Beyond Incentive Compatibility: Rational Harm-Proof Transaction Fee Mechanisms[*]

Forest Zhang[†], Elain Park[‡], and Ke Wu[§]

University of Michigan, Ann Arbor

Abstract

On a blockchain, users compete for scarce block space in an auction run by the miner to get their transactions confirmed in the block. This auction is called *transaction fee mechanism* (TFM). Recent work [Rou21, CS23, SCW23] has been focused on *incentive compatibility* (IC), requiring that honest behavior maximizes the payoff for each type of strategic player: users, the miner, or miner–user coalitions. In this work, we introduce *rational-harm proofness (RHP)*, which rules out any deviation that harms honest parties without also reducing the deviator's own utility. RHP closes a gap left by IC: IC does not forbid utility neutral yet externally harmful deviations. For example, in a second-price auction, the second-highest bidder can increase the winner's payment without affecting their own payoff. Such deviation is eliminated by RHP.

We characterize TFMs satisfying RHP alongside incentive compatibility for users (UIC) and miners (MIC). For finite block size, we develop a *complete characterization* in two models:

- In the *plain model*—where a single miner unilaterally implements the auction—we prove a *tetrilemma* (3-out-of-4 impossibility): among the four desired properties {positive miner revenue, UIC, MIC, RHP against miner–user coalitions}, no mechanism achieves all four simultaneously. Meanwhile, any three are jointly achievable in the plain model.
- In the *MPC-assisted model*—where a committee of miners jointly implement the auction via multi-party computation (MPC)—we construct a randomized TFM with a positive miner revenue that achieves UIC, MIC, and RHP against all three types of strategic players. We further show that randomness is necessary: any deterministic TFM satisfying UIC and RHP in this model must confirm no transactions when the number of users exceeds the block size.

Finally, we show that IC and RHP are *incomparable*: for each strategic role, there are mechanisms satisfying one but not the other in both models. Our results broaden the design objectives for TFMs: beyond incentive compatibility, mechanisms should also preclude costless harm to honest participants.

[*] Author order is randomized.

[†] `forestz@umich.edu`

[‡] `elainpk@umich.edu`

[§] `kewucse@umich.edu`

1 Introduction

Block space on a blockchain is a scarce resource, so users compete with each other to get their transactions confirmed in a block. This process can essentially be viewed as an auction where the block producer, also called the miner, sells the limited k number of block slots. Each user bids to get one slot for its transaction, and a *transaction fee mechanism* (TFM) decides which transactions are confirmed, how much each confirmd transaction pays, and how much revenue the miner receives.

Recent progress [LSZ19, Yao, BEOS19, BCD$^+$, Rou21, FMPS21] observed that, due to the open and decentralized environment of blockchain, TFM design departs significantly from classical auction design. For example, classical auctions typically assume a trusted auctioneer and focus on the strategies of individual users. In contrast, on a blockchain, the auctioneer is a strategic miner who may deviate from the prescribed protocol to gain more revenue. Additionally, smart contracts make it easy for miners and users to enter binding side contracts and split their joint gains off-chain. A growing body of work therefore explores *incentive compatibility* (IC) properties of TFM so that the honest behavior is provably best for all participants. Ideally, a "dream" TFM should satisfy 1.) *user incentive compatibility* (UIC): bidding one's true valuation is optimal for each individual user; 2.) *miner incentive compatibility* (MIC): the miner is incentivized to execute the mechanism honestly; and 3.) *side-contract proofness* (SCP): no miner-user coalition can deviate to jointly profit.

However, incentive compatibility alone does not protect honest participants. There are deviations that leave a strategic player's utility unchanged compared to honest behavior, yet strictly harm others. Standard incentive compatibility notions do not rule out such utility-neutral but externally harmful actions. For example, the gold-standard second-price auction [Vic61] is known to be incentive compatible for individual users: the highest bidder wins and pays the second-highest price. However, the second-highest bidder can raise their bid slightly without changing their own utility (the bidder still loses and pays nothing) but increase the winner's payment. Such manipulations are costless for the deviator but impose real externalities on others. In repeated, competitive environments like blockchains, such "neutral-to-me but harmful-to-you" moves can compound over time. For example, repeatedly inflating a rival's payments reduces their budget and delays confirmations across blocks. This motivates us to ask

Can we design TFMs that achieve desired incentive compatibilities while also protecting honest participants from being harmed by rational deviations?

Related concepts have been studied in classical mechanism design under non-bossiness and robustness to secondary goals [SS81, Tho16, LSZ23, DPTM24, LSZ23]. However, those notions apply to individual agents and do not capture the decentralized setting of TFMs with strategic miners and miner-user coalitions. In this work, we initiate a systematic study of the above question in the context of TFMs. We formalize the requirement that no strategic players (whether a user, miner, or miner-user coalition) can harm any honest participants unless they also harm the deviator themselves as a property called **rational-harm proofness** (RHP). We then characterize when TFMs can achieve RHP alongside the standard incentive guarantees. For practical viability, we also target *positive miner revenue* so that participation is profitable for miners.

Our results apply to two settings: (1) the **plain model**, where a single miner unilaterally determines the block contents, capturing today's mainstream blockchain architecture; and (2) the **MPC-assisted model** [SCW23], where a committee of miners jointly execute a cryptographic primitive, called the multi-party computation (MPC), to implement the TFM.

Prior work have shown an important obstacle: UIC and SCP cannot be achieved simultaneously in either the plain [CS23] or the MPC-assisted model [SCW23]. Since UIC is tightly coupled to user experience, we therefore prioritize UIC and ask how far one can go on RHP while maintaining UIC and MIC (See Section 1.2 for a discussion on practical implication of of such TFMs). We present a complete characterization of when UIC, MIC, and RHP can be achieved together with positive miner revenue in both models. The detailed results are given below.

1.1 Our Results

Model at a Glance. We begin with a high-level model of TFM to contextualize our results. Formal definitions appear in Section 3. We focus on a single block of finite size k. One can view the mechanism as selling k identical block slots. Each user i has a private true value v_i indicating the maximum they are willing to pay for a block slot. Users submit bids for inclusion, and a miner proposes a block of up to k bids. A TFM specifies, for any vector of submitted bids: 1.) which bids (transactions) are included and confirmed in the block; 2.) the payment charged to each confirmed bid; and 3.) the miner's revenue. Importantly, in blockchain, not all payments need to go to the miner: some or even all payments can be *burned*. Once the set of up to k bids to include are fixed[1], the confirmation, payment, and miner revenue rules are executed on-chain.

In the *plain model*, a single miner unilaterally determines which bids to include. In the *MPC-assisted model*, a committee of M miners jointly execute the mechanism via an MPC, as if a trusted party were running the mechanism honestly on all bids. The MPC ensures that if at least one miner is honest, then no coalition of up to $M-1$ miners can tamper with the outcome. In particular, the outcome is either correctly computed on all received bids or the protocol aborts—resulting in no bids being confirmed and no miner revenue. The auction is not rerun if the protocol aborts as we focus on myopic players who derive utility only from a single block.

We require the mechanism to satisfy three basic properties: 1.) *Individual rationality*: no bid pays more than its value; 2.) *Budget feasibility*: the miner's revenue does not exceed the total payments paid by users, i.e., the mechanism does not create money.[2]; 3.) *Weak symmetry*: metadata such as the identity or timestamp of a bid is used only for tie-breaking; otherwise, outcomes depend only on bid values. For later reference, we note that all of our impossibility results hold assuming only weak symmetry, but all of our feasibility results are strongly symmetric, meaning that all bids with the same value always receive the same treatment, without a tie-breaking based on metadata.

Strategy Space. We focus on direct-revelation mechanisms, in which honest users submit a single bid equal to their true values. An honest miner includes up to k bids as prescribed in the mechanism, without censoring honest bids or injecting fake ones.

Strategic players may deviate to increase their utility. We consider an *ex post* setting where a strategic player can observe all honest bids before choosing a strategy. We consider three types of strategic players: 1.) A single user. A strategic user can bid an arbitrary value for its transaction, refuse to bid, and/or inject any number of *fake bids* since pseudonyms are easy to register on-chain. 2.) Strategic miner(s). In the plain model, the single miner can inject fake bids, drop honest users' bids, and select any set of up to k bids for the block. In the MPC-assisted model, a coalition of up to $M-1$ miners can inject fake bids or cause the protocol to abort, but they cannot selectively censor honest users' bids or alter the outcome if the protocol does not abort. 3.) A miner-user coalition, containing the miner and some users in the plain model, or up to $M-1$ miners together with some users. A coalition's strategy is a combination of its members' strategies.

[1] There is a subtle distinction between a bid being included in a block versus being confirmed. A bid must be included to have a chance of confirmation, but depending on the mechanism, not all included bids are necessarily confirmed; see Section 3 for detailed explanation.

[2] The miner gets a fixed block reward independent from the transaction fees, so we do not model the block reward.

Utility. A user with true value v gets utility $v - p$ if its transaction is confirmed and pays p, and 0 if not confirmed. A miner's utility is its revenue minus any cost it pays for its own fake bids. A coalition's utility is the sum of its members' utilities.

Conceptual Contribution: Defining RHP. Our first contribution is conceptual: we elevate the principle of "no costless harm" to a first-class requirement for TFM design. We formalize a new safety notion, *rational-harm proofness (RHP)*, which guarantees that rational players cannot make any honest participant strictly worse off unless the deviators also hurt themselves in the process. As an implication, if a strategy is profitable for the deviators, it constitutes a Pareto improvement relative to honest behavior from the perspective of all participants.

Formally, for any strategic player $\mathcal{C}$, whether an individual user, miner, or miner-user coalition, let $H_\mathcal{C}$ denote $\mathcal{C}$'s honest strategy. For any true value profile $\mathbf{v} = (v_1, ..., v_n)$ of all the users, let $\mathsf{util}_i(\mathbf{v}; S_\mathcal{C})$, $\mathsf{mutil}_j(\mathbf{v}; S_\mathcal{C})$, and $\mathsf{util}_\mathcal{C}(\mathbf{v}; S_\mathcal{C})$ denote the expected utility of a user i, a miner j (in the plain model there is only one miner), and the coalition $\mathcal{C}$, respectively, in the randomized experiment where players in $\mathcal{C}$ adopts some (possibly randomized) strategy $S_\mathcal{C}$ and players outside $\mathcal{C}$ behaves honestly.

Definition 1.1 (Rational-harm proofness)**.** We say that a TFM satisfies rational-harm proofness (RHP) against a strategic player or coalition, denoted as $\mathcal{C}$, iff for any true value vector $\mathbf{v}$ of users, for any strategy $S_\mathcal{C}$ for coalition $\mathcal{C}$ such that $\mathsf{util}_\mathcal{C}(\mathbf{v}; S_\mathcal{C}) \geq \mathsf{util}_\mathcal{C}(\mathbf{v}; H_\mathcal{C})$, it must be that

$$\mathsf{util}_i(\mathbf{v}; S_\mathcal{C}) \geq \mathsf{util}_i(\mathbf{v}; H_\mathcal{C})$$

for any honest user $i \notin \mathcal{C}$. Additionally, for any miner $j \notin \mathcal{C}$, it must be that

$$\mathsf{mutil}_j(\mathbf{v}; S_\mathcal{C}) \geq \mathsf{mutil}_j(\mathbf{v}; H_\mathcal{C}).$$

Based on the three types of strategic players, we say that a TFM satisfies

- *User rational-harm proofness (URHP)* iff Definition 1.1 holds when $\mathcal{C} = \{i\}$ contains an individual user i. Note that in this case, a strategic user cannot harm an honest miner as well.
- *Miner rational-harm proofness (MRHP)* iff Definition 1.1 holds when $\mathcal{C}$ only contains the single miner in the plain model, or any coalition of up to $M - 1$ miners in the MPC-assisted model.
- *d-coalition rational-harm proofness (d-CRHP)* iff Definition 1.1 holds when $\mathcal{C}$ contains the miner and at most d users in the plain model, or $\mathcal{C}$ includes up to $M - 1$ miners and at most d users in the MPC-assisted model. We require that coalitions contain at least one user.

RHP is broadly applicable and extends beyond TFMs. For example, in auctions or online marketplaces, RHP eliminates strategies where a participant maliciously raise prices or block outcomes to hurt competitors without affecting their own payoff. At its core, RHP protects honest participants and ensures that they will not be harmed by any rational actors who seeks to maximizes their own payoffs.

We also define a weaker notion called *weak RHP*, which only requires that any *strictly profitable* deviation for $\mathcal{C}$ does not harm honest participants. This is strictly weaker than RHP and strictly weaker than IC, since IC rule out the existence of any strictly profitable deviations.

Landscape in the Plain Model. We first consider the plain model, where a single miner determines which bids to include. This setting reflects the architecture of most mainstream blockchains. The game proceeds as follows: honest users submit their bids truthfully; any strategic players then choose their strategies after observing the honest bids; the miner selects up to k bids for the block. The blockchain protocol then honestly confirms some of these bids and determines payments and miner revenue according to the mechanism's rules.

Our goal is to design a TFM that satisfies *all* of the following properties in the plain model:

- UIC: a user's utility is maximized by bidding its true value. This ensures a good user experience.
- MIC: the single miner maximizes its utility by following the prescribed mechanism honestly.
- URHP, MRHP, CRHP: no rational players can harm honest players without harming themselves.
- Positive miner revenue: the miner gets a positive revenue in some scenarios, so that honest participation is financially profitable.

We do not ask for SCP (incentive compatibility against miner-user coalitions) due to the impossibility that UIC and SCP cannot both be true [CS23, SCW23]. Instead, we focus on CRHP as the collusion-resilience notion.

Unfortunately, it is impossible to achieve all the desired criteria in the plain model. In fact, we develop the following **"tetrilemma: 3-out-of-4 impossibility"**: among the four properties of UIC, MIC, 1-CRHP, and positive revenue, at most three can simultaneously be true in the plain model. Formally, on the impossibility side,

Theorem 1.2 (Impossibility in the plain model)**.** *Let k be the finite block size. In the plain model, any UIC, MIC, and* 1*-CRHP TFM must confirm no bids when there are more than $k + 1$ users. This implies that the miner revenue must always be zero regardless of the input bids.*

In fact, even if we relax 1-CRHP to weak 1-CRHP, the miner revenue still must always be zero.

Theorem 1.3 (Impossibility under weak-CRHP)**.** *The miner revenue in any UIC, MIC, and weak* 1*-CRHP TFM in the plain model must always be zero.*

Corollary 1.4. *Only trivial TFMs, where no bids are ever confirmed, satisfy UIC, MIC, MRHP, and (weak)* 1*-CRHP in the plain model.*

On the positive side, we identify concrete TFMs that achieve any three out of the four desired properties, indicating that the above trade-off is tight.

Mechanism 1.5 (All-or-nothing posted-price)**.** *Let $\mathsf{res} \geq 0$ be the reserved price and $0 \leq \epsilon \leq \mathsf{res}$ be the revenue parameter. When $\epsilon = 0$, we call this mechanism "burning all-or-nothing". If there are no more than k bids and every bid is strictly higher than res, then include and confirm all bids. Otherwise, include and confirm no bids. Each confirmed bid pays the reserved price res and miner gets ϵ per each confirmed bid.*

Mechanism 1.6 (Posted-price with random selection)**.** *Let $\mathsf{res} > 0$ be the reserved price and $0 < \epsilon \leq \mathsf{res}$ be the revenue parameter. Randomly choose k bids that are strictly higher than res to confirm. If less than k bids are strictly higher than res, confirm all of them. Each confirmed bid pays the reserved price res and miner gets ϵ per each confirmed bid.*

Mechanism 1.7 (First-price auction)**.** *The highest k bids gets confirmed and pays their bids. Break ties arbitrarily. All payments go to the miner.*

Theorem 1.8 (Feasibility in the plain model)**.** *Let k denote the block size. In the plain model,*

- *Mechanism 1.5 with $\epsilon = 0$ (burning all-or-nothing) achieves UIC, MIC, and d-CRHP for any $d \geq 1$, but miner revenue is always zero.*
- *Mechanism 1.5 with $\epsilon > 0$ (all-or-nothing posted-price) achieves UIC, d-CRHP for any $d \geq 1$, and positive revenue.*
- *Mechanism 1.6 (posted-price with random selection) achieves UIC, MIC, and positive revenue.*
- *Mechanism 1.7 (first-price auction) achieves MIC, d-CRHP for any $d \geq 1$, and positive revenue.*

Table 1 summarizes the properties of Mechanism 1.5, 1.6, 1.7 in the plain model. Interestingly, Mechanism 1.5 exhibits a sharp phase-transition depending on the revenue parameter ϵ. With $\epsilon = 0$, the mechanism is MIC but not MRHP: the miner is happy to follow the protocol, yet can take actions that reduce some honest users' utility without harming itself since its own revenue is 0 anyway. In contrast, with $\epsilon > 0$, it is MRHP (the miner cannot harm honest parties without harming itself) but not MIC: the miner now has a profitable deviation as explained below.

Consider a scenario with more than k users whose true values all exceed res. In an honest execution of Mechanism 1.5 with $\epsilon > 0$, no bids would be confirmed and miner gets no revenue. However, the miner can deviate by pretending that it only sees k bids (of its choice) that are above res by dropping other users' bids. The miner could even run an off-chain auction to decide which k bids it "sees". Under this deviation, exactly k bids get confirmed, each paying res and the miner gets $k \cdot \epsilon$ revenue on-chain, plus any potential off-chain revenue. Crucially, this deviation does not hurt any honest user since no one is confirmed in the honest case. The deviation in fact allows k users to benefit from the deviation and enjoy positive utility. Thus MRHP is not violated , even though the mechanism's credibility is undermined by the miner's profitable deviation. This example shows why we seek to enforce IC in addition to RHP: RHP alone does not prevent such undetectable but profit-increasing miner deviations. We further discuss these considerations in Section 1.2.

Table 1: Properties of Mechanism 1.5, 1.6, and 1.7 in the Plain Model. Here AoN stands for all-or-nothing. The last column of Rev indicates whether the mechanism achieves positive miner revenue.

