v)/v = 0$. □

Next, we will show that for any user i, no matter how user i bids, the joint utility of the rest of the world does not change. Formally, given a bid vector $\mathbf{b}$ and a user i, let $\mathsf{SW}(\mathbf{b})$ be social welfare, i.e., the sum of the utilities of all users and the miner. Let $\mathsf{SW}_{-i}(\mathbf{b}) = \mathsf{SW}(\mathbf{b}) - \mathsf{util}_i(\mathbf{b})$, i.e., the joint utility of the miner and all users except user i.

Lemma B.6. *Consider any (possibly randomized) TFM satisfying UIC and global SCP. Then, for any bid vector $\mathbf{b}_{-i}$ of an arbitrary length, for any user i, and for any b_i, b_i', it holds that*

$$\mathsf{SW}_{-i}(\mathbf{b}_{-i}, b_i) = \mathsf{SW}_{-i}(\mathbf{b}_{-i}, b_i').$$

Proof. Because the mechanism satisfies UIC, user i's payment is subject to Myerson's lemma. Thus, given any bid vector $\mathbf{b}_{-i}$ from other users, user's payment is $p_i(\mathbf{b}_{-i}, b_i) = b_i \cdot x_i(\mathbf{b}_{-i}, b_i) - \int_0^{b_i} x_i(\mathbf{b}_{-i}, t)dt$. Suppose user i's true value is b_i. For any b_i', if user i bids b_i' instead of b_i, user i's loss can be represented as

$$\begin{aligned}\mathsf{util}_i(\mathbf{b}_{-i}, b_i) - \mathsf{util}_i(\mathbf{b}_{-i}, b_i') &= [b_i \cdot x_i(\mathbf{b}_{-i}, b_i) - p_i(\mathbf{b}_{-i}, b_i)] - [b_i \cdot x_i(\mathbf{b}_{-i}, b_i') - p_i(\mathbf{b}_{-i}, b_i')] \\ &\le [b_i' - b_i][x_i(\mathbf{b}_{-i}, b_i') - x_i(\mathbf{b}_{-i}, b_i)].\end{aligned}$$

Next, By global SCP, for any bid vector $\mathbf{b}_{-i}$, it must be

$$\mathsf{SW}_{-i}(\mathbf{b}_{-i}, b_i) + \mathsf{util}_i(\mathbf{b}_{-i}, b_i) \ge \mathsf{SW}_{-i}(\mathbf{b}_{-i}, b_i') + \mathsf{util}_i(\mathbf{b}_{-i}, b_i').$$

Notice that the argument above holds for any b_i and b_i'. Thus, for any b_i and b_i', we have

$$\mathsf{SW}_{-i}(\mathbf{b}_{-i}, b_i') - \mathsf{SW}_{-i}(\mathbf{b}_{-i}, b_i) \le [b_i' - b_i][x_i(\mathbf{b}_{-i}, b_i') - x_i(\mathbf{b}_{-i}, b_i)]. \tag{5}$$

Now, we can divide the interval $[b_i, b_i']$ into L equally sized segments $b_i^{(0)}, \ldots, b_i^{(L)}$ where $b_i^{(0)} = b_i$ and $b_i^{(L)} = b_i'$. By Equation (5), we have

$$\begin{aligned}\mathsf{SW}_{-i}(\mathbf{b}_{-i}, b_i') - \mathsf{SW}_{-i}(\mathbf{b}_{-i}, b_i) &= \sum_{j=0}^{L-1} \mathsf{SW}_{-i}(\mathbf{b}_{-i}, b_i^{(j+1)}) - \mathsf{SW}_{-i}(\mathbf{b}_{-i}, b_i^{(j)}) \\ &\le \sum_{j=0}^{L-1} (b_j^{(j+1)} - b_j^{(j)}) \cdot \left[x_i(\mathbf{b}_{-i}, b_i^{(j+1)}) - x_i(\mathbf{b}_{-i}, b_i^{(j)})\right] \\ &= \frac{b_i' - b_i}{L} \cdot \left[x_i(\mathbf{b}_{-i}, b_i') - x_i(\mathbf{b}_{-i}, b_i)\right].\end{aligned}$$

By taking the limit for $L \to \infty$, we have $\mathsf{SW}_{-i}(\mathbf{b}_{-i}, b_i') - \mathsf{SW}_{-i}(\mathbf{b}_{-i}, b_i) \le 0$. Because the argument above holds for any b_i and b_i', it must be $\mathsf{SW}_{-i}(\mathbf{b}_{-i}, b_i) = \mathsf{SW}_{-i}(\mathbf{b}_{-i}, b_i')$. □

Lemma B.7. *Consider any (possibly randomized) TFM satisfying UIC and global SCP. Then, for any bid vector $\mathbf{b}_{-i}$ of an arbitrary length, for any user i, and for any b_i, it holds that*