Mechanism	**UIC**	**URHP**	**MIC**	**MRHP**	**SCP**	**CRHP**	**Rev**
Mechanism 1.5 ($\epsilon = 0$)	✓	✓	✓	✗	✗	✓	✗
Mechanism 1.5 ($\epsilon > 0$)	✓	✓	✗	✓	✗	✓	✓
Mechanism 1.6	✓	✓	✓	✗	✗	✗	✓
Mechanism 1.7	✗	✗	✓	✓	✓	✓	✓

Landscape in the MPC-assisted Model. In the MPC-assisted model, M miners jointly execute the TFM and share the total revenue. Strategic miners can still inject fake bids, but the MPC prevents them from altering the computed outcome once all bids are submitted. It is convenient to think of an ideal functionality $\mathcal{F}$ that honestly implements the mechanism as follows:

Fix an arbitrary strategic player $\mathcal{C}$, which in the MPC-assisted model could be an individual user, a coalition of up to $M - 1$ miners, or a miner-user coalition of up to $M - 1$ miners and some users. Honest users submit their bids to $\mathcal{F}$. After seeing these honest bids, the strategic coalition $\mathcal{C}$ decides what bids to submit. Then $\mathcal{F}$ computes the TFM outcome (which bids are confirmed, payments, miner revenue) based on all submitted bids. If $\mathcal{C}$ includes more than $M/2$ of the miners, $\mathcal{F}$ first sends the outcome to $\mathcal{C}$; if $\mathcal{C}$ approves, the outcome is released to everyone, otherwise the auction aborts. No one gets confirmed, and miner gets no revenue. *No rerun occurs of the auction if the protocol aborts.* If $\mathcal{C}$ has fewer than $M/2$ miners, $\mathcal{F}$ directly sends the outcome to all players. This ideal functionality can be instantiated by real-world protocols [GMW87, Can01].

In the MPC-assisted model, we can simultaneously achieve all the desire properties, UIC, MIC, all three RHP variabnts, as well as positive miner revenue.

Theorem 1.9 (Feasibility in the MPC-assisted model)**.** *Mechanism 1.6 (posted-price with random selection) achieves UIC, MIC, URHP, MRHP, d-CRHP for any $d \geq 1$, and positive miner revenue in the MPC-assisted model.*

A natural question is whether the use of randomness is necessary for the above result. Can we achieve all properties with a deterministic mechanism? Interestingly, Mechanism 1.5 with $0 < \epsilon$ does achieve all properties in the MPC-assisted model, but it is somewhat trivial in the sense that it confirms no bids whenever there are more than k users. In fact, this triviality is unavoidable for any deterministic mechanism:

Theorem 1.10 (Characterization of deterministic mechanism in the MPC-assisted model)**.** *Let k denote the block size. Mechanism 1.5 (all-or-nothing posted-price) with $0 < \epsilon \leq \mathsf{res}$ achieves UIC, MIC, URHP, MRHP, d-CRHP for any $d \geq 1$, and positive revenue in the MPC-assisted model.*

Meanwhile, no deterministic UIC and URHP TFM in the MPC-assisted model can confirm any bid when there are more than k users.

IC and RHP are Incomparable. Perhaps surprisingly, we find that incentive compatibility (IC) and rational-harm proofness (RHP) are fundamentally *incomparable* properties: neither implies the other. Intuitively, IC guarantees no player has a profitable deviation, but it does not rule out deviations that leave the deviator's utility unchanged while harming others. Conversely, RHP prevents any deviation that harms others unless it also hurts the deviator, but a deviator might still profit via a strategy that harms no one else.

For example, the classic second-price auction is UIC but not URHP: the second-highest bidder can slightly raise their bid, leaving their own outcome unchanged while increasing the winner's payment. Conversely, a *burning first-price auction* (same as Mechanism 1.7 except that all payments are burned so the miner earns zero revenue) is URHP: no user can make an honest user pay more or lose confirmation without reducing their own utility, and the miner cannot be harmed because it always earns 0. However, this mechanism is not UIC: a user can profit by underbidding just enough to still be among the top k bids, thereby paying less. In general, we prove:

Theorem 1.11 (IC and RHP are incomparable)**.** *In both the plain and the MPC-assisted model, IC and RHP are incomparable. For each type $X \in \{$user, miner, miner-user coalition$\}$, there exist mechanisms that are X-IC but not X-RHP, and mechanisms that are X-RHP but not X-IC.*

We defer the full catalog of mechanisms demonstrating this incomparability to Section 6, where we summarize the results in Table 2 for the plain model and Table 3 for the MPC-assisted model.

1.2 Design Implications and Rationale

Prioritize UIC over SCP. Prior work has shown that UIC and SCP (incentive compatibility against miner-user coalitions) cannot simultaneously hold in either model. In this work, we choose to prioritize UIC over SCP for user experience. Non-UIC mechanisms like the first-price auction used in Bitcoin force users to guess the right bid to avoid overpaying, leading to volatile fees and poor user experience. By contrast, a UIC mechanism offers a simple experience: users can simply bid their true value. In fact, Ethereum's adoption of the EIP-1559 fee mechanism replaced the first-price auction with a posted-price model specifically to simplify fee estimation for users [BCD$^+$].

Scope of the Posted-Price with Random Selection. As Theorem 1.9 shows, Mechanism 1.6 satisfies UIC, MIC, and all variants of RHP while achieving positive revenue in the MPC-assisted model. It does *not* satisfy SCP. To see this, suppose fewer than k users have value above the reserve res, resulting in some slack capacity in the block, and a colluding user has true value in the range $(\mathsf{res} - \epsilon/M, \mathsf{res})$. A miner colluding with this user can profit from the following deviation: The user overbids to get confirmed and pay res, and the miner gains an extra ϵ/M revenue from this confirmation. This deviation strictly improves the coalition's joint utility, although it is only beneficial if the colluding user's true value is very close to res. In general, any profitable deviation of miner-user coalitions in this mechanism falls into these two types: 1.) filling otherwise unused block space as above, or 2.) having some low-value colluders drop out to raise the confirmation probability of higher-value colluders. Crucially, none of these deviations harm any honest user, so CRHP is preserved. In essence, Mechanism 1.6 offloads any residual efficiency improvements to strategic miner-user coalitions. The coalition's may gain either by utilizing slack capacity or redistributing confirmations among colluding users, but they must not hurt outsiders. Meanwhile, individual user or coalition of up to $M-1$ miners have no incentive to deviate.

Choosing the Reserve and Revenue Parameters. For real-world deployment, where the mechanism runs repeatedly block-by-block, the reserve price res in Mechanism 1.6 can be adjusted dynamically, as in Ethereum's EIP-1559 [BCD$^+$]. The protocol tracks block demand and raises or lowers res for the next block to target a stable throughput. The revenue parameter ϵ should be chosen *sufficiently small*: a small ϵ discourages large-scale collusion and keeps block-space usage aligned with genuine demand. In particular, users with true value below $\mathsf{res}-\epsilon$ will not find collusion worthwhile, and those slightly below res may occasionally collude and be confirmed in harmless, Pareto-improving ways as described above. Consequently, observed demand approximates the number of users with value at least $\mathsf{res} - \epsilon$. Thus, ϵ acts as a tunable parameter that permits harmless deviations without distorting the overall fee signal.

1.3 Related Work

Transaction fee mechanisms. Since the foundational works of Roughgarden [Rou21] and Chung and Shi [CS23], the design of incentive-compatible TFMs has drawn significant attention [SCW23, WSC24, BGR24, GY23, GY24, CRS24, GTW24, CSLZZ25]. For finite block size, [CS23] showed that UIC and 1-SCP are incompatible in the plain model. Subsequently, [SCW23] introduced the MPC-assisted model and proved that UIC and 2-SCP remain incompatible even assuming cryptography, and [CRS24] showed that no TFM can satisfy UIC, MIC, and *off-chain agreement proofness* (OCA-proofness) in the plain model, where OCA-proofness is a collusion-resistant notion which requires that a global coalition of the miner and all users cannot extract profit from the protocol by deviating. Collectively, these results reveal a fundamental tension between IC and collusion-resilience in TFMs. Our work adds a complementary angle by focusing on preventing costless harm. We focus on CRHP as a coalition-resistant constraint and show that UIC, MIC, and CRHP can *all* be achieved simultaneously in both the plain and the MPC-assisted models, although feasibility in the plain model must come at the cost of zero revenue.

Non-bossiness and robustness to secondary goals. Our RHP notion is conceptually related to the classical idea of *non-bossiness* [SS81, DPTM24], which prevents an agent from changing others' outcomes without changing their own (see Thomson's survey [Tho16] for variants and applications). By contrast, RHP operates at the *utility* level: it rules out any deviation that makes an honest participant worse off unless the deviators also hurt themselves. Non-bossiness is distinct from incentive compatibility: an allocation can be strategy-proof yet bossy, and vice versa. Similarly, we showed in Theorem 1.11 that IC and RHP are fundamentally incomparable properties.

There is a rich literature on making mechanisms robust to *secondary goals* beyond direct payoff [Car19, JM05, LRS21, LSZ23]. For example, [JM05] studied auctions with externalities where bidders' utilities depend on others' allocations or information, and recent work formalizes robustness to hidden secondary goals via non-bossy mechanisms [LSZ23]. We refer the readers to Carroll's survey of robust mechanism design [Car19]. These works illustrate that even canonical strategy-proof mechanisms may fail when agents care about externalities. Our RHP property can be viewed as robustness guarantee against costless harm by adversarial deviators. Notably, our RHP definition explicitly accounts for miner deviations and miner–user collusion, not just single user strategies.

Miner deviations and credible auctions. Akbarpour and Li [AL20] introduce *credible auctions*, requiring that the auctioneer (miner) has no incentive to *safely* deviate from the prescribed mechanism, where a deviation is *safe* if it admits a plausible explanation to all users. Credibility is related to, but distinct from, MIC: as shown in [CS23], a TFM can satisfy MIC yet fail credibility. Moreover, [AL20] proves a trilemma: a revenue-optimal auction cannot be simultaneously *credible*, *strategy-proof*, and *static*. By comparison, our impossibility results target a different scope: we do not impose revenue optimality as an objective. Subsequent work [FW20, EFW22] demonstrates that cryptographic techniques can circumvent the trilemma impossibility, but these constructions do not ensure harm-prevention as provided by RHP or explicitly address the miner-user coalition in TFMs.

2 Technical Overview

This section highlights the core ideas behind our impossibility results. The main technical challenge in analyzing RHP is that we must track not only how a deviation changes the deviator's own utility, but also its effect on every honest participant's utility. This is different from the analysis of incentive compatibility, which considers only the deviator's payoff. We outline below how our key impossibility theorems are proved. The complementary positive results are established by verifying properties of the candidate mechanisms and are presented in the technical sections.

Notation. For a bid vector $\mathbf{b} = (b_1, ..., b_n)$, we use $\max(\mathbf{b})$ to denote $\max_{i\in[n]}(b_i)$. We use $\mathbf{b}_{-i}$ to denote $(b_1, \ldots, b_{i-1}, b_{i+1} \ldots, b_n)$, and we use $(\mathbf{b}_{-i}, b_i)$ and $\mathbf{b}$ interchangeably. For any $i \in [n]$, we use $x_i(\mathbf{b})$, $p_i(\mathbf{b})$, and $\mathsf{util}_i(\mathbf{b})$ to denote the confirmation probability, expected payment, and expected utility, respectively, of user i's true value b_i, when everyone behaves honestly with an input bid vector $\mathbf{b}$. We use $\mu(\mathbf{b})$ to denote the expected miner revenue in the honest case.

We focus on bids with *unique* values since we only require the mechanisms to be weakly symmetric for the impossibility. We call a bid b_i *unique in* $\mathbf{b}$ if $b_j \neq b_i$ for all $j \neq i$. For a unique bid b_i, we will sometimes abuse notation and refer to its value b_i directly: We use $\widetilde{x}_{b_i}(\mathbf{b})$, $\widetilde{p}_{b_i}(\mathbf{b})$, and $\widetilde{\mathsf{util}}_{b_i}(\mathbf{b})$ to denote the confirmation probability, expected payment, and expected honest utility of the unique bid with value b_i in $\mathbf{b}$, respectively, in the honest execution. For two sets S and T, we use $S \prec T$ to show that $\sup S < \inf T$.

2.1 Impossibility of UIC+MIC+CRHP in the Plain Model

We start by sketching the proof of Theorem 1.2, which states that any TFM satisfying UIC, MIC, and 1-CRHP in the plain model cannot confirm any bids when the input bid vector $|\mathbf{b}| > k$. We first explain the proof for deterministic TFMs to illustrate the main idea, and then discuss how to generalize it to randomized TFMs.

Impossibility for Deterministic TFM. The proof uses an induction on a carefully constructed sequence of bid vectors. We use a simple *toy example* to illustrate the key idea behind this sequence.

Toy Example: Suppose a TFM satisfies 1-CRHP. Consider two users with true values $\mathbf{b} = (b_1, b_2)$. Assume b_1 and b_2 are unique in $\mathbf{b}$, and in the honest execution user 2's bid is confirmed ($x_2(\mathbf{b}) = 1$) and pays $p_2(\mathbf{b}) < b_1$. So user 2 gets a positive utility. Under these conditions, user 1 must also be confirmed in the honest execution. If not, consider a coalition of the miner and user 1. In the honest outcome, their joint utility is just the miner's revenue $\mu(\mathbf{b})$. Now the coalition performs this **censor-then-replace** attack: the miner ignores user 2's bid b_2, and user 1 changes its own bid to b_2 while also injecting a fake bid of value b_1. The resulting bid vector is still $\mathbf{b}$ but now both b_1 and b_2 come from user 1. By weak symmetry, the bid of value b_2 (now coming from user 1) would still be confirmed and charged $p_2(\mathbf{b})$, while the bid of value b_1 is not confirmed. Under this strategy, the coalition's joint utility strictly increases: the miner's revenue is unchanged, and user 1 now gets utility $b_1 - p_2(\mathbf{b}) > 0$, while the honest user 2 is harmed. This contradicts 1-CRHP. Thus b_1 must be confirmed in the honest execution.

The above toy example can be generalized to a useful "above-payment-guarantee" lemma: *In an honest execution of a (weak) 1-CRHP TFM, if some unique bid b_i earns positive utility, then* any *unique bid in b that is higher than $p_i(\mathbf{b})$ must also be confirmed.* Intuitively, if a higher bid is not confirmed, then it can collude with the miner and perform the censor-then-replace attack on b_i, which contradicts 1-CRHP. Here, we require uniqueness of the bid to get rid of the potential tie breaking that can depend on the metadata.

Induction. Next, we show how to prove Theorem 1.2 for deterministic TFM. Assume for sake of contradiction that there exists a deterministic TFM satisfying UIC, MIC, and 1-CRHP in which some bid v_{i^*} is confirmed when the input bid vector is $\mathbf{v}$ with $|\mathbf{v}| > k$. We will construct a sequence of bid vectors $\mathbf{v}^{(m)}$ for $m = 1, 2, ...$, and prove by induction that m bids must be confirmed in $\mathbf{v}^{(m)}$. This directly contradict the block size k when $m = k + 1$. The sequence is constructed as follows:

$$\begin{aligned} b_1 &= \max(\mathbf{v}) + 1, & \mathbf{v}^{(1)} &= (\mathbf{v}_{-i^*}, b_1), \\ b_m &= \frac{1}{2}(\max(\mathbf{v}) + b_{m-1}), & \mathbf{v}^{(m)} &= \left(\mathbf{v}^{(m-1)}, b_m\right) \quad \text{for } m > 1. \end{aligned}$$

The inductive invariant states that

$$b_m \text{ must be confirmed in } \mathbf{v}^{(m)} \text{ and its payment } \widetilde{p}_{b_m}(\mathbf{v}^{(m)}) \le \max(\mathbf{v}).$$

Effectively, in each step, we add a bid b_m that lies between the payment $\widetilde{p}_{b_{m-1}}\left(\mathbf{v}^{(m-1)}\right) \le \max(\mathbf{v})$ and b_{m-1} to $\mathbf{v}^{(m-1)}$. By the inductive statement, b_m gets a positive utility in $\mathbf{v}^{(m)}$, and all b_z for $z \in [m]$ is higher than b_m's payment. Thus, by the above-payment-guarantee lemma, all b_z for $z \in [m]$ must be confirmed.

The induction proof uses Myerson's lemma (Lemma 3.5). Roughly speaking, for a deterministic UIC TFM, for any fixed other users' bids $\mathbf{b}_{-i}$, there exists a critical value v^* such that any $b_i > v^*$ gets confirmed and pays v^*, i.e., $x_i(b_i, \mathbf{b}_{-i}) = 1$ and $p_i(b_i, \mathbf{b}_{-i}) = v^*$ for any $b_i > v^*$.

Base case: The base case follows directly from UIC and the Myerson's lemma: Raising user i^*'s bid in $\mathbf{v}$ to $\max(\mathbf{v}) + 1$ must maintain its confirmation probability and payment. Thus, $\widetilde{x}_{b_1}(\mathbf{v}^{(1)}) = 1$ and $\widetilde{p}_{b_1}(\mathbf{v}^{(1)}) = p_{i^*}(\mathbf{v}) \le \max(\mathbf{v})$ where the inequality follows from individual rationality.

Inductive step: Now assume we have $\mathbf{v}^{(m-1)}$ where b_{m-1} is confirmed and pays no more than $\max(\mathbf{v})$. By Myerson's Lemma, it suffices to show that for any b such that $\max(\mathbf{v}) < b < b_{m-1}$, bid b must be confirmed in $(\mathbf{v}^{(m-1)}, b)$. This implies that $\widetilde{x}_{b_m}(\mathbf{v}^{(m)}) = 1$ and $\widetilde{p}_{b_m}(\mathbf{v}^{(m)}) \leq \max(\mathbf{v})$.

For the sake of contradiction, assume that there exists some $\max(\mathbf{v}) < b^* < b_{m-1}$ such that b^* is not confirmed in $(\mathbf{v}^{(m-1)}, b^*)$. Imagine a world where $\mathbf{v}^{(m-1)}$ is the honest bid vector. Since $|\mathbf{v}^{(m)}| \geq n > k$, some user j must be unconfirmed in the honest case. Consider a coalition consisting of the miner and user j who injects a fake bids b^*. Since b^* is unconfirmed, the miner revenue must remain unchanged by MIC: $\mu(\mathbf{v}^{(m-1)}) = \mu(\mathbf{v}^{(m-1)}, b^*)$. Therefore, this strategy does not change the coalition's utility compared to the honest case.[3] That said, by 1-CRHP, this strategy should harm no honest bid in $\mathbf{v}^{(m-1)}$, including b_{m-1}. Therefore, b_{m-1} should still get a positive utility in $(\mathbf{v}^{(m-1)}, b^*)$ and pays no more than $\max(\mathbf{v})$. However, by the above-payment-guarantee, since b^* is greater that the payment of b_{m-1} in $(\mathbf{v}^{(m-1)}, b^*)$, we know that b^* must be confirmed, which contradicts our assumption. This completes the induction proof.

Generalizing to Randomized TFMs. The above proof idea needs significant adaptation for randomized mechanisms. If the TFM uses randomness, then having more than k bids, each with a positive *confirmation probability* is no longer an immediate contradiction. In a randomized TFM, different random outcomes might confirm different sets of at most k bids. Therefore, we derive a stronger condition called the *k-winners lemma*: If a TFM is MIC and (weak) 1-CRHP, then it must be that

$$|W(\mathbf{b})| \leq k, \text{ where } W(\mathbf{b}) := \{i \colon \mathsf{util}_i(\mathbf{b}) > 0\}. \tag{1}$$

Here, $W(\mathbf{b})$ denotes the set of *winners* who have positive expected utility in the honest case when the input bid vector is $\mathbf{b}$. Note that positive *utility* is a strictly stronger requirement than positive *confirmation probability.* We prove this lemma formally in Section 4.2. For the overview, we assume it is true and show how to prove Theorem 1.2.

For possibly randomized TFMs, the "above-payment-guarantee" can be generalized to "above-ratio-guarantee": *In a 1-CRHP TFM , if a unique bid b_i has positive expected utility in* $\mathbf{b}$*, then any unique bid in* $\mathbf{b}$ *above the ratio $p_i(\mathbf{b})/x_i(\mathbf{b})$ must have a positive confirmation probability.* In the deterministic case, this ratio $p_i(\mathbf{b})/x_i(\mathbf{b})$ is simply the payment itself. We define, for any bid b_i such that $x_i(\mathbf{b}) > 0$, the *ratio* as

$$r_i(\mathbf{b}) := p_i(\mathbf{b})/x_i(\mathbf{b}), \qquad \text{and if } b_i \text{ unique, } \widetilde{r}_{b_i}(\mathbf{b}) := \widetilde{p}_{b_i}(\mathbf{b})/\widetilde{x}_{b_i}(\mathbf{b}).$$

If a unique bid $b > \widetilde{r}_b(\mathbf{b})$, then b gets a positive utility in $\mathbf{b}$.

To prove Theorem 1.2, suppose for the sake of contradiction, that for some UIC, MIC, and 1-CRHP TFM, there exists some $\mathbf{v}$ where $|\mathbf{v}| = n > k$ and $x_{i^*}(\mathbf{v}) > 0$ for some $i^* \in [n]$. As in the deterministic case, we will construct a sequence of bid vectors $\mathbf{v}^{(m)} = (\mathbf{v}^{(m-1)}, b_m)$, each adding a new bid b_m to $\mathbf{v}^{(m-1)}$, where $\widetilde{r}_{b_{m-1}}(\mathbf{v}^{(m-1)}) < b_m < b_{m-1}$. We then prove that $W(\mathbf{v}^{(m)}) \geq m$ by induction, which leads to a contradiction of the k-winner lemma for $m > k$.

Challenges. However, simply adding one bid in each step is **not** enough. Using the above-ratio-guarantee, we can argue similarly as the deterministic case that b_m gets a positive confirmation probability in $(\mathbf{v}^{(m-1)}, b_m)$. Then raising b_m to some b'_m ends up with a positive *utility* for b'_m in $(\mathbf{v}^{(m-1)}, b'_m)$ by a corollary of Myerson's lemma.

The challenge is that raising b_m to b'_m can alter payments for others: this naive approach might make earlier winners $b_1, ..., b_{m-1}$ lose their utility, breaking the inductive invariant. Note that 1-CRHP does not rule out such strategies since raising b_m to b'_m may harm the coalition themselves

[3] This is why this proof does not work for the impossibility w.r.t. weak 1-CRHP.

in a randomized TFMs. We overcome this by carefully choosing a set S_m of two possible values $\alpha^{(m)} < \beta^{(m)}$ that the new bid b_m can take. Effectively, we build a sequence of *sets* of bid vectors $V^{(m)} = \{\mathbf{v}_{-i^*}\} \times S_1 \times \cdots \times S_m = \{(\mathbf{v}_{-i^*}, b_1, ..., b_m) : b_j \in S_j \text{ for } j \in [m]\}$. The induction will show that for each m, we have some $\mathbf{b} \in V^{(m)}$ such that $W(\mathbf{b}) \geq m$.

Base case: Pick $\max(\mathbf{v}) < \alpha^{(1)} < \beta^{(1)}$ to be two values large enough such that b_1 has a positive utility in $(\mathbf{v}_{-i^*}, b_1)$ for any $b_1 \in S_1 = \{\alpha^{(1)}, \beta^{(1)}\}$. Such values must exist since $x_{i^*}(\mathbf{v}) > 0$. Let $V^{(1)} = \{\mathbf{v}_{-i^*}\} \times S_1$. Then for any $\mathbf{b} \in V^{(1)}$, we have $W(\mathbf{b}) \geq 1$.

Inductive step: We demonstrate the main idea with $m = 2$, and summarize the general construction afterwards. Let $r^{(1)} := \max_{\mathbf{v}^{(1)} \in V^{(1)}} r_{n+1}(\mathbf{v}^{(1)}, \alpha^{(1)})$ and $\alpha^{(2)}$ and $\beta^{(2)}$ be such that $r^{(1)} < \alpha^{(2)} < \beta^{(2)} < \alpha^{(1)}$. By a similar argument as in the deterministic case, $\alpha^{(1)}$ must have a positive confirmation probability in $(\mathbf{v}^{(1)}, \alpha^{(1)})$ for any $\mathbf{v}^{(1)} \in V^{(1)}$, and therefore, $r^{(1)}$ is well-defined. The value $r^{(1)}$ is an upper bound for the ratio $\widetilde{r}_b(\mathbf{v}^{(1)}, b)$ for any bid $b < \alpha^{(1)}$ with a positive confirmation probability. This maximum guarantees that the new bid $b_2 \in S_2 = \{\alpha^{(2)}, \beta^{(2)}\}$ *always* lies between $b_1 \geq \alpha^{(1)}$ and its corresponding ratio, no matter which value b_1 takes on. Let $V^{(2)} = \{\mathbf{v}_{-i^*}\} \times S_1 \times S_2$. We will show that

$$\text{for any } \mathbf{v}^{(2)} = (\mathbf{v}_{-i^*}, b_1, b_2) \in V^{(2)},\ \widetilde{\mathsf{util}}_{b_2}(\mathbf{v}^{(2)}) > 0 \text{ for any } b_2 \in S_2; \tag{2}$$

$$\text{for any } i \in [2],\ \widetilde{\mathsf{util}}_{\beta^{(i)}}\left(\mathbf{v}_{-i^*}, \beta^{(1)}, \beta^{(2)}\right) > 0, \tag{3}$$

which completes the induction proof for $m = 2$ since $|W(\mathbf{v}_{-i^*}, \beta^{(1)}, \beta^{(2)})| \geq 2$.

Property (2) directly follows from UIC, since any bid $b > r^{(1)}$ must have a positive utility in $(\mathbf{v}^{(1)}, b)$ for any $\mathbf{v}^{(1)} \in V^{(1)}$. Otherwise, when the honest bid vector is $(\mathbf{v}^{(1)}, b)$, the user with true value b is incentivized to overbid to $\alpha^{(1)}$ to gain a positive utility.

To see why (3) is true, consider $\mathbf{b}' = (\mathbf{v}_{-i^*}, \alpha^{(1)}, \beta^{(2)})$. Since $\beta^{(2)}$ has a positive utility in $\mathbf{b}'$ by (2), and $\alpha^{(1)} > \beta^{(2)} > r^{(1)} \geq \widetilde{r}_{\beta^{(2)}}(\mathbf{b}')$, we know $\alpha^{(1)}$ must have a positive confirmation probability in $\mathbf{b}'$by the above-ratio-guarantee. Therefore, raising $\alpha^{(1)}$ to $\beta^{(1)}$ gives $\beta^{(1)}$ a positive utility. This completes the proof for $m = 2$.

In general, at each step m, we define $r^{(m-1)} = \max_{\mathbf{v}^{(m-1)} \in V^{(m-1)}} r_{n+m-1}\left(\mathbf{v}^{(m-1)}, \alpha^{(m-1)}\right)$ as an upper bound, over all possible $\mathbf{v}^{(m-1)} \in V^{(m-1)}$, of the ratio that a new bid smaller than $\alpha^{(m-1)}$ could have if it gets a positive confirmation probability. We then pick $r^{(m-1)} < \alpha^{(m)} < \beta^{(m)} < \alpha^{(m-1)}$ and define $S_m = \{\alpha^{(m)}, \beta^{(m)}\}$ and $V^{(m)} = V^{(m-1)} \times S_m$. Roughly speaking, in the m-th step, for any $(b_1, ..., b_m) \in S_1 \times \cdots \times S_m$, the new bid b_m must get a positive utility in $(\mathbf{v}_{-i^*}, b_1, ..., b_m)$. Now consider specifically $\mathbf{b}' = (\mathbf{v}_{-i^*}, \beta^{(1)}, ..., \beta^{(m)})$. Using the above-ratio-guarantee and our choice of the sets S_j, we can conclude that for any $n \leq j \leq n + m - 1$, we have $x_j(\mathbf{b}'_{-j}, \alpha^{(j-n+1)}) > 0$. Therefore, $\mathsf{util}_j(\mathbf{b}') > 0$ for each $n \leq j \leq n + m - 1$. When $m = k + 1$, this directly contradicts the k-winners lemma. This completes the impossibility proof for randomized TFMs in the plain model. We omit the details of the inductive reasoning here. See Section 4.2 for the complete proof.

2.2 Impossibility under Weak-CRHP in the Plain Model

In this section, we outline the proof of Theorem 1.3, which shows that even if the TFM satisfies only weak 1-CRHP, no positive miner revenue is possible alongside UIC and MIC in the plain model. Recall that weak 1-CRHP means that strictly profitable deviations must not harm honest players, but it does not rule out utility-neutral deviations that are harmful to others.

This relaxation complicates the proof as mentioned in Footnote 3. In the 1-CRHP setting, each step of the induction used a contradiction argument to show that a newly added bid b_m must receive a positive confirmation probability: if b_m were not confirmed, a miner–user coalition could instead inject b_m without changing joint utility, and under 1-CRHP such a deviation cannot harm any honest player. Under only weak 1-CRHP, however, this utility-neutral strategy can harm honest players, so the earlier argument breaks down.

To overcome this, we adopt a slightly different approach. The key observation is that by MIC, if we start with some bid vector $\mathbf{v}$ with a positive miner revenue, then adding any set of additional bids $\mathbf{b}$ will still produce $\mu(\mathbf{v}, \mathbf{b}) > 0$. By budget feasibility, this implies that some bid must have a positive confirmation probability in $(\mathbf{v}, \mathbf{b})$. We use the following lemma to argue that this confirmed bid must be one of the newly added bids.

Lemma 2.1 (no-persistent-old-winner, informal)**.** *Suppose $\mathbf{v} = (v_1, \ldots, v_n)$ and a sequence of sets $S_1, ..., S_k \subset \mathbb{R}_{\geq 0}$ such that $|S_i| \geq 2$, and that $\{\max(\mathbf{v})\} \prec S_1 \prec \cdots \prec S_k$. Then there is no $i \in [n]$ such that $x_i(\mathbf{v}, \mathbf{b}) > 0$ for all $\mathbf{b} = (b_1, \ldots, b_m) \in S_1 \times \cdots \times S_m$.*

Suppose, for the sake of contradiction, that there exists a UIC, MIC, weak 1-CRHP TFM and an initial bid vector $\mathbf{v}$ with $\mu(\mathbf{v}) > 0$. We inductively construct a sequence of bid-vector sets $V^{(m)}$ by adding k new sets, instead of one set, at each step, i.e. $V^{(m)} = V^{(m-1)} \times A_1^{(m)} \times \cdots \times A_k^{(m)}$. We show by induction that for each m there exists some $\mathbf{b} \in V^{(m)}$, such that $\mu(\mathbf{b}) > 0$ and $W(\mathbf{b}) \geq m$. For $m = k + 1$ this yields a contradiction to the k-winner lemma (which remains valid under weak 1-CRHP). In the induction proof, we will make use of the *Graham–Rothschild theorem*:

Lemma 2.2 ([GRS90])**.** *For all $p, \xi, a > 0$, there exists N such that, if $T \subseteq A_1 \times \cdots \times A_p$, each $|A_i| = N$ for $i \in [p]$, and $|T| \geq \xi \cdot N^p$, then there exist subsets $B_i \subseteq A_i$ with $|B_i| \geq a$ such that*

$$B_1 \times \cdots \times B_p \subseteq G.$$

Looking ahead, in the induction, each newly added set $A_j^{(m)}$ in the m-th step has size N_m. The sequence of N_m is defined backwards starting from $N_{k+2} = 2$, such that at the end of the m-th step, there exist m sets $S_1^{(m)}, ..., S_m^{(m)}$, each with size N_{m+1}, such that any b_i in $\mathbf{b} = (b_1, ..., b_m) \in S_1^{(m)} \times ... \times S_m^{(m)}$ must have a positive confirmation probability in $(\mathbf{v}^{(m)}, \mathbf{b})$. Therefore, b_i must have a positive utility in all possible $(\mathbf{v}^{(m)}, \mathbf{b})$ as long as $b_i > \min S_i^{(m)}$. We show how to choose N_m given N_{m+1} in the base case and inductive step below.

<u>*Base case:*</u> Let $\{\max(\mathbf{v})\} \prec A_1^{(1)} \prec ... \prec A_k^{(1)}$ be k sets, each of size N_1, where N_1 is the N in Lemma 2.2 with $p = k$, $\xi = \frac{1}{n+k}$, and $a = N_2$. By MIC, $\mu(\mathbf{v}, \mathbf{b}) > 0$ for any $\mathbf{b} \in A_1^{(1)} \times \cdots \times A_k^{(1)}$. Therefore, some bid must have a positive confirmation probability in each $(\mathbf{v}, \mathbf{b})$. Based on the smallest index of the bid being confirmed, we get a natural partition of $A_1^{(1)} \times \cdots \times A_k^{(1)}$. Let

$$T_i = \{\mathbf{b} \in A_1^{(1)} \times \cdots \times A_k^{(1)} : \text{ the smallest index } i \text{ such that } x_i(\mathbf{v}, \mathbf{b}) > 0 \text{ is } i\} \text{ for } i \in [n + k].$$

At least one of these sets $|T_{i^*}| \geq \frac{1}{n+k} \cdot N_1^k$ for some $i^* \in [n + k]$. By Lemma 2.2, there exists $B_j^{(1)} \subseteq A_j^{(1)}$, each $|B_j^{(1)}| \geq N_2$, such that $B_1^{(1)} \times \cdots \times B_k^{(1)} \subseteq T_{i^*}$, i.e., the i^*-th bid is confirmed in $(\mathbf{v}, \mathbf{b})$ for any $\mathbf{b} \in B_1^{(1)} \times \cdots \times B_k^{(1)}$. By Lemma 2.1, i^* cannot be one of the old bids from $\mathbf{v}$, i.e., $i^* > n$. Let $\ell = i^* - n$.

Pick an arbitrary $\mathbf{b}^* \in B_1^{(1)} \times \cdots \times B_k^{(1)} \subseteq T_{i^*}$, and let $\mathbf{v}^{(1)} = (\mathbf{v}, \mathbf{b}^*_{-\ell})$. Let $S_1^{(1)}$ be a set with N_2 values such that $A_k^{(1)} \prec S_1^{(1)}$, and the ratio $r^{(1)}$ (defined analogously as in Section 2.1) $r^{(1)} = \widetilde{r}_{\max S_1^{(1)}}(\mathbf{v}^{(1)}, \max S_1^{(1)})$. By Myerson's lemma, $\widetilde{x}_{b_1}(\mathbf{v}^{(1)}, b_1) > 0$ for any $b_1 \in S_1^{(1)}$.

<u>*Inductive step:*</u> Pick k sets $A_1^{(2)}, ..., A_k^{(2)}$ such that $\{r^{(1)}, \max(\mathbf{v}^{(1)})\} \prec A_1^{(2)} \prec \cdots \prec A_k^{(2)} \prec S_1^{(1)}$ and each $|A_j^{(2)}| = N_2$. This sequence of sets exist because $\{r^{(1)}, \max(\mathbf{v}^{(1)})\} \prec S_1^{(1)}$ by construction. Here,

N_2 is the N in Lemma 2.2 with $p = k+1$, $\xi = \frac{1}{n+2k}$, and $a = N_3$. Still, by MIC, $\mu(\mathbf{v}^{(1)}, \mathbf{b}, b_1) > 0$ for any $(\mathbf{b}, b_1) \in A_1^{(2)} \times \cdots \times A_k^{(2)} \times S_1^{(1)}$, meaning that some bid must have a positive confirmation probability. Consider the following natural partition of $A_1^{(2)} \times \cdots \times A_k^{(2)} \times S_1^{(1)}$ for $i \in [n+2k]$:

$$T_i = \left\{ (\mathbf{b}, b_1) \in A_1^{(2)} \times \cdots \times A_k^{(2)} \times S_1^{(1)} : \text{ the smallest index } i \text{ such that } x_i(\mathbf{v}, \mathbf{b}, b_1) > 0 \text{ is } i \right\}.$$

By a similar reasoning as in the base case, there must exist an $i^* \in [n+2k]$, such that there exists $B_j^{(2)} \subseteq A_j^{(2)}$ and $S_1^{(2)} \subseteq S_1^{(1)}$, each of size N_3, such that $B_1^{(2)} \times \cdots \times B_k^{(2)} \times S_1^{(2)} \subseteq T_{i^*}$, i.e., the i^*'th bid has a positive confirmation probability in $(\mathbf{v}^{(1)}, \mathbf{b}, b_1)$ for all $(\mathbf{b}, b_1) \in B_1^{(2)} \times \cdots \times B_k^{(2)} \times S_1^{(2)}$. By Lemma 2.1, i^* cannot be smaller than $n+k$. By a careful analysis based on the censor-then-replace attack that censors all bids not in $\mathbf{v}^{(1)}$ and replaces the b_1 bid, we show that $i^* < n+2k$. That is, i^* must represent one of the new bids from $B_1^{(2)} \times \cdots \times B_k^{(2)}$. Let $\ell = i^* - |\mathbf{v}^{(1)}|$.

Pick an arbitrary $(\mathbf{b}^*, b_1) \in B_1^{(2)} \times \cdots \times B_k^{(2)} \times S_1^{(2)} \subseteq T_{i^*}$, and let $\mathbf{v}^{(2)} = (\mathbf{v}^{(1)}, \mathbf{b}^*_{-\ell})$. Let $S_2^{(2)}$ be a set with N_3 values such that $A_k^{(2)} \prec S_2^{(2)} \prec S_1^{(2)}$, and the ratio $r^{(2)}$ is defined analogously as the base case. Then b_1 and b_2 must have a positive confirmation probability in $(\mathbf{v}^{(2)}, b_1, b_2)$ for any $(b_1, b_2) \in S_1^{(2)} \times S_2^{(2)}$.