$$\mathsf{SW}_{-i}(\mathbf{b}_{-i}) = \mathsf{SW}_{-i}(\mathbf{b}_{-i}, b_i).$$

Proof. First, we show that $\mathsf{SW}_{-i}(\mathbf{b}_{-i}) = \mathsf{SW}_{-i}(\mathbf{b}_{-i}, 0_i)$. To see this, if $\mathsf{SW}_{-i}(\mathbf{b}_{-i}) < \mathsf{SW}_{-i}(\mathbf{b}_{-i}, 0_i)$, then consider a world where users' true values are $\mathbf{b}_{-i}$. The global coalition can inject a fake bid 0_i, and the social welfare increases. On the other hand, if $\mathsf{SW}_{-i}(\mathbf{b}_{-i}) > \mathsf{SW}_{-i}(\mathbf{b}_{-i}, 0_i)$, then consider a world where users' true values are $(\mathbf{b}_{-i}, 0_i)$. Because user i's utility is zero under an honest execution, we have $\mathsf{SW}_{-i}(\mathbf{b}_{-i}, 0_i) = \mathsf{SW}(\mathbf{b}_{-i}, 0_i)$. If user i drops out, the social welfare becomes $\mathsf{SW}_{-i}(\mathbf{b}_{-i})$, which is larger than the honest execution. Thus, it must be $\mathsf{SW}_{-i}(\mathbf{b}_{-i}) = \mathsf{SW}_{-i}(\mathbf{b}_{-i}, 0_i)$. The lemma directly follows from the argument above and Lemma B.6. □

Theorem B.8. *No non-trivial, possibly randomized truthful TFM can simultaneously satisfy UIC, MIC, and global SCP (all in the ex post sense) when the block size is finite.*

Proof. For the sake of contradiction, suppose the mechanism is non-trivial; that is, there exists a scenario $\mathbf{b} = (b_1, \ldots, b_t)$ such that some user i has positive allocation probability $p_i(\mathbf{b})$. We first show that for a sufficiently large v_0, the allocation probability under a single bid v_0 is positive. Let v_0 be any real number such that $p_i(\mathbf{b}) \cdot v_0 > \sum_{j=1}^{t} b_j$. Then, imagine that there is only one user with true value v_0. If the allocation probability $x(v_0)$ is zero, the global coalition can replace the user's primary bid with b_i, inject the fake bids $\mathbf{b}_{-i}$, and pretend that the bid vector is $\mathbf{b}$. This way, the user would get positive allocation probability $p_i(\mathbf{b})$. Because its expected payment is at most $\sum_{j=1}^{t} b_j$, the user's expected utility (and thus the social welfare) is positive. This violates global SCP, so the allocation probability $x(v_0)$ must be positive.

Let u_v denote the user's utility when there is only one user submitting its true value v. By Myerson's lemma, it must be $u_v = \int_0^v x(t)dt$. Because $x(v_0) > 0$, we can find a sufficiently large value v_1 such that $u_{v_1}/v_1 > x(v_0)/2$. Moreover, because $x(v)$ is non-decreasing, we have $u_v/v \geq u_{v_1}/v_1$ for any $v \geq v_1$.

Next, by weak symmetry, when n users all bid v, the social welfare is $\mathsf{SW}(\mathbf{v}_n)$ regardless of the users' identities. Thus, by Lemma B.7, for any user identity i, we have $\mathsf{SW}(\mathbf{v}_n) = \mathsf{SW}_{-i}(\mathbf{v}_{n+1})$. Because $\mathsf{SW}(\mathbf{v}_{n+1}) = \mathsf{SW}_{-i}(\mathbf{v}_{n+1}) + \mathsf{util}_i(\mathbf{v}_{n+1})$, we have $\mathsf{util}_i(\mathbf{v}_{n+1}) = \mathsf{SW}(\mathbf{v}_{n+1}) - \mathsf{SW}(\mathbf{v}_n)$. In other words, given any v and n, every user's utility is the same regardless of the identity. Given any v and n, the sum of all users' utilities is $\mathsf{SW}(\mathbf{v}_n) - \mu(\mathbf{v}_n)$, so a user's utility is $\mathsf{util}(\mathbf{v}_n) = \frac{1}{n}[\mathsf{SW}(\mathbf{v}_n) - \mu(\mathbf{v}_n)]$. Therefore, we have

$$\mathsf{SW}(\mathbf{v}_{n+1}) = \frac{n+1}{n} \cdot \mathsf{SW}(\mathbf{v}_n) - \frac{1}{n} \cdot \mu(\mathbf{v}_{n+1}). \tag{6}$$

By Lemma B.5, for any natural number N, we can find a sufficiently large value $v \geq v_1$ such that $\mu_{n+1}(v)/v < x(v_0)/4$ for all $n + 1 \leq N$. Because $u_v/v \geq u_{v_1}/v_1$ for any $v \geq v_1$, we have

$$\frac{u_v}{v} \geq \frac{u_{v_1}}{v_1} > \frac{x(v_0)}{2} > \frac{2\mu_{n+1}(v)}{v}.$$

Thus, we have $\mu_{n+1}(v) < u_v/2$ for all $n + 1 \leq N$, and Equation (6) implies $\mathsf{SW}(\mathbf{v}_{n+1}) \geq \frac{n+1}{n} \cdot \mathsf{SW}(\mathbf{v}_n) - \frac{1}{2n} \cdot u_v$. When there is only one user submitting its true value v, the social welfare $\mathsf{SW}(\mathbf{v}_1)$ is at least u_v. By the proof of induction, it is easy to see that $\mathsf{SW}(\mathbf{v}_n) \geq n \cdot u_v/2$. Because the block size k is finite, the social welfare $\mathsf{SW}(\mathbf{v}_n)$ must be upper bounded by kv. Thus, we have $u_v/v \leq 2k/n$ for all $n \leq N$. However, as we have shown, for a sufficiently large value v_1, we have $u_v/v \geq u_{v_1}/v_1 > x(v_0)/2$. Consequently, if we choose N such that $N > 4k/x(v_0)$, we have $u_v/v \leq 2k/N < x(v_0)/2$, which leads to a contradiction.

□