In general, for all $m \geq 2$, the inductive step proceeds analogously. Define N_i backward as follows: set $N_{k+2} := 2$, and for each $i = k+1, k, \ldots, 1$, let N_i be the number N given by Lemma 2.2 with parameters $p = k+i-1$, $\xi = \frac{1}{n+ki}$, and $a = N_{i+1}$. This choice ensures that at the end of the $(k+1)$-th step we obtain sets $S_1^{(k+1)}, \ldots, S_{k+1}^{(k+1)}$, each of size two, such that for every $\mathbf{b} = (b_1, \ldots, b_{k+1}) \in S_1^{(k+1)} \times \cdots \times S_{k+1}^{(k+1)}$, each coordinate b_i has a positive confirmation probability in $(\mathbf{v}^{(k+1)}, \mathbf{b})$. Consequently, taking $b_i^* := \max S_i^{(k+1)}$ for all $i \in [k+1]$ yields a bid vector $(\mathbf{v}^{(k+1)}, b_1^*, \ldots, b_{k+1}^*)$ in which all b_i^* enjoy positive utility, by a similar reasoning in Section 2.1, which contradicts the k-winner lemma.

Due to the heavy technical nature of this construction, we refer the reader to Section 4.3 for the complete proof.

2.3 Impossibility of UIC+URHP for Deterministic MPC-assisted TFMs

The proof of Theorem 1.10 is very similar to the deterministic impossibility in the plain model. However, we cannot directly use the above-payment-guarantee lemma because it utilizes a censor-then-replace attack, which cannot be performed in the MPC-assisted model. Instead we prove the above-payment-guarantee for deterministic UIC and URHP mechanisms in the MPC-assisted model. The crux of the proof is to start with a slightly weaker claim: *for a deterministic TFM that satisfies UIC and URHP, if b_i gets a positive utility in $\mathbf{b}$, then any unique bid higher than b_i must also be confirmed.* Then we only need to show that any unique b_j between $p_i(\mathbf{b})$ and b_i must be confirmed to get the above-payment-guarantee lemma.

Suppose for the sake of contradiction that in some UIC and URHP TFM in the MPC-assisted model, some b_j between $p_i(\mathbf{b})$ and b_i was not confirmed. Consider some b_i' such that $p_i(\mathbf{b}) < b_i' < b_j < b_i$. Using Myerson's lemma and the above weaker claim, we can show that both b_i' and b_j get a positive utility in $(\mathbf{b}_{-i}, b_i')$. That said, if the honest bid vector was $(\mathbf{b}_{-i}, b_i')$, user i can harm user j by raising its bid to b_i. This strategy keeps user i's utility unchanged compared to the honest case but strictly harm user j, which contradicts URHP. This proves the above-payment-guarantee for UIC and URHP TFMs in the MPC-assisted model. The rest of the proof follows by a similar reasoning of the deterministic case in Section 2.1.

3 Model and Preliminary

3.1 Notation

Let $\mathbb{N}$ denote the set of natural numbers and $\mathbb{R}_{\geq 0}$ denote the non-negative real numbers. Given a bid vector $\mathbf{b} = (b_1, \ldots, b_n)$ and any $i \in [n]$, we use $\mathbf{b}_{-i}$ to denote $(b_1, \ldots, b_{i-1}, b_{i+1}, \ldots, b_n)$. We use $(\mathbf{b}_{-i}, b'_i)$ to denote the vector obtained by replacing b_i with b'_i in $\mathbf{b}$, and treat this as equivalent to $(b_1, \ldots, b_{i-1}, b'_i, b_{i+1}, \ldots, b_n)$. For vectors $\mathbf{a}$ and $\mathbf{b}$, we write $\mathbf{a} \subseteq \mathbf{b}$ if every entry of $\mathbf{a}$ appears in $\mathbf{b}$. We use $|\mathbf{b}|$ to denote the number of elements in vector $\mathbf{b}$. For any vector $\mathbf{v} = (v_1, v_2, \ldots, v_n)$, we define $\max(\mathbf{v}) := \max_{i \in [n]} v_i$. For a vector $\mathbf{v}$ and $S \subseteq [n]$, let $\mathbf{v}_S := (v_i)_{i \in S}$. For two sets S and T, we use $S \prec T$ to show that $\sup S < \inf T$.

3.2 Transaction Fee Mechanism

We consider a transaction fee mechanism (TFM) where each block has a finite capacity of k transactions. Assume each user has a true value of $v_i \in \mathbb{R}_{\geq 0}$ that measures the maximum amount a user is willing to pay to get its transaction confirmed in the block. Each user i submits their transaction together with a *bid* b_i. We assume each transaction takes one slot in the block. In this paper, "bid" and "transaction" are used interchangeably. A TFM with capacity k is defined as a tuple $(\mathbf{I}, \mathbf{C}, \mathbf{P}, \mathbf{R})$ where:

- **Inclusion rule** $\mathbf{I}(\cdot)$: takes a bid vector $\mathbf{b}$ as input, and outputs a *block* $B \subseteq \mathbf{b}$ of at most k bids to include.
- **Confirmation rule** $\mathbf{C}(\cdot)$: takes as input a block B of the included bids and chooses a subset of included bids to confirm. Specifically, $\mathbf{C}(B)$ outputs a vector $(x_1, ..., x_{|B|}) \in \{0,1\}^{|B|}$, indicating whether each bid is confirmed.
- **Payment rule** $\mathbf{P}(\cdot)$: takes a block B as input and outputs a vector of $(p_1, ..., p_{|B|}) \in \mathbb{R}_{\geq 0}^{|B|}$, indicating the price paid by each transaction in B.
- **Revenue rule** $\mathbf{R}(\cdot)$: takes a block B as input, and outputs the miner's revenue $\mu \in \mathbb{R}_{\geq 0}$.

A feasible TFM $(\mathbf{I}, \mathbf{C}, \mathbf{P}, \mathbf{R})$ should satisfy the following properties:

- *Space feasibility:* the size of the block $|B| \leq k$.
- *Individual rationality:* a user's payment shall not exceed the bid amount, i.e., for any $b_i \in B$, its payment $p_i \leq b_i$.
- *budget feasibility*: the miner's revenue cannot exceed the total payment collected from all confirmed bids, i.e., $\mu \leq \sum_{i=1}^{|B|} p_i$. When the miner's revenue is strictly less than the total payments, we say that the difference is *burnt*.

Among the four rules, the confirmation rule, payment rule, and revenue rule are all executed by the blockchain based on the block, and we treat these three rules as always correctly implemented based on the input block B. The inclusion rule is either implemented by the miner in the plain model or implemented by an MPC in the MPC-assisted model.

Weak Symmetry. Given a bid $\mathbf{b} = (b_1, ..., b_n)$, let x_i and p_i denote the random variable representing the probability of bid b_i getting confirmed and the payment it needs to pay in an honest execution. We say that a TFM satisfies *weak symmetry* iff the distribution of the set $\{(b_i, x_i, p_i)\}_{i \in [n]}$ is the same for input bid vector $\pi(\mathbf{b})$ for any permutation π on $\mathbf{b}$.

An operational view of weak symmetry assumes the following: Given a bid vector $\mathbf{b}$ where each bid may carry some metadata such as identity, public keys of the user, or timestamp. The honest mechanism first sorts the bids based on the amount, and can perform arbitrary tie-breaking rule based on the metadata if multiple bids have the same amount. After the sorting step, all the four rules $(\mathbf{I}, \mathbf{C}, \mathbf{P}, \mathbf{R})$ depend only on bid amounts and positions in the sorted vector. Weak symmetry is a natural requirement in practice and does not require two bids of the same amount to always receive the same treatment.

3.3 Game Induced by TFM

Henceforth, let $\mathcal{C}$ denotes the strategic players. Specifically, $\mathcal{C}$ can be a strategic user, strategic miner(s) in control of the current block, or a coalition of the miner(s) and one or more users.

3.3.1 TFM Game in the Plain Model.

In the plain model, the TFM game is defined as follows:

1. Each honest users not in $\mathcal{C}$ submits their bid represented by a single real-value bid. Let $\mathbf{b}_{-\mathcal{C}}$ denote the bids from these honest users.
2. The coalition $\mathcal{C}$ decides their bids $\mathbf{b}_{\mathcal{C}}$ according to $\mathbf{b}_{-\mathcal{C}}$.
3. The miner selects up to k bids from $(\mathbf{b}_{\mathcal{C}}, \mathbf{b}_{-\mathcal{C}})$ to form a block B.
4. The blockchain protocol executes the confirmation rule $\mathbf{C}$, payment rule $\mathbf{P}$, and miner revenue rule $\mathbf{R}$ to the block B created by the miner.

Strategy Space. In this paper, we focus on direct-revelation mechanism, i.e., for an user i with true value v_i, the honest strategy $H_i(v_i) = v_i$ is to submit a single bid representing its true value. We focus on *ex-post* incentives: strategic parties may condition their actions based on the honest users' bids, i.e., we do not hide honest bids from the strategic players. A strategic user may choose to submit a bid vector $\mathbf{b}_i^*$ that contains zero to multiple bids which do not necessarily reflect their true value. We call all the additional bids that the user injects as *fake* bids.

A miner's honest behavior $H_{\mathcal{M}}$ is to implement the prescribed inclusion rule on the input bid vector $\mathbf{b}$ without submitting any fake bids, i.e., if users' honest bid vector is $\mathbf{b}$, honest miner strategy $H_{\mathcal{M}}(\mathbf{b})$ outputs a block $\mathbf{I}(\mathbf{b})$. In the plain model, a strategic miner may choose to deviate from the prescribed inclusion rule arbitrarily by dropping bids, injecting fake bids, and arbitrarily choosing the randomness used in the inclusion rule, i.e., the strategy $S_{\mathcal{M}}(\mathbf{b})$ outputs an arbitrary block B^* of size at most k. A strategic coalition can adopt a combination of the strategies of its members.

3.3.2 MPC-Assisted Model

In the MPC-assisted model, M miners jointly run a multi-party computation (MPC) to realize an ideal functionality that honestly implements the inclusion rule, and the M miners share the total revenue. Concretely, there is an ideal functionality $\mathcal{F}_{\text{TFM}}$ that, on input $\mathbf{b}$, implements the honest inclusion, confirmation, payment and revenue rule on $\mathbf{b}$. who always performs correctly. In this paper, we analyze incentives in the "ideal" world with $\mathcal{F}_{\text{TFM}}$. Following the universal composibility paradigm in the cryptography literature, all the guarantees hold with an overwhelming probability when instantiating with real-world cryptography [GMW87, Can01].

In the MPC-assisted model, the game runs as follows:

1. Users not in $\mathcal{C}$ submits a bid represented by a single real value to $\mathcal{F}_{\text{TFM}}$. Let $\mathbf{b}_{-\mathcal{C}}$ denote the bids from users outside $\mathcal{C}$.
2. $\mathcal{F}_{\text{TFM}}$ sends $\mathbf{b}_{-\mathcal{C}}$ to $\mathcal{C}$ and receives a bid vector $\mathbf{b}_{\mathcal{C}}$ from the coalition.
3. $\mathcal{F}_{\text{TFM}}$ implements the inclusion, confirmation, payment, and revenue rule on $(\mathbf{b}_{-\mathcal{C}}, \mathbf{b}_{\mathcal{C}})$. The outputs include a vector indicating whether each bid is confirmed or not, a payment vector of every bid's payment, and the total miner revenue.
4. If the number of miners in $\mathcal{C}$ is at least $M/2$, sends the outcome to $\mathcal{C}$ and receives a response. If $\mathcal{C}$ responded fail, send fail to everyone, meaning that the auction aborts. No one gets confirmed and miners get no revenue. Otherwise if $\mathcal{C}$ responses ok, send the outcome to every player.
5. If the number of miners in $\mathcal{C}$ is less than $M/2$, send the outcome to every player.

Strategy Space. User's strategy space is the same as in the plain model. An honest miner does not submit any bids, whereas strategic miner(s) may inject one or more fake bids. If there are at least $M/2$ miners in the strategic coalition, they can also choose to fail the auction. Unlike in the plain model, now strategic miner(s) can no longer drop honest users' bids or arbitrarily choose which subset of bids to include since now the inclusion rule is implemented by the ideal functionality $\mathcal{F}_{\text{TFM}}$. A strategic coalition can adopt a combination of the strategies of its members.

3.4 Utility and Incentive Compatibility

Utility. Each user i has a true value $v_i \in \mathbb{R}_{\geq 0}$ if its primary bid representing its transaction gets confirmed. All the fake bids have true value 0. Let p_i denote the total payment user i needs to pay. Then user i's utility is $v_i - p_i$ if its primary bid gets confirmed and $-p_i$ otherwise. A miner's utility is its revenue $\mu - p_{\mathcal{M}}$, where $p_{\mathcal{M}}$ denotes the total payment from the miner if they inject any fake bids. A coalition's joint utility is the sum of all coalition members' utilities.

Below, for a strategic player or coalition, denoted as $\mathcal{C}$, we use $H_{\mathcal{C}}$ to denote $\mathcal{C}$'s honest strategy of $\mathcal{C}$. Let $\mathbf{v}$ represent the vector of true values of all users. We use $\mathsf{util}_i(\mathbf{v}; S_{\mathcal{C}})$ to denote the expected utility of user i in the following randomized experiment:

- In the mechanism, players in $\mathcal{C}$ adopts strategy S, while all other players act honestly, where all users' true values are represented by $\mathbf{v}$.
- Output utility of player i.

A miner j's expected utility $\mathsf{mutil}_j(\mathbf{v}; S_{\mathcal{C}})$ and coalition's joint utility $\mathsf{util}_{\mathcal{C}}(\mathbf{v}; S_{\mathcal{C}})$ are defined analogously w.r.t the above randomized experiment.

Definition 3.1 (Incentive compatibility)**.** Given a TFM $(\mathbf{I}, \mathbf{C}, \mathbf{P}, \mathbf{R})$, we say that the TFM satisfies incentive compatibility (IC) w.r.t. a strategic player or coalition, denoted as $\mathcal{C}$, iff for any true value vector $\mathbf{v}$ of users, for any strategy $S_{\mathcal{C}}$ of coalition $\mathcal{C}$, we have

$$\mathsf{util}_{\mathcal{C}}(\mathbf{v}; S_{\mathcal{C}}) \leq \mathsf{util}_{\mathcal{C}}(\mathbf{v}; H_{\mathcal{C}}).$$

Specifically, we say that a TFM satisfies

- *User incentive compatibility (UIC)* if the above holds when $\mathcal{C}$ contains an individual user, and the miner acts honestly.

- *Miner incentive compatibility (MIC)* in the plain model if the above holds when $\mathcal{C}$ only contains the miner.

  In the MPC-assisted model, we say that a TFM satisfies MIC if the above holds when $\mathcal{C}$ contains up to $M-1$ miners jointly run the MPC implementing $\mathcal{F}_{\text{TFM}}$.

- *d-side-contract-proofness (d-SCP)* in the plain model for some integer $d \geq 1$ if the above holds when $\mathcal{C}$ contains the miner and at least one but no more than d number of users.

  In the MPC-assisted model, we say that a TFM satisfies d-SCP for some integer $d \geq 1$ if the above holds when $\mathcal{C}$ contains up to $M-1$ miners running the MPC and at least one but no more than d number of users.

In this work, we introduce a new notion which requires that a rational strategic player cannot harm other honest players without harming themselves. We formalize this as *Rational-Harm Proofness* (RHP).

Definition 3.2 (Restatement of Definition 1.1)**.** Given a TFM $(\mathbf{I}, \mathbf{C}, \mathbf{P}, \mathbf{R})$, we say that the TFM satisfies rational-harm proofness (RHP) w.r.t. a strategic player or coalition, denoted as $\mathcal{C}$, iff for any true value vector $\mathbf{v}$ of users, for any strategy $S_{\mathcal{C}}$ of coalition $\mathcal{C}$ such that $\mathsf{util}_{\mathcal{C}}(\mathbf{v}; S_{\mathcal{C}}) \geq \mathsf{util}_{\mathcal{C}}(\mathbf{v}; H_{\mathcal{C}})$, it must be that

$$\mathsf{util}_i(\mathbf{v}; S_{\mathcal{C}}) \geq \mathsf{util}_i(\mathbf{v}; H_{\mathcal{C}})$$

for any honest user $i \notin \mathcal{C}$. Additionally, for any miner $j \notin \mathcal{C}$, it must be that

$$\mathsf{mutil}_j(\mathbf{v}; S_{\mathcal{C}}) \geq \mathsf{mutil}_j(\mathbf{v}; H_{\mathcal{C}})$$

Equivalently, any strategy that benefits the coalition $\mathcal{C}$ or maintains the same utility as in the honest behavior must not harm any other player as well. Specifically, we say that a TFM satisfies

- *User rational-harm proofness (URHP)* if the above holds when $\mathcal{C}$ contains an individual user.

- *Miner rational-harm proofness (MRHP)* in the plain model if the above holds when $\mathcal{C}$ only contains the miner.

  In the MPC-assisted model, we say that a TFM satisfies MRHP if the above holds when $\mathcal{C}$ contains up to $M-1$ miners running the MPC implementing $\mathcal{F}_{\text{TFM}}$.

- *d-coalition rational-harm proofness (d-CRHP)* if the above holds when $\mathcal{C}$ contains the miner and at least one but no more than d number of users.

  In the MPC-assisted model, we say that a TFM satisfies d-CRHP for some integer $d \geq 1$ if the above holds when $\mathcal{C}$ contains up to $M-1$ miners running the MPC and at least one but no more than d number of users.

We also define a weaker notion called *weak RHP*: any strategy that *strictly* benefits the coalition must not harm any players outside the coalition.

Definition 3.3 (Weak RHP)**.** A TFM is *weakly RHP* w.r.t. $\mathcal{C}$ if Definition 3.2 holds with the strict inequality $u_{\mathcal{C}}(\mathbf{v}; S_{\mathcal{C}}) > u_{\mathcal{C}}(\mathbf{v}; H_{\mathcal{C}})$ in place of $\geq$. Equivalently, every *strictly profitable* deviation for $\mathcal{C}$ must be a Pareto improvement.

Fact 3.4. *If a TFM is incentive compatible w.r.t. some $\mathcal{C}$, then it is weak RHP w.r.t. $\mathcal{C}$.*

Proof. Since incentive compatibility rule out the existence of any strictly profitable strategies of $\mathcal{C}$ compared to honest behavior, it automatically achieves weak RHP w.r.t. $\mathcal{C}$. □

3.5 Myerson's Lemma

Our impossibility will rely on the famous Myerson's Lemma. Below, we use $x_i(\mathbf{b})$ and $p_i(\mathbf{b})$ to denote user i's probability of getting confirmed and its expected payment under input bid vector $\mathbf{b}$ when everyone behaves honestly. We also abuse notation and redefine $\mathsf{util}_i(\mathbf{b}) := b_i \cdot x_i(\mathbf{b}) - p_i(\mathbf{b})$ as the honest expected utility assuming $\mathbf{b}$ is the true value vector honest bid vector.

Lemma 3.5 (Myerson's Lemma [Mye81])**.** *A TFM satisfies UIC if and only if:*

- **Monotone allocation***: For any user i, any other users' bids $\mathbf{b}_{-i}$, any $b'_i > b_i$, it must be $x_i(\mathbf{b}_{-i}, b'_i) \geq x_i(\mathbf{b}_{-i}, b_i)$.*
- **Unique payment***: For any user i, any other users' bids $\mathbf{b}_{-i}$, and bid b_i from user i, user i's expected payment can be uniquely determined as*

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

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

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

4 Characterization in the Plain Model

4.1 Feasibility Results

In this section, we present the mechanisms referenced in Theorem 1.9 and show how each mechanism achieves any three of the four target properties: UIC, MIC, CRHP, and positive miner revenue. For completeness, we restate each mechanism with respect to the four rules for defining a TFM.

4.1.1 UIC + MIC + CRHP

All-or-nothing posted-price //Reserve price $\mathsf{res} \geq 0$.
//Revenue parameter $0 \leq \epsilon \leq \mathsf{res}$.

- **Inclusion & Confirmation**: If there are no more than k bids and all bids are strictly greater than res, include and confirm all bids. Otherwise, include and confirm no bids.
- **Payment & Revenue:** Each confirmed bid pays the reserve price res, and the miner receives a revenue share of ϵ per confirmed bid.

Lemma 4.1. *The above all-or-nothing posted-price TFM with $\epsilon = 0$ satisfies UIC, MIC, d-CRHP for any $d \geq 1$ in the plain model. Additionally, it also satisfies URHP.*

Proof. UIC follows from the fact that for any fixed user i, for fixed other users' bids $\mathbf{b}_{-i}$, user i's allocation is monotone and the price is exactly as defined in Lemma 3.5. MIC follows from the fact that miner receives no revenue. In the rest of this proof, we focus on proving d-CRHP for any $d \geq 1$ and URHP.

d-CRHP: Note that in this mechanism, either all bids are confirmed or no bids are confirmed. Suppose for the sake of contradiction that there exists a true value vector $\mathbf{v}$, a miner-user coalition $\mathcal{C}$ consisting of the miner and a non-empty set of users $\mathcal{U}$, and a strategy $S_\mathcal{C}$, such that $\mathsf{util}_\mathcal{C}(\mathbf{v}; S_\mathcal{C}) \geq \mathsf{util}_\mathcal{C}(\mathbf{v}; H_\mathcal{C})$, but the strategy $S_\mathcal{C}$ hurts some honest user $i \notin \mathcal{C}$. Because all bids pay res regardless of the bid vector, the only way to harm an honest user i is to make its bid unconfirmed whereas it gets confirmed in the honest case. That said, in the honest case, there must be no more than k bids and all bids are positive. This is only possible when all users, including those in $\mathcal{U}$ have a valuation greater than res and are also confirmed in the honest case. If a strategy $S_\mathcal{C}$ makes user i's bid unconfirmed, all bids are unconfirmed, which harms the coalition's joint utility as well. This contradicts the assumption that $S_\mathcal{C}$ does not harm the coalition.

URHP: Suppose for the sake of contradiction that there exists a true value vector $\mathbf{v}$, a use i, and a strategy S_i, such that $\mathsf{util}_i(\mathbf{v}; S_i) \geq \mathsf{util}_i(\mathbf{v}; H_i)$, but the strategy S_i hurts either some honest user $j \neq i$ or the miner. Similarly to the above reasoning, to harm a user j, the strategy S_i must unconfirm all bids, which harms user i as user i is confirmed in the honest case and $v_i > \mathsf{res}$. The miner cannot be harmed since its revenue is 0 anyway. □

Not 1-SCP: When there are more than k users, meaning that no bids are confirmed in the honest case, the miner can ignore all bids except a colluding player with positive valuation to guarantee that the colluding user is confirmed. This strictly improves the coalition's joint utility.

Not MRHP: If there are less than k users in the honest case, the miner can harm a user i with a positive true value by ignoring user i's bid. This does not influence miner's utility since it's 0 anyway, but harms the honest user i.

4.1.2 UIC + CRPH + Positive Revenue

Lemma 4.2. *The all-or-nothing posted-price TFM (Section 4.1.1) with $\epsilon > 0$ satisfies UIC, d-CRHP for any $d \geq 1$ and positive revenue in the plain model. Additionally, it satisfies URHP and MRHP.*

Proof. UIC and d-CRHP follow from the proof of Lemma 4.1. Positive miner revenue does not change any of the reasoning for these properties. The rest of the proof focuses on URHP and MRHP.

URHP: By the same reasoning as in Lemma 4.1, a strategic user cannot harm other users. A strategic user cannot harm the honest miner either, since otherwise the user must make all users unconfirm compared to the honest case, which harms the strategic user themselves.

MRHP: Notice that in this mechanism, for a user to be harmed, all users must be confirmed in the honest case and all users must be unconfirmed in the strategic case. Thus, the miner receives at least $\epsilon > 0$ revenue in the honest case and 0 in the strategic case if the strategy harms any user. □

Not MIC or 1-SCP: If $|S| > k$ and every bid is above res, the miner gets 0 revenue. by ignoring $|S| - k$ bids from S, the miner can achieve exactly $|S| = k$ and earn $k\epsilon > 0$ revenue. The above mechanism does not achieve 1-SCP for the same reason as the miner can collude with any of the included players for the same effect.

4.1.3 UIC + MIC + Positive Revenue

Posted-price with random selection //Reserve price $\mathsf{res} > 0$.
//Revenue parameter $0 < \epsilon \leq \mathsf{res}$.

- **Inclusion & Confirmation**: Let S be the set of bids strictly above the reserve price res. If $|S| \leq k$, include and confirm all bids in S. Otherwise, randomly choose k bids in S to include and confirm.
- **Payment & Revenue**: Each confirmed bid pays the reserve price res, and the miner receives a revenue share of ϵ per confirmed bid.

Lemma 4.3. *The above posted-price with random selection TFM satisfies UIC, MIC, and positive revenue in the plain model. Additionally, it achieves URHP.*

Proof. The mechanism's allocation and payment satisfies Myerson's lemma, and therefore achieves UIC. In the rest of this proof, we focus on MIC and URHP.

MIC: The miner receives res for each confirmed bid from users. Injecting fake bids does not increase the miner's utility. The miner maximizes its revenue from the users by confirming $\min(k, |S|)$ bids in S, which is exactly the honest inclusion rule.

URHP: Suppose for the sake of contradiction that there exists a true value vector $\mathbf{v}$, a use i, and a strategy S_i, such that $\mathsf{util}_i(\mathbf{v}; S_i) \geq \mathsf{util}_i(\mathbf{v}; H_i)$, but the strategy S_i hurts either some honest user $j \neq i$ or the miner. All confirmed bids pay the same price, so the only way to harm an honest user j is to is to lower their confirmation probability in comparison to the honest case. This is only possible if $v_j > \mathsf{res}$ and the strategic player injects fake bids above res to dilute the confirmation probability. However, then user i would have to pay res and gain nothing if a fake bid if confirmed, which reduces its expected utility compared to honest case. This contradicts our assumption.

The miner can only be harmed if user i was confirmed in the honest case but chooses not to bid to decrease the number of bids above res. As an implication, their honest utility is $v_i - \mathsf{res} > 0$. However, this means that not bidding harms the strategic user themselves.

□

Not 1-SCP: A miner can program the randomness to ensure that its colluding user gets confirmed as long as its colluding user has a true value higher than the reserved price res.

Not MRHP or 1-CRHP: By a similar strategy as above. The miner (or miner-user coalition) can program the randomness to reduce an honest user's probability of getting confirmed while keeping the miner (or the coalition)'s utility at least as good as the honest case.

4.1.4 MIC + CRHP + Positive Revenue

First-price auction

- **Inclusion & Confirmation:** Include and confirm the top k bids, breaking ties arbitrarily.
- **Payment & Revenue:** All confirmed bids pay their bid price and all payments go to the miner.

Lemma 4.4. *The above first price auction satisfies MIC, d-CRHP for any $d \geq 1$, and has positive miner revenue in the plain model. Additionally, it satisfies MRHP and d-SCP for any $d \geq 1$.*

Proof. **MIC:** The miner receives all payments, so they would have to increase payments from users to increase their utility. In the honest inclusion rule, the top bids are confirmed and paid, and there is no way to have payments higher than the bids.

d-**CRHP & MRHP:** If everyone bids honestly, the total utility of all users is 0, so there is trivially no way to harm other users.

d-**SCP:** In the honest case, because all payments go to the miner, the utility gained by a coalition and any set of users is equal to the summation of the top k valuations. It is impossible to gain more utility than this for any coalition due to the block size limit, budget feasibility, and individual rationality (See Section 3.2). □

Not UIC or URHP: Since users pay their own bid, there is an incentive for strategic underbidding, where a user can bid below their true valuation to increase utility while still being confirmed. This strategy both benefits the strategic user and harms the honest miner by reducing revenue.

4.2 Impossibilities in the Plain Model

We first introduce the following useful lemmas in our impossibility proofs. Recall that, for a bid vector $\mathbf{b} = (b_1, \ldots, b_n)$, for $i \in [n]$, $x_i(\mathbf{b})$, $p_i(\mathbf{b})$ and $\mathsf{util}_i(\mathbf{b})$ denote the confirmation probability, expected payment, and expected utility for bid b_i in $\mathbf{b}$ when the honest inclusion rule is followed. Also, $\widetilde{x}_{b_i}(\mathbf{b})$, $\widetilde{p}_{b_i}(\mathbf{b})$, and $\widetilde{\mathsf{util}}_{b_i}(\mathbf{b})$ are defined analogously for a unique bid b_i in bid vector $\mathbf{b}$.

Lemma 4.5. *For any TFM that satisfies UIC, given any $\mathbf{b}_{-i}$, let $b^* = \inf\{b : x_i(\mathbf{b}_{-i}, b) > 0\} < +\infty$. Then for any $v^* > v \geq b^*$, we have $\mathsf{util}_i(\mathbf{b}_{-i}, v^*) > \mathsf{util}(\mathbf{b}_{-i}, v)$.*

Proof. Suppose that for some $\mathbf{b}_{-i}$ and v, we have $x_i(\mathbf{b}_{-i}, v) > 0$. Then for any $v^* > v$,

$$\begin{aligned}
\mathsf{util}_i(\mathbf{b}_{-i}, v^*) &= v^* \cdot x_i(\mathbf{b}_{-i}, v^*) - p_i(\mathbf{b}_{-i}, v^*) \\
&= \int_0^{v^*} x_i(\mathbf{b}_{-i}, t)dt \qquad \text{by Myerson's Lemma (Lemma 3.5).} \\
&> \int_0^{v} x_i(\mathbf{b}_{-i}, t)dt = \mathsf{util}_i(\mathbf{b}_{-i}, v)
\end{aligned}$$

□

Lemma 4.6. *For any TFM that satisfies UIC, given any $\mathbf{b}_{-i}$, let $b^* = \inf\{b : x_i(\mathbf{b}_{-i}, b) > 0\} < +\infty$. Then for any $v^* \geq v > b^*$, the ratio $\frac{p_i(\mathbf{b}_{-i},v^*)}{x_i(\mathbf{b}_{-i},v^*)} \geq \frac{p_i(\mathbf{b}_{-i},v)}{x_i(\mathbf{b}_{-i},v)}$.*

Proof. Suppose that for some $\mathbf{b}_{-i}$, we have $x_i(\mathbf{b}_{-i}, v) > 0$. Then for any $v^* \geq v$,

$$\begin{aligned}
\frac{p_i(\mathbf{b}_{-i}, v^*)}{x_i(\mathbf{b}_{-i}, v^*)} &= v^* - \frac{\int_0^{v^*} x_i(\mathbf{b}_{-i}, t)dt}{x_i(\mathbf{b}_{-i}, v^*)} \qquad \text{by Myerson's Lemma (Lemma 3.5)} \\
&= \left(v - \frac{\int_0^{v} x_i(\mathbf{b}_{-i}, t)dt}{x_i(\mathbf{b}_{-i}, v^*)}\right) + \left((v^* - v) - \frac{\int_v^{v^*} x_i(\mathbf{b}_{-i}, t)dt}{x_i(\mathbf{b}_{-i}, v^*)}\right) \\
&\geq \left(v - \frac{\int_0^{v} x_i(\mathbf{b}_{-i}, t)dt}{x_i(\mathbf{b}_{-i}, v^*)}\right) + \left((v^* - v) - \frac{(v^* - v)\left(x_i(\mathbf{b}_{-i}, v^*) - x_i(\mathbf{b}_{-i}, v)\right)}{x_i(\mathbf{b}_{-i}, v^*)}\right) \\
&\geq v - \frac{\int_0^{v} x_i(\mathbf{b}_{-i}, t)dt}{x_i(\mathbf{b}_{-i}, v^*)} \\
&\geq v - \frac{\int_0^{v} x_i(\mathbf{b}_{-i}, t)dt}{x_i(\mathbf{b}_{-i}, v)} = \frac{p_i(\mathbf{b}_{-i}, v)}{x_i(\mathbf{b}_{-i}, v)}.
\end{aligned}$$

□

Lemma 4.7. *Suppose a TFM* $(\mathbf{I}, \mathbf{C}, \mathbf{P}, \mathbf{R})$ *satisfies MIC in the plain model. For any* $\mathbf{b} \subseteq \mathbf{b}'$, *it must be that* $\mu(\mathbf{b}') \geq \mu(\mathbf{b})$.

Proof. For the sake of contradiction, suppose this was not the case, so $\mu(\mathbf{b}') < \mu(\mathbf{b})$. Then, when the miner sees the honest bid vector $\mathbf{b}'$, they censor the bids not in $\mathbf{b}$ perform the honest inclusion rule as if the input bid vector is $\mathbf{b}$ to gain more revenue. This contradicts MIC. □

Lemma 4.8. *Suppose a TFM* $(\mathbf{I}, \mathbf{C}, \mathbf{P}, \mathbf{R})$ *satisfies MIC in the plain model. Then for any bid vectors* $\mathbf{b}, \mathbf{b}', \mathbf{v}$, *we have* $\mu(\mathbf{b}, \mathbf{v}) \geq \mu(\mathbf{b}', \mathbf{v}) - \sum_{i=1}^{|\mathbf{b}'|} p_i(\mathbf{b}', \mathbf{v})$.

Proof. For the sake of contradiction, suppose $\mu(\mathbf{b}, \mathbf{v}) < \mu(\mathbf{b}', \mathbf{v}) - \sum_{i=1}^{|\mathbf{b}'|} p_i(\mathbf{b}', \mathbf{v})$. Consider the world where $(\mathbf{b}, \mathbf{v})$ is the honest bid vector. The miner's honest utility is $\mu(\mathbf{b}, \mathbf{v})$. Consider the following miner strategy: it deviates from the honest inclusion rule by ignoring the bids in $\mathbf{b}$ and submitting the fake bids in $\mathbf{b}'$ to achieve the bid vector $(\mathbf{b}', \mathbf{v})$. Inserting the fake bids costs the miner $\sum_{i=1}^{|\mathbf{b}'|} p_i(\mathbf{b}', \mathbf{v})$ so the miner's utility is

$$\mu(\mathbf{b}', \mathbf{v}) - \sum_{i=1}^{|\mathbf{b}'|} p_i(\mathbf{b}', \mathbf{v}) > \mu(\mathbf{b}, \mathbf{v}), \qquad \text{by assumption}$$

which contradicts MIC. □

In a (possibly randomized) mechanism, the inclusion rule, confirmation rule, payment rule and miner revenue rule can be randomized. However, the randomness used in the inclusion rule is freely chosen by the miner, whereas the randomness used in other rules come from the blockchain. Let Ω and $\mathcal{D}$ be the sample space and distribution of the randomness specified by the honest inclusion rule. For each $r \in \Omega$, define $x_i(\mathbf{b} \mid r)$, $p_i(\mathbf{b} \mid r)$, and $\mu(\mathbf{b} \mid r)$ as the expected confirmation probability of bid b_i, the expected payment of bid b_i, and the expected miner revenue, respectively, conditioned on the randomness used in the inclusion rule being r. The expectation in each case is now taken over the randomness used in the confirmation, payment, and miner revenue rules, respectively.

By definition,

$$x_i(\mathbf{b}) = \mathop{\mathbb{E}}_{r \leftarrow \mathcal{D}}[x_i(\mathbf{b} \mid r)], \qquad p_i(\mathbf{b}) = \mathop{\mathbb{E}}_{r \leftarrow \mathcal{D}}[p_i(\mathbf{b} \mid r)], \qquad \mu(\mathbf{b}) = \mathop{\mathbb{E}}_{r \leftarrow \mathcal{D}}[\mu(\mathbf{b} \mid r)].$$

We define $\mathsf{util}_i(\mathbf{b} \mid r)$ analogously as user i's honest expected utility conditioned on randomness r used in the inclusion rule when the input bid vector is $\mathbf{b}$. We similarly define $\widetilde{\mathsf{util}}_v(\mathbf{b})$ as the expected utility of a *unique* bid v in $\mathbf{b}$. Let

$$W(\mathbf{b}) \colon = \{i \in [n] : \mathsf{util}_i(\mathbf{b}) > 0\}$$

to be the set of users whose honest expected utility is strictly positive under the bid vector $\mathbf{b}$.

Lemma 4.9. *Suppose a TFM* $(\mathbf{I}, \mathbf{C}, \mathbf{P}, \mathbf{R})$ *satisfies weak* 1*-CRHP in the plain model. For any bid vector* $\mathbf{b} = (b_1, \ldots, b_n)$ *where there exists a unique* b_i *in* $\mathbf{b}$ *such that* $\mathsf{util}_i(\mathbf{b}) > 0$, *for all unique* b_j *in* $\mathbf{b}$ *where* $b_j > \frac{p_i(\mathbf{b})}{x_i(\mathbf{b})}$, *we have* $x_j(\mathbf{b}) > 0$.

In words, if there is a unique bid b_i *in* $\mathbf{b}$ *with a positive utility, all unique bids greater than* b_i *in* $\mathbf{b}$ *must have positive confirmation probability as well.*

Proof. Seeking contradiction, suppose there exists a bid vector $\mathbf{b} = (b_1, \ldots, b_n)$ where b_i is a unique bid in $\mathbf{b}$ and $\mathsf{util}_i(\mathbf{b}) > 0$, but $x_j(\mathbf{b}) = 0$ for some unique $b_j > \frac{p_i(\mathbf{b})}{x_i(\mathbf{b})}$ in $\mathbf{b}$.

Consider a world where $\mathbf{b}$ is the honest bid vector. Consider a coalition consisting of user j with true value b_j and the miner, whose joint expected utility in the honest case equals $\mu(\mathbf{b})$. The coalition can perform the following strategy:

- The miner ignores the bid b_i from user i;
- User j underbids to b_i and injects a fake bid of value b_j under an arbitrary fake identity.
- The miner performs the honest inclusion rule on the remaining bids.

Because the bids b_i and b_j are unique, by weak symmetry, the fake bid at b_j will not be confirmed, while user j's real bid at b_i will have positive confirmation probability with expected payment $p_i(\mathbf{b})$. Under the deviation, the coalition's joint utility equals

$$\mu(b) + b_j \cdot x_i(\mathbf{b}) - p_i(\mathbf{b}) > \mu(b) + \frac{p_i(\mathbf{b})}{x_i(\mathbf{b})} \cdot x_i(\mathbf{b}) - p_i(\mathbf{b}) > \mu(b),$$

where the rightmost expression is the coalition's expected joint utility in the honest case. However, user i in the honest case has positive utility, but is now unconfirmed. Thus, this strategy increases the coalition's joint utility while harming the honest user i, contradicting 1-CRHP. □

Lemma 4.10. *Suppose a TFM* $(\mathbf{I}, \mathbf{C}, \mathbf{P}, \mathbf{R})$ *satisfies MIC in the plain model. Then, for any* $\mathbf{b}$, *we have* $\mu(\mathbf{b} \mid r) = \mu(\mathbf{b})$ *almost surely:*

$$\Pr_{r \leftarrow \mathcal{D}}[\mu(\mathbf{b} \mid r) = \mu(\mathbf{b})] = 1.$$

Proof. By MIC, it must be that $\mu(\mathbf{b} \mid r) \leq \mu(\mathbf{b})$ for any $r \in \Omega$. Otherwise, the miner can fix r as the randomness used in the inclusion rule instead of sampling the randomness from $\mathcal{D}$ and strictly increases its expected revenue. Since $\mu(\mathbf{b}) = \mathop{\mathbb{E}}_{r \leftarrow \mathcal{D}}[\mu(\mathbf{b} \mid r)]$, we have $\Pr_{r \leftarrow \mathcal{D}}[\mu(\mathbf{b} \mid r) = \mu(\mathbf{b})] = 1$. □

Lemma 4.11. *Suppose a TFM* $(\mathbf{I}, \mathbf{C}, \mathbf{P}, \mathbf{R})$ *satisfies MIC and weak* 1*-CRHP, and the block size is bounded by* k *in the plain model. Then, for every bid vector* $\mathbf{b} = (b_1, \ldots, b_n)$,

$$|W(\mathbf{b})| \leq k,$$

i.e., at most k *users can have positive expected utility under the honest mechanism.*

Proof. Assume, for the sake of contradiction, that there exists a bid vector $\mathbf{b}$ where $|W(\mathbf{b})| > k+1$. Since the block size is at most k, the inclusion rule cannot include more than k bids for any randomness $r \in \Omega$. Therefore, there must exists some $i \in W(\mathbf{b})$ and a set $\Omega_i := \{r \in \Omega : x_i(\mathbf{b} \mid r) = 0\}$ such that $\Pr_{r \leftarrow \mathcal{D}}[\Omega_i] > 0$, this implies that

$$\Pr_{r \leftarrow \mathcal{D}}[\mathsf{util}_i(\mathbf{b} \mid r) = 0] > 0 \tag{4}$$

Recall that by definition, $\mathsf{util}_i(\mathbf{b}) = \mathbb{E}_{r \leftarrow \mathcal{D}}[\mathsf{util}_i(\mathbf{b} \mid r)] > 0$. Together with Equation (4), this means that $\Pr_{r \leftarrow \mathcal{D}}[\Omega_i'] > 0$, where $\Omega_i' = \{r \in \Omega : \mathsf{util}_i(\mathbf{b} \mid r) > \mathsf{util}_i(\mathbf{b})\}$.

For every fixed randomness, at most k bids can be included. Thus, for each $r \in \Omega_i'$, there exists some $j \in W(\mathbf{b})$ such that $x_j(\mathbf{b} \mid r) = 0$. By the pigeonhole principle, there exists some $j^* \in W(\mathbf{b})$ such that

$$\Pr_{r \leftarrow \mathcal{D}}[\mathsf{util}_i(\mathbf{b} \mid r) > \mathsf{util}_i(\mathbf{b}) \text{ and } x_{j^*}(\mathbf{b} \mid r) = 0] > 0.$$

Define set $\Omega^* := \{r \in \Omega : \mathsf{util}_i(\mathbf{b} \mid r) > \mathsf{util}_i(\mathbf{b})$ and $x_{j^*}(\mathbf{b} \mid r) = 0$ and $\mu(\mathbf{b} \mid r) = \mu(\mathbf{b})\}$. By Lemma 4.10, we have $\Pr_{r \leftarrow \mathcal{D}}[\Omega^*] > 0$.

Consider the following strategy by a coalition of the miner and user i as follows: the miner picks an arbitrary randomness $r^* \in \Omega^*$ in the inclusion rule. By the choice of Ω^*, the coalition's expected joint utility becomes:

$$\mu(\mathbf{b} \mid r^*) + \mathsf{util}_i(\mathbf{b} \mid r^*) = \mu(\mathbf{b}) + \mathsf{util}_i(\mathbf{b} \mid r^*) > \mu(\mathbf{b}) + \mathsf{util}_i(\mathbf{b}),$$

which strictly improves the coalition's expected joint utility. Meanwhile, user j^*, who, in the honest case, has positive expected utility, is now excluded and receives zero utility, violating 1-CRHP. □

Throughout the proof, we will rely on bids having strictly positive utility. Therefore, in our argument, we often need to remove the infimum value from a set, which might have zero utility. For this purpose, we define an operator $\mathsf{rmInf}(\cdot)$. For any set $S \subset \mathbb{R}$, $\mathsf{rmInf}(S) = S \setminus \inf S$. Recall that for two sets S and S', we say $S \prec S'$ if $\sup S \leq \inf S'$.

Lemma 4.12. *Let $(\mathbf{I}, \mathbf{C}, \mathbf{P}, \mathbf{R})$ be a TFM that satisfies UIC, MIC, and weak 1-CRHP. Suppose there exists a bid vector $\mathbf{v} = (v_1, \ldots, v_n)$ and a sequence of sets $S_1, ..., S_m \subseteq \mathbb{R}_{\geq 0}$ for some $m \in \mathbb{N}$, such that for any $\mathbf{b} = (b_1, \ldots, b_m) \in S_1 \times \cdots \times S_m$, we have each b_ℓ being unique in $(\mathbf{v}, \mathbf{b})$ and $\widetilde{x}_{b_1}(\mathbf{v}, \mathbf{b}) > 0$. Additionally, suppose*

- *for all $\ell \in [m]$, $|S_\ell| \geq 2$;*
- *$S_1 \prec \cdots \prec S_m$.*

Then, there exists a bid vector $\mathbf{b}^*$ such that $|W(\mathbf{b}^*)| \geq m$.

Proof. We prove this by direct construction of $\mathbf{b}^*$.

Claim 4.13. *For $b_1 \in \mathsf{rmInf}(S_1)$ and $\mathbf{b}' \in (S_2 \times \cdots \times S_m)$, we have $\widetilde{\mathsf{util}}_{b_1}(\mathbf{v}, b_1, \mathbf{b}') > 0$.*

Proof. By assumption, for $\widehat{b}_1 \in S_1$ where $\widehat{b}_1 < b_1$, which exists by the first bullet point, we have $\widetilde{x}_{\widehat{b}_1}(\mathbf{v}, \widehat{b}_1, \mathbf{b}') > 0$. Thus, by Lemma 4.5, $\widetilde{\mathsf{util}}_{b_1}(\mathbf{v}, b_1, \mathbf{b}') > \widetilde{\mathsf{util}}_{\widehat{b}_1}(\mathbf{v}, \widehat{b}_1, \mathbf{b}') \geq 0$. □

Pick an arbitrary $\mathbf{b} = (b_1, \ldots, b_m) \in \mathsf{rmInf}(S_1) \times \cdots \times \mathsf{rmInf}(S_m)$. Let $\mathbf{b}^* = (\mathbf{v}, \mathbf{b})$. By assumption, all b_ℓ for $\ell \in [m]$ are unique in $\mathbf{b}^*$.

For any fixed $\ell \in [m]$, pick an arbitrary $\widehat{b}_\ell \in S_\ell$ where $\widehat{b} < b_\ell$. Since $\mathsf{rmInf}(S_\ell)$ is non-empty, such $\widehat{b}_\ell$ must exist. Further more, by assumption $\widehat{b}_\ell$ is unique in the bid vector $(\mathbf{b}^*_{-(n+\ell)}, \widehat{b}_\ell)$.

For $\ell \neq 1$, by Claim 4.13 and Lemma 4.9, we have

$$\widetilde{x}_{b'_\ell}\left(\mathbf{b}^*_{-(n+\ell)}, \widehat{b}_\ell\right) > 0.$$

By Lemma 4.5, we conclude that

$$\widetilde{\mathsf{util}}_{b_\ell}(\mathbf{b}^*) > \widetilde{\mathsf{util}}_{b'_\ell}(\mathbf{b}^*_{-(n+\ell)}, \widehat{b}_\ell) \geq 0$$

for each $\ell \in [m]$. Thus, $|W(\mathbf{b}^*)| \geq m$. □

Our impossibility relies on the following key technical lemma.

Lemma 4.14. *Suppose a TFM $(\mathbf{I}, \mathbf{C}, \mathbf{P}, \mathbf{R})$ satisfies UIC, MIC, and 1-CRHP, and let $\mathbf{v}$ be a bid vector such that $|\mathbf{v}| \geq k + 1$ and index i^* where $x_{i^*}(\mathbf{v}) > 0$. Then, for any $k \in \mathbb{N}$, there exists a bid vector $\mathbf{b}^*$ such that*

$$W(\mathbf{b}^*) > k.$$

Combining Lemma 4.14 and Lemma 4.11, we immediately get the following impossibility result:

Theorem 4.15. *Let k denote the block size. If a TFM $(\mathbf{I}, \mathbf{C}, \mathbf{P}, \mu)$ satisfies UIC, MIC, and 1-CRHP, for every bid vector $\mathbf{v}$ where $|\mathbf{v}| \geq k+1$, no bids have positive confirmation probability.*

Proof. By Lemma 4.14 and Lemma 4.11. □

4.2.1 Proof of Lemma 4.14

Suppose, for the sake of contradiction, that there exists $\mathbf{v}$ where $|\mathbf{v}| \geq k+1$ and $x_i(\mathbf{v}) > 0$ for an index $i \in [n]$. We will prove a stronger result using induction to construct $\mathbf{b}^*$ such that $W(\mathbf{b}^*) > k$.

For any $m \in [k+1]$, our induction constructs a sequence of positive real numbers $\epsilon^{(1)}, ..., \epsilon^{(m)}$, a sequence of sets $S^{(1)}, \ldots, S^{(m)} \subset \mathbb{R}_{\geq 0}$ such that properties (Prop′-1)-(Prop′-3) below hold.

(Prop′-1) Each $S^{(\ell)}$ contains two values: $S^{(\ell)} = \{\alpha^{(\ell)}, \beta^{(\ell)}\}$ with $\alpha^{(\ell)} < \beta^{(\ell)}$.

(Prop′-2) $\max(\mathbf{v}) + \epsilon^{(m)} < \alpha^{(m)} < \beta^{(m)} < \alpha^{(m-1)} < \beta^{(m-1)} < \cdots < \alpha^{(1)} < \beta^{(1)}$.
This implies that $\left\{\max(\mathbf{v}) + \epsilon^{(m)}\right\} \prec S^{(m)} \prec \cdots \prec S^{(1)}$

(Prop′-3) For all $\mathbf{b} = (b_1, \ldots, b_m) \in S^{(1)} \times \cdots \times S^{(m)}$, b_m is unique in $(\mathbf{v}_{-i}, \mathbf{b})$. Moreover,

$$\widetilde{x}_{b_m}(\mathbf{v}_{-i}, \mathbf{b}) > 0 \quad \text{and} \quad \frac{\widetilde{p}_{b_m}(\mathbf{v}_{-i}, \mathbf{b})}{\widetilde{x}_{b_m}(\mathbf{v}_{-i}, \mathbf{b})} \leq \alpha^{(m)} - \epsilon^{(m)}.$$

As an implication, $\widetilde{\mathsf{util}}_{b_m}(\mathbf{v}_{-i}, \mathbf{b}) > 0$.

We first show how the lemma follows assuming that the above properties holds for any $m \in [k+1]$ and give the induction proof afterwards. Let $m = k+1$. Consider $\mathbf{v}_{-i}$ and the sequence of sets $S^{(1)}, ..., S^{(k+1)}$. By (Prop′-1), $|S^{(\ell)}| \geq 2$ for all $\ell \in [k+1]$. From (Prop′-2), we also know that $S^{(1)} \prec \cdots \prec S^{(k+1)}$, and that for $\mathbf{b} = (b_1, \ldots, b_m) \in S^{(1)} \times \cdots \times S^{(m)}$, each b_ℓ is unique in $(\mathbf{v}_{-i}, \mathbf{b})$. Furthermore, $\widetilde{x}_{b_m}(\mathbf{v}_{-i}, \mathbf{b}) > 0$ by (Prop′-3). Thus, the assumptions of Lemma 4.12 hold for $\mathbf{v}_{-i}$ and the sets $S^{(k+1)}, ..., S^{(1)}$, which means that some $\mathbf{b}^*$ exists where $|W(\mathbf{b}^*)| \geq k+1$.

Induction Proof. In the rest of the proof, we focus on the induction for proving properties (Prop′-1) - (Prop′-3). For clarity, in our induction, we use superscripts to refer to the propositions for a particular $m \in [k+1]$, i.e. (Prop′-1)$^{(m)}$ represents property (Prop′-1) for m.

Base Case: $m = 1$. Let $v^* = \max(\mathbf{v}) + 1$. Consider a bid vector $\mathbf{v}' = (\mathbf{v}_{-i}, v^*)$. Since $x_i(\mathbf{v}) > 0$, by Lemma 4.5, we know that $\mathsf{util}_i(\mathbf{v}_{-i}, v^*) > \mathsf{util}_i(\mathbf{v}) \geq 0$, and thus $\tau := \frac{p_i(\mathbf{v}')}{x_i(\mathbf{v}')} < v^*$. Define

$$\epsilon^{(1)} := \frac{1}{2}(v^* - \max(\mathbf{v}, \tau)) > 0 \quad \text{and} \quad S^{(1)} := \left\{\alpha^{(1)} = v^* - \epsilon^{(1)}, \beta^{(1)} = v^* - \frac{1}{2}\epsilon^{(1)}\right\}.$$

We now prove the properties:

- (Prop′-1)$^{(1)}$ and (Prop′-2)$^{(1)}$ are true by construction.
- Pick an arbitrary $b_1 \in S^{(1)}$. By Myerson's Lemma (Lemma 3.5), $\widetilde{x}_{b_1}(\mathbf{v}_{-i}, b_1) \geq x_i(\mathbf{v}) > 0$. Furthermore, we have

$$\begin{aligned} \frac{\widetilde{p}_{b_1}(\mathbf{v}_{-i}, b_1)}{\widetilde{x}_{b_1}(\mathbf{v}_{-i}, b_1)} &\leq \frac{p_i(\mathbf{v}')}{x_i(\mathbf{v}')} && \text{by Lemma 4.6} \\ &= \tau \leq \max(\mathbf{v}, \tau) \leq \alpha^{(1)} - \epsilon^{(1)}. \end{aligned}$$

  Thus, (Prop′-3)$^{(1)}$ is satisfied.

Inductive Steps: $1 < m \leq k+1$. Suppose $S^{(1)}, \ldots, S^{(m-1)}$ and $\epsilon^{(m-1)}$ are the sequence of sets and the real number respectively that satisfy (Prop′-1)$^{(m-1)}$-(Prop′-3)$^{(m-1)}$. We now construct $S^{(m)}$ and $\epsilon^{(m)}$. Recall that $\alpha^{(m-1)} = \min S^{(m-1)}$.

Claim 4.16. *For all* $\mathbf{b} = (b_1, \ldots, b_{m-1}) \in S_1^{(m-1)} \times \cdots \times S_{m-1}^{(m-1)}$ *and any* $\tau \in (\alpha^{(m-1)} - \epsilon^{(m-1)}, \alpha^{(m-1)})$, *let* $\mathbf{b}_\tau = ((\mathbf{v}_{-i}, \tau), \mathbf{b})$. *Then* τ *is a unique bid in* $\mathbf{b}_\tau$ *and* $\widetilde{x}_\tau(\mathbf{b}_\tau) > 0$.

Proof. By (Prop′-2)$^{(m-1)}$, τ and b_{m-1} are both unique bids in $\mathbf{b}_\tau$. In the rest of this proof, we focus on proving that $\widetilde{x}_\tau(\mathbf{b}_\tau) > 0$. Suppose for the sake of contradiction that $\widetilde{x}_\tau(\mathbf{b}_\tau) = 0$. We first show that

$$\widetilde{\mathsf{util}}_{b_{m-1}}(\mathbf{b}_\tau) \geq \widetilde{\mathsf{util}}_{b_{m-1}}(\mathbf{v}_{-i}, \mathbf{b}). \tag{5}$$

Suppose for the sake of contradiction that $\widetilde{\mathsf{util}}_{b_{m-1}}(\mathbf{b}_\tau) < \widetilde{\mathsf{util}}_{b_{m-1}}(\mathbf{v}_{-i}, \mathbf{b})$. Consider a world where $(\mathbf{v}_{-i}, \mathbf{b})$ is the honest input bid vector. Since $|\mathbf{v}| > k$, by Lemma 4.11, there must exist some user j whose expected utility is 0 in the honest case. Now consider a coalition made up of the miner and user j that performs the following strategy: they inject a fake bid at τ, faking a world with input bid vector $\mathbf{b}_\tau$, and performs the honest mechanism on $\mathbf{b}_\tau$. The fake bid τ must be unconfirmed by assumption. In the honest case, their expected joint utility is just the expected miner revenue $\mu(\mathbf{v}_{-i}, \mathbf{b})$. The coalition's strategic utility is $\mu(\mathbf{b}_\tau) + \mathsf{util}_j(\mathbf{b}_\tau) \geq \mu(\mathbf{v}_{-i}, \mathbf{b})$ by Lemma 4.7 and Lemma 4.8. However, user i is harmed by the assumption that $\widetilde{\mathsf{util}}_{b_{m-1}}(\mathbf{b}_\tau) < \widetilde{\mathsf{util}}_{b_{m-1}}(\mathbf{v}_{-i}, \mathbf{b})$, which contradicts 1-CRHP.

Now we are ready to prove the claim statement. Consider a world where $\mathbf{b}_\tau$ is the honest bid vector. Let u denote the user whose true value is b_{m-1} and let u' denote the user who's valuation is τ. Then user u's expected utility in the honest case is

$$\begin{aligned}
\widetilde{\mathsf{util}}_{b_{m-1}}(\mathbf{b}_\tau) &\geq \widetilde{\mathsf{util}}_{b_{m-1}}(\mathbf{v}_{-i}, \mathbf{b}) && \text{by Equation (5)} \\
&> 0 && \text{by (Prop}'\text{-3)}^{(m-1)}
\end{aligned}$$

Consider a coalition made of the user u' and the miner. Since $\widetilde{x}_\tau(\mathbf{b}_\tau) = 0$ by the assumption, the coalition's expected joint utility in the honest case is simply the expected miner revenue $\mu(\mathbf{b}_\tau)$. Consider the following strategy of the coalition:

- The miner ignores the bid b_{m-1} from user u;
- User u' raises its bid to b_{m-1}. The miner then performs the honest inclusion rule on $(\mathbf{v}_{-i}, \mathbf{b})$ where the bid at b_{m-1} now comes from the coalition.

Under this strategy, the coalition's expected joint utility becomes

$$\begin{aligned}
&\mu\left(\mathbf{v}_{-i}, \mathbf{b}\right) + \left(\tau \cdot \widetilde{x}_{b_{m-1}}\left(\mathbf{v}_{-i}, \mathbf{b}\right) - \widetilde{p}_{b_{m-1}}\left(\mathbf{v}_{-i}, \mathbf{b}\right)\right) \\
> &\mu\left(\mathbf{v}_{-i}, \mathbf{b}\right) + \left(\gamma - \epsilon^{(m-1)}\right) \cdot \widetilde{x}_{b_{m-1}}\left(\mathbf{v}_{-i}, \mathbf{b}\right) - \widetilde{p}_{b_{m-1}}\left(\mathbf{v}_{-i}, \mathbf{b}\right) && \text{by choice of } \tau \\
\geq &\mu\left(\mathbf{v}_{-i}, \mathbf{b}\right) + \left(\frac{\widetilde{p}_{b_{m-1}}\left(\mathbf{v}_{-i}, \mathbf{b}\right)}{\widetilde{x}_{b_{m-1}}\left(\mathbf{v}_{-i}, \mathbf{b}\right)}\right) \cdot \widetilde{x}_{b_{m-1}}\left(\mathbf{v}_{-i}, \mathbf{b}\right) - \widetilde{p}_{b_{m-1}}\left(\mathbf{v}_{-i}, \mathbf{b}\right) && \text{by (Prop}'\text{-3)}^{(m-1)} \\
= &\mu\left(\mathbf{v}_{-i}, \mathbf{b}\right) \geq \mu(\mathbf{b}_\tau) && \text{by Lemma 4.8}
\end{aligned}$$

The last step comes from the assumption that $\widetilde{x}_\tau(\mathbf{b}_\tau) = 0$. This strategy strictly benefits the coalition while harming honest user u, which contradicts 1-CRHP. □

Consider the value

$$\delta := \min_{\mathbf{b} \in S^{(1)} \times \cdots \times S^{(m-1)}} \left(\alpha^{(m-1)} - \frac{p_i((\mathbf{v}_{-i}, \alpha^{(m-1)}), \mathbf{b})}{x_i((\mathbf{v}_{-i}, \alpha^{(m-1)}), \mathbf{b})} \right). \tag{6}$$

By Claim 4.16 and Lemma 4.5, $\frac{p_i((\mathbf{v}_{-i}, \alpha^{(m-1)}), \mathbf{b})}{x_i((\mathbf{v}_{-i}, \alpha^{(m-1)}), \mathbf{b})} < \alpha^{(m-1)}$ for all $\mathbf{b} \in S^{(1)} \times \cdots \times S^{(m-1)}$. Since there is only a finite number of possible $\mathbf{b}$ by $(\mathsf{Prop'\text{-}1})^{(m-1)}$, we have $\delta > 0$. Define

$$\epsilon^{(m)} := \frac{1}{2} \min\left(\delta, \epsilon^{(m-1)}\right) \quad \text{and} \quad S^{(m)} := \left\{ \alpha^{(m)} = \alpha^{(m-1)} - \epsilon^{(m)}, \beta^{(m)} = \alpha^{(m-1)} - \frac{1}{2}\epsilon^{(m)} \right\}$$

- $(\mathsf{Prop'\text{-}1})^{(m)}$ follows by construction.
- By construction, $\beta^{(m)} < \alpha^{(m-1)}$. Moreover,

  $$\begin{aligned} \alpha^{(m)} = \alpha^{(m-1)} - \epsilon^{(m)} &\geq \max(\mathbf{v}) + \epsilon^{(m-1)} - \epsilon^{(m)} && \text{by } (\mathsf{Prop'\text{-}3})^{(m-1)} \\ &\geq \max(\mathbf{v}) + \epsilon^{(m)}. && \text{by choice of } \epsilon^{(m)} \end{aligned}$$

  $(\mathsf{Prop'\text{-}2})^{(m)}$ thus follows by combining $\max(\mathbf{v}) + \epsilon^{(m)} \leq \alpha^{(m)} < \beta^{(m)} < \alpha^{(m-1)}$ and $(\mathsf{Prop'\text{-}2})^{(m-1)}$.
- Let $\mathbf{b} = (b_1, \ldots, b_m) \in S^{(1)} \times \cdots \times S^{(m)}$. By Claim 4.16, $\widetilde{x}_{b_m}(\mathbf{v}_{-i^*}, \mathbf{b}) > 0$. Furthermore, we have

  $$\begin{aligned} \frac{\widetilde{p}_{b_m}(\mathbf{v}_{-i^*}, \mathbf{b})}{\widetilde{x}_{b_m}(\mathbf{v}_{-i^*}, \mathbf{b})} &\leq \frac{p_{n+m}(\mathbf{v}_{-i^*}, \mathbf{b}_{-m}, \alpha^{(m-1)})}{x_{n+m}(\mathbf{v}_{-i^*}, \mathbf{b}_{-m}, \alpha^{(m-1)})} && \text{by Lemma 4.6} \\ &\leq \alpha^{(m-1)} - \delta && \text{by Equation (6)} \\ &= \alpha^{(m)} - \epsilon^{(m)}. && \text{by choice of } \epsilon^{(m)} \end{aligned}$$

  Thus, $(\mathsf{Prop'\text{-}3})^{(m)}$ is satisfied.

Corollary 4.17. *Let k denote the block size. If a TFM $(\mathbf{I}, \mathbf{C}, \mathbf{P}, \mu)$ satisfies UIC, MIC, and 1-CRHP, then the miner revenue must be zero, i.e., $\mu(\mathbf{b}) = 0$ for any $\mathbf{b}$.*

Proof. Seeking contradiction, suppose that there was a transaction fee mechanism $(\mathbf{I}, \mathbf{C}, \mathbf{P}, \mathbf{R})$ that satisfies UIC, MIC, and 1-CRHP, and has finite block size k, and for a bid vector $\mathbf{b}$, $\mu(\mathbf{b}) > 0$. Now consider $\mathbf{b}' = (\mathbf{b}, \underbrace{0, \ldots, 0}_{k+1}) = (b'_1, \ldots, b'_n)$. By Lemma 4.7, $\mu(\mathbf{b}') \geq \mu(\mathbf{b}) > 0$. Thus, by budget feasibility (see Section 3.2), there exists $i \in [n]$ where $x_i(\mathbf{b}') = 1$. Additionally, $n > k$, so by Theorem 4.15, this is impossible, leading to a contradiction. □

4.3 Impossibility Under Weak-CRHP

In this section, we present the proof of the stronger impossibility result under weak-CRHP. The proof relies on the following technical lemma.

Lemma 4.18. *Suppose a TFM $(\mathbf{I}, \mathbf{C}, \mathbf{P}, \mathbf{R})$ satisfies UIC, MIC, and weak 1-CRHP, and let $\mathbf{v}$ be a bid vector such that $\mu(\mathbf{v}) > 0$. Then, there exists a bid vector $\mathbf{b}^*$ such that*

$$W(\mathbf{b}^*) > k.$$

Combining Lemma 4.18 and Lemma 4.11, we immediately get the following impossibility result:

Theorem 4.19 (Restatement of Theorem 1.3)**.** *Let k denote the block size. If a TFM $(\mathbf{I}, \mathbf{C}, \mathbf{P}, \mathbf{R})$ satisfies UIC, MIC, and weak 1-CRHP, then miner revenue must be zero, i.e. $\mu(\mathbf{b}) = 0$ for any $\mathbf{b}$.*

Proof. Note that Lemma 4.11 holds for weak 1-CRHP TFMs. The theorem thus follows from Lemma 4.18 and Lemma 4.11. □

Proof of Lemma 4.18 As with previous proofs, we start with a bid vector $\mathbf{v} = (v_1, \ldots, v_n)$ such that $\mu(\mathbf{v}) > 0$ and construct inductively a sequence of sets of vectors, such that in the m-th step, we have some vectors $\mathbf{v}^{(m)}$ such that $W(\mathbf{v}^{(m)}) \geq m$. Then when $m = k+1$, this contradicts Lemma 4.11. To aide in our proof, we utilize the following implication of the Graham–Rothschild theorem [GRS90] to determine the size of our sets.

Lemma 4.20. *(from [GRS90]) For all $p, \xi, a > 0$ there exists N such that, if $G \subseteq A_1 \times \cdots \times A_p$, $|A_i| = N$ for each i, and if $|G| \geq \xi N^p$, then there exist subsets $B_i \subseteq A_i$ with $|B_i| \geq a$ such that*

$$B_1 \times \cdots \times B_p \subseteq G.$$

Like in the proof of Lemma 4.14, in each inductive step, we build a sequence of sets $S_1^{(m)}, ..., S_m^{(m)}$, each of size at least two. To ensure that we can pick sets of size at least two in step $k+1$, we define the following sequence N_m for $m \in [k+2]$ backwards:

$$\begin{aligned} &N_{k+2} := 2, \\ &N_m := N \text{ in Lemma 4.20 with } (p = k+m-1,\ \xi{=}\tfrac{1}{n+m\cdot k},\ a{=}N_{\ell+1}),\ \text{ for } 1 \leq m \leq k+1. \end{aligned} \tag{7}$$

Specifically, for any $m \in [k+1]$, we construct a sequences of sets $S_1^{(m)}, \ldots, S_m^{(m)}$, bid vectors $\mathbf{v}^{(m)}$, and real numbers $\epsilon^{(m)}$ such that they satisfy the properties listed in (Prop*-1)–(Prop*-5) below.

(Prop*-1) For all $\ell \in [m]$, $|S_\ell^{(m)}| = N_{m+1}$.

(Prop*-2) $\mathbf{v} \subseteq \mathbf{v}^{(1)} \subseteq \cdots \subseteq \mathbf{v}^{(m)}$, and $|\mathbf{v}^{(m)}| = n + (k-1)m$.

(Prop*-3) For $m \geq 2$ and $\ell \in [m-1]$,

$$S_\ell^{(m)} \subseteq S_\ell^{(m-1)} \subseteq ... \subseteq S_\ell^{(\ell)}.$$

(Prop*-4) $S_m^{(m)} \prec \cdots \prec S_1^{(m)}$. Furthermore,

$$\max(\mathbf{v}^{(m)}) \leq \min S_m^{(m)} - \epsilon^{(m)}$$

(Prop*-5) Fix any $\ell \in [m]$ and let $\mathbf{b} = (b_1, \ldots, b_\ell) \in S_1^{(m)} \times \cdots \times S_\ell^{(m)}$. Then, the following hold:

$$\widetilde{x}_{b_\ell}\left(\mathbf{v}^{(\ell)}, \mathbf{b}\right) > 0, \quad \text{and} \quad \frac{\widetilde{p}_{b_\ell}\left(\mathbf{v}^{(\ell)}, \mathbf{b}\right)}{\widetilde{x}_{b_\ell}\left(\mathbf{v}^{(\ell)}, \mathbf{b}\right)} \leq \min S_m^{(m)} - \epsilon^{(m)}.$$

We first show how the lemma follows assuming that the above properties holds for any $m \in [k+1]$ and give the induction proof afterwards. Let $m = k+1$. Consider $\mathbf{v}^{(k+1)}$ and the sequence of sets $S_1^{(k+1)}, ..., S_k^{(k+1)}$. By (Prop*-1) and Equation (7), $|S^{(\ell)}| \geq 2$ for all $\ell \in [k+1]$. From (Prop*-4), we also know that $S_1^{(k+1)} \prec ... \prec S_k^{(k+1)}$, and that for $\mathbf{b} = (b_1, \ldots, b_m) \in S_1^{(k+1)}, ..., S_k^{(k+1)}$, each b_ℓ is unique in $(\mathbf{v}^{(k+1)}, \mathbf{b})$. Furthermore, $\widetilde{x}_{b_m}(\mathbf{v}^{(k+1)}, \mathbf{b}) > 0$ by (Prop*-5). Thus, the assumptions of Lemma 4.12 hold for $\mathbf{v}^{(k+1)}$ and the sets $S_{k+1}^{(k+1)}, ..., S_1^{(k+1)}$, which means that some $\mathbf{b}^*$ exists where $|W(\mathbf{b}^*)| \geq k+1$.

Induction Proof. In the rest of the proof, we focus on the induction for proving properties (Prop*-1) - (Prop*-5). For clarity, in our induction, we use superscripts to refer to the propositions for a particular $m \in [k+1]$, i.e. $(\text{Prop*-1})^{(m)}$ represents property (Prop*-1) for m. *Base Case: $m = 1$.* Define a sequence of sets A_z for $z \in [k]$ such that

$$\{\max(\mathbf{v}) + z + 1\} \prec A_z \prec \{\max(\mathbf{v}) + z + 2\}$$

and $|A_z| = N_1$ for all $z \in [k]$. By construction $A_1 \prec \cdots \prec A_k$. We also define $\phi(\cdot)$ as follows:

$$\phi\left(\widehat{\mathbf{b}}\right) := \min\Big\{i : x_i(\widehat{\mathbf{b}}) > 0\Big\}. \tag{8}$$

Consider the bid vector $(\mathbf{v}, \mathbf{v}')$ for $\mathbf{v}' \in A_1 \times \cdots \times A_{k+1}$. By Lemma 4.7, $\mu(\mathbf{v}, \mathbf{v}') \geq \mu(\mathbf{v}) > 0$, and therefore, by budget feasibility, there exists an index i such that $x_i(\mathbf{v}, \mathbf{v}') > 0$. Thus, $\phi(\mathbf{v}, \mathbf{v}')$ is well defined. The $n+k$ possible values of $\phi(\mathbf{v}, \mathbf{v}')$ naturally give the following partition of $A_1 \times \cdots \times A_k$:

$$T_i := \Big\{\mathbf{v}' \in A_1 \times \cdots \times A_k : \phi(\mathbf{v}, \mathbf{v}') = i\Big\}$$

for $i \in [n+k]$.

Since this is a partition of k sets $A_1, \ldots, A_k$ all of size N_1, there must exist a smallest index i^* such that $|T_{i^*}| \geq \frac{1}{n+k} \cdot N_1^k$. By Lemma 4.20 and the definition of N_1 and N_2, we know that there exists subsets $B_z \subseteq A_z$ for all $z \in [k]$ with $|B_z| = N_2$ such that $B_1 \times \cdots \times B_k \subseteq T_{i^*}$.

Claim 4.21. *$i^* > n$.*

Proof. Suppose for the sake of contradiction that $i^* \leq n$. Let $V = (v_{i^*}, \inf B_1) \setminus \{v_1, \ldots, v_n\}$, since $\inf B_1 \geq \inf A_1 > \max(\mathbf{v})$, this set has infinite elements. From Myerson's Lemma (Lemma 3.5), for $\mathbf{v}' \in B_1 \times \cdots \times B_k$ and $v^* \in V$, $x_{i^*}((\mathbf{v}_{-i^*}, v^*), \mathbf{v}') \geq \widetilde{x}_{i^*}(\mathbf{v}, \mathbf{v}') > 0$ where the last step comes from the definition of i^*. Thus, by construction, the hypothesis of Lemma 4.12 is satisfied with respect to $\mathbf{v}_{-i^*}$ and the sets $V, B_1, \ldots, B_k$, and we can conclude that there exists $\mathbf{b}^*$ such that $|W(\mathbf{b}^*)| \geq k+1$, which contradicts Lemma 4.11. □

Thus, we know that for every bid vector $(\mathbf{v}, \widehat{\mathbf{v}})$ where $\widehat{\mathbf{v}} \in T_{i^*}$, we have $x_{i^*}(\mathbf{v}, \widehat{\mathbf{v}}) > 0$, so the bid falls in $B_{i^*-n} \subseteq A_{i^*-n}$ has a positive confirmation probability. Pick an arbitrary $\widehat{\mathbf{v}} \in T_{i^*}$. Let $v^* = \max(\mathbf{v}) + k + 3 > \widehat{v}_{i^*-n}$. Consider a bid vector $\mathbf{v}^* = ((\mathbf{v}, \widehat{\mathbf{v}})_{-i^*}, v^*)$. Since $x_{i^*}(\mathbf{v}, \widehat{\mathbf{v}}) > 0$, by Lemma 4.5, we know that $\mathsf{util}_{i^*}(\mathbf{v}^*) > \mathsf{util}_{i^*}(\mathbf{v}, \widehat{\mathbf{v}}) \geq 0$, and thus $\tau := \frac{p_i(\mathbf{v}^*)}{x_i(\mathbf{v}^*)} < v^*$. We define the following:

- $\mathbf{v}^{(1)} := (\mathbf{v}, \widehat{\mathbf{v}})_{-i^*}$;
- $\epsilon^{(1)} := \frac{1}{2}(v^* - \max(\mathbf{v}^{(1)}, \tau)) > 0$;
- $S_1^{(1)}$ s.t. $\big\{v^* - \epsilon^{(1)}\big\} \prec S_1^{(1)} \prec \{v^*\}$ and $|S_1^{(1)}| = N_2$.

We now prove the properties:

- $(\text{Prop*-1})^{(1)}$ and $(\text{Prop*-2})^{(1)}$ follow by construction.
- $(\text{Prop*-3})^{(1)}$ is trivially satisfied.
- $(\text{Prop*-4})^{(1)}$ follows from construction and the fact that

$$\begin{aligned}\min S_1^{(1)} - \epsilon^{(1)} &\geq v^* - (v^* - \max(\mathbf{v}^{(1)}, \tau)) \\ &= \max(\mathbf{v}^{(1)}, \tau) \\ &\geq \max(\mathbf{v}^{(1)}).\end{aligned}$$

- (Prop*-5)$^{(m)}$, note that for any $b_1 \in S_1^{(1)}$, and $b' \in B_{i^*-n}$, we know that $\min S_1^{(1)} \geq v^* - \epsilon^{(1)} \geq \max(\mathbf{v}^{(1)}) \geq \max(B_{i^*-n})$, so

$$\begin{aligned} \widetilde{x}_{b_1}(\mathbf{v}^{(1)}, b_1) &\geq \widetilde{x}_{v_{i^*}}(\mathbf{v}^{(1)}, b') && \text{by Myerson's Lemma (Lemma 3.5)} \\ &> 0. && \text{by definition of } i^* \end{aligned}$$

Furthermore,

$$\begin{aligned} \frac{\widetilde{p}_{b_1}(\mathbf{v}^{(1)}, b_1)}{\widetilde{x}_{b_1}(\mathbf{v}^{(1)}, b_1)} &\leq \frac{\widetilde{p}_{v^*}(\mathbf{v}^{(1)}, v^*)}{\widetilde{x}_{v^*}(\mathbf{v}^{(1)}, v^*)} && \text{by Lemma 4.6} \\ &= \tau \\ &\leq \max(\mathbf{v}^{(1)}, \tau) \\ &\leq \min S_1^{(1)} - \epsilon^{(1)}. \end{aligned}$$

Thus, (Prop*-5)$^{(m)}$ is satisfied.

Inductive Step: $m > 1$. Suppose $S_1^{(m-1)}, \dots, S_{m-1}^{(m-1)}$, $\mathbf{v}^{(1)}, \dots, \mathbf{v}^{(m-1)}$, and $\epsilon^{(m-1)}$ are the sequences of sets, bid vectors, and the real number respectively that satisfy (Prop*-1)$^{(m-1)}$-(Prop*-5)$^{(m-1)}$. For simplicity, let $\gamma := \min S_{m-1}^{(m-1)}$, and

$$\alpha := \gamma - \epsilon^{(m-1)}, \quad \Delta := \frac{\epsilon^{(m-1)}}{k+2}. \tag{9}$$

Then for $z \in [k]$, redefine A_z such that

$$\{\alpha + z \cdot \Delta\} \prec A_z \prec \{\alpha + (z+1) \cdot \Delta\} \tag{10}$$

and $|A_z| = N_m$ for $z \in [k]$. By construction, we have $A_1 \prec \cdots \prec A_k$.

Consider the bid vector $\mathbf{v}^* = (\mathbf{v}^{(m-1)}, \mathbf{v}')$ where $\mathbf{v}' \in \left(\prod_{z=1}^k A_z\right) \times \left(\prod_{\ell=1}^{m-1} S_\ell^{(m-1)}\right)$. Then $|\mathbf{v}^*| = n + k \cdot m$. Recall the definition of $\phi(\cdot)$ from Equation (8). Again, by Lemma 4.7 and (Prop*-2)$^{(m-1)}$, $\mu(\mathbf{v}^*) \geq \mu(\mathbf{v}) > 0$. Therefore, by budget feasibility, there exists $i \in [n + k \cdot m]$ such that $x_i(\mathbf{v}^*) > 0$, and thus, $\phi(\mathbf{v}^*)$ is well defined. The $n + m \cdot k$ possible values of $\phi(\mathbf{v}^*)$ naturally give the following partition of $\left(\prod_{z=1}^k A_z\right) \times \left(\prod_{\ell=1}^{m-1} S_\ell^{(m-1)}\right)$.

$$T_i := \left\{ \mathbf{v}' \in \left(\prod_{z=1}^k A_z\right) \times \left(\prod_{\ell=1}^{m-1} S_\ell^{(m-1)}\right) : \phi(\mathbf{v}^{(m-1)}, \mathbf{v}') = i \right\}$$

for $i \in [n + k \cdot m]$. The sequence $\{T_i\}_{i \in [n+k \cdot m]}$ is a partition of the product of $k + m - 1$ sets $A_1, \dots, A_k, S_1, \dots S_{m-1}$, each of size N_m by (Prop*-1)$^{(m-1)}$. Thus, by the pigeon hole principle, let i^* be the smallest index such that $|T_{i^*}| \geq \frac{1}{n+m \cdot k} \cdot N_m^{k+m-1}$. By Lemma 4.20 and the definitions of N_m and N_{m+1}, there exist subsets

$$B_z \subseteq A_z \text{ for all } z \in [k], \quad \text{and} \quad B_z \subseteq S_{z-k}^{(m-1)} \text{ for all } z \in \{k+1, \dots, k+m-1\},$$

each satisfying $|B_z| = N_{m+1}$, such that

$$B_1 \times \cdots \times B_{k+m-1} \subseteq T_{i^*}.$$

Note that in any vector $\mathbf{v}^* = (\mathbf{v}^{(m-1)}, \mathbf{v}')$ where $\mathbf{v}' \in \left(\prod_{z=1}^k A_z\right) \times \left(\prod_{\ell=1}^{m-1} S_\ell^{(m-1)}\right)$, the indices can be understood as follows:

- The bids $\mathbf{v}_i^*$ for $1 \le i \le n + (k-1)(m-1)$ come from $\mathbf{v}^{(m-1)}$: $\mathbf{v}_i^* = \mathbf{v}_i^{(m-1)}$.
- The bids $\mathbf{v}_i^*$ for $n + (k-1) \cdot (m-1) + 1 \le i \le n + (k-1) \cdot (m-1) + k$ come from $A_1 \times \cdots \times A_k$: $\mathbf{v}_i^* \in A_{i-n-(k-1)(m-1)}$.
- The bids $\mathbf{v}_i^*$ for $n + (k-1) \cdot (m-1) + k + 1 \le i \le n + k \cdot m$ come from $S_1^{(m-1)} \times \cdots \times S_{m-1}^{(m-1)}$: $\mathbf{v}_i^* \in S_{i-n-(k-1)(m-1)-k}^{(m-1)}$.

Let $L := n + (k-1) \cdot (m-1)$ and $R := n + (k-1) \cdot (m-1) + k$ be the left and right boundary indices of bids that come from A_1 through A_k.

Claim 4.22. *$L < i^*$.*

Proof. Suppose for the sake of contradiction that $i^* \le L$. Let $\mathbf{b} = (b_1, \dots, b_{m-1}) \in B_{k+1} \times \cdots \times B_{k+m-1}$, and let $V = (\mathbf{v}_{i^*}^{(m-1)}, \inf B_1) \setminus \left\{\mathbf{v}_1^{(m-1)}, \dots, \mathbf{v}_L^{m-1}, b_1, \dots, b_{m-1}\right\}$, since $\inf B_1 \ge \inf A_1 > \max(\mathbf{v}^{(m-1)})$, this set has infinite elements. Similarly to the base case, by construction, the hypothesis of Lemma 4.12 is satisfied w.r.t $(\mathbf{v}_{-i^*}^{(m-1)}, b_1, \dots, b_{m-1})$ and the sets $V, B_1, \dots, B_k$. Thus, we can conclude that there exists $\mathbf{b}^*$ such that $|W(\mathbf{b}^*)| > k$, which contradicts Lemma 4.11. □

Claim 4.23. *$i^* \le R$.*

Proof. Suppose, for the sake of contradiction, that $i^* > R$. Let $j^* = i^* - R$ be the index of the bid with positive confirmation probability that comes from set $S_{j^*}^{(m-1)}$. Pick an arbitrary $\mathbf{v}' = (a_1, \dots, a_k, b_1, \dots, b_{m-1}) \in \prod_{i=1}^{k+m-1} \mathsf{rmInf}(B_i)$. Recall that by the definition of T_{i^*} and $\phi(\cdot)$ (see Equation (8)), for any $\widehat{\mathbf{v}} \in \left(\prod_{i=1}^{k+m-1} B_i\right) \subseteq T_{i^*}$, we have:

$$x_{i^*}\left(\mathbf{v}^{(m-1)}, \widehat{\mathbf{v}}\right) > 0.$$

Then, by Lemma 4.5, user i^* in the honest bid vector $(\mathbf{v}, \mathbf{v}')$ has positive utility, since $b_{j^*} \neq \min B_{i^*-n}$.

Consider a world where $(\mathbf{v}^{(m-1)}, \mathbf{v}')$ is the honest bid vector. Let i^* be the user with valuation b_{j^*} and u be the user with valuation a_k. In the honest case, user i^* gets a positive utility by our choice of b_{j^*}, while user u has zero probability of confirmation by definition of $\phi(\cdot)$ (Equation (8)).

Now, consider a coalition consisting of the miner and user u. Their honest joint utility is simply the miner revenue $\mu(\mathbf{v}^{(m-1)}, \mathbf{v}')$. The coalition could perform the following strategy:

- The miner censors the following bids, pretending it has not received these bids from users:
  - b_{j^*} from user i^*;
  - all bids $a_1, \dots, a_{k-1}$;
  - all bids in $\mathbf{v}^{(m-1)}$ that are not in $\mathbf{v}^{(j^*)}$.
  - all bids b_ℓ for $\ell \in \{j^* + 1, \dots, m-1\}$.
- User u raises their bid to b_{j^*}. The miner then performs the honest inclusion rule on the resulting bid vector $\mathbf{v}^* = (\mathbf{v}^{(j^*)}, b_1, ..., b_{j^*})$.

So, the strategy removes the bids at indices $n + (k-1) \cdot j^* + 1$ through $i^* - 1$ in $(\mathbf{v}^{(m-1)}, \mathbf{v}')$. Recall that by definition of $\phi(\cdot)$ (Equation (8)), we have

$$x_i(\mathbf{v}^{(m-1)}, \mathbf{v}') = 0 \text{ for } 1 \le i \le i^* - 1. \tag{11}$$

Thus, all of indices of the bids that were censored in the above strategy had 0 confirmation probability. Therefore, the expected miner revenue under the strategy is at least

$$\begin{aligned} \mu(\mathbf{v}^*) &\geq \mu(\mathbf{v}^{(m-1)}, \mathbf{v}') + \sum_{i=n+(k+1)\cdot j^*+1}^{i^*-1} p_i(\mathbf{v}^{(m-1)}, \mathbf{v}') && \text{by Lemma 4.8} \\ &= \mu(\mathbf{v}^{(m-1)}, \mathbf{v}'). \end{aligned}$$

Additionally, as we are only censoring bids and replacing one of them, $\mathbf{v}^* \subseteq (\mathbf{v}^{(m-1)}, \mathbf{v}')$. Thus, by Lemma 4.7, $\mu(\mathbf{v}^*) \leq \mu(\mathbf{v}^{(m-1)}, \mathbf{v}')$. Therefore, we have $\mu(\mathbf{v}^*) = \mu(\mathbf{v}^{(m-1)}, \mathbf{v}')$.

Meanwhile, since $(b_1, ..., b_{j^*}) \in S_1^{(m-1)} \times \ldots S_{j^*}^{(m-1)}$, notice how $\mathbf{v}^*$ matches the form of $(\text{Prop}^*\text{-5})^{(m-1)}$ for $\ell = j^*$. This means user u's utility in $\mathbf{v}^*$ under the above strategy is

$$\begin{aligned} & a_k \cdot \widetilde{x}_{\mathbf{b}_{j^*}}(\mathbf{v}^*) - \widetilde{p}_{\mathbf{b}_{j^*}}(\mathbf{v}^*) \\ &> (\gamma - \epsilon^{(m-1)}) \cdot \widetilde{x}_{\mathbf{b}_{j^*}}(\mathbf{v}^*) - \widetilde{p}_{\mathbf{b}_{j^*}}(\mathbf{v}^*) && \text{by Equation (10)} \\ &\geq \frac{\widetilde{p}_{\mathbf{b}_{j^*}}(\mathbf{v}^*)}{\widetilde{x}_{\mathbf{b}_{j^*}}(\mathbf{v}^*)} \cdot \widetilde{x}_{\mathbf{b}_{j^*}}(\mathbf{v}^*) - \widetilde{p}_{\mathbf{b}_{j^*}}(\mathbf{v}^*) && \text{by } (\text{Prop}^*\text{-5})^{(m-1)} \\ &= 0. \end{aligned}$$

Thus, u has positive utility under the described strategy. Therefore, the coalition's joint utility under the strategy strictly increases while harming user i^*, who had a positive utility in the honest case. This contradicts 1-CRHP. □

By Claim 4.22 and Claim 4.23, for every $(\mathbf{v}^{(m-1)}, \mathbf{v}')$ where $\mathbf{v}' \in T_{i^*}$, we know that a bid with positive confirmation probability exists within the set $B_{i^*-L} \subseteq A_{i^*-L}$. Pick an arbitrary $\widehat{\mathbf{v}} \in B_1 \times \cdots \times B_k$, and define $\mathbf{v}^{(m)} = \left(\mathbf{v}^{m-1)}, \widehat{\mathbf{v}}\right)_{-i^*}$. Recall that $\gamma := \min S_{m-1}^{(m-1)}$. Consider the value

$$\delta := \min_{\mathbf{b} \in S_1^{(m-1)} \times \cdots \times S_{m-1}^{(m-1)}} \left(\gamma - \frac{p_R(\mathbf{v}^{(m)}, \gamma, \mathbf{b})}{x_R(\mathbf{v}^{(m)}, \gamma, \mathbf{b})} \right). \tag{12}$$

Let γ' be defined such that $\max(\mathbf{v}^{(m)}) < \gamma' < 0$ and let $\widehat{\mathbf{b}} = ((\mathbf{v}^{(m)}, \hat{\mathbf{v}})_{-i^*}, \gamma'), \mathbf{b})$. We know that γ' is unique in $(\mathbf{v}^{(m)}, \gamma', \mathbf{b})$ and unique in $\widehat{\mathbf{b}}$. By the definition of i^* and Myerson's Lemma (Lemma 3.5), $x_{i^*}(\widehat{\mathbf{b}}) > 0$. Then, by weak symmetry, we know that $\mathsf{util}_R(\mathbf{v}^{(m-1)}, \gamma', \mathbf{b}) = \mathsf{util}_{i^*}(\widehat{\mathbf{b}}) > 0$.

Finally, by Lemma 4.5, $\mathsf{util}_R(\mathbf{v}^{(m-1)}, \gamma, \mathbf{b}) > \mathsf{util}_R(\mathbf{v}^{(m-1)}, \gamma', \mathbf{b}) > 0$, implying that $\frac{p_R(\mathbf{v}^{(m-1)}, \gamma, \mathbf{b})}{x_R(\mathbf{v}^{(m-1)}, \gamma, \mathbf{b})} < \gamma$ for all $\mathbf{b} \in S_1^{(m-1)} \times \cdots \times S_{m-1}^{(m-1)}$. Since there is only a finite number of possible $\mathbf{b}$ by $(\text{Prop}'\text{-1})^{(m-1)}$, we have $\delta > 0$. Thus, we can define the following:

- $\mathbf{v}^{(m)} = \left(\mathbf{v}^{m-1)}, \widehat{\mathbf{v}}\right)_{-i^*}$;
- $S_1^{(m)}, \ldots, S_{m-1}^{(m)} = B_{k+1}, \ldots, B_{k+m-1}$;
- $\epsilon^{(m)} := \frac{1}{2} \min\left(\delta, \epsilon^{(m-1)}\right)$;
- $S_m^{(m)}$ s.t. $\left\{\gamma - \epsilon^{(m)}\right\} \prec S_m^{(m)} \prec \{\gamma\}$ and $|S_m^{(m)}| = N_{m+1}$.

We now prove the properties:

- $(\text{Prop}^*\text{-1})^{(m)}$, $(\text{Prop}^*\text{-2})^{(m)}$, and $(\text{Prop}^*\text{-3})^{(m)}$ are satisfied by construction.

- (Prop*-4)$^{(m)}$ follows from construction and the fact that

$$\begin{aligned}
\min S_m^{(m)} - \epsilon^{(m)} &= \gamma - \min(\epsilon^{(m-1)}, \delta) \\
&\geq \gamma - \epsilon^{(m-1)} \\
&\geq \gamma - \Delta && \text{by Equation (9)} \\
&\geq \max A_k && \text{by Equation (10)} \\
&\geq \max\,(\mathbf{v}^{(m)}).
\end{aligned}$$

- By (Prop*-3)$^{(m)}$, every set in (Prop*-5)$^{(m)}$ is contained in the corresponding set from (Prop*-5)$^{(m-1)}$. Therefore, for each $\ell \in [m-1]$, (Prop*-5)$^{(m)}$ follows from (Prop*-5)$^{(m-1)}$.

  For $\ell = m$, let $\mathbf{b} = (b_1, \cdots b_m) \in S_1^{(m)} \times \cdots \times S_m^{(m)}$, and $v_{i^*} \in B_{i^*-L}$. Note that for any $b_m \in S_m^{(m)}$, and $b' \in B_{i^*-L}$, we know that $b_m \geq \min S_m^{(m)} \geq \gamma - \epsilon^{(m)} \geq \max(\mathbf{v}^{(m)}) \geq \max(B_{i^*-L})$. Then

$$\begin{aligned}
\widetilde{x}_{b_m}(\mathbf{v}^{(m)}, \mathbf{b}) &\geq \widetilde{x}_{v_{i^*}}(\mathbf{v}^{(m)}, \mathbf{b}_{-m}, v_{i^*}) && \text{by Myerson's Lemma (Lemma 3.5)} \\
&> 0. && \text{by definition of } i^*
\end{aligned}$$

  Furthermore,

$$\begin{aligned}
\frac{\widetilde{p}_{b_m}(\mathbf{v}^m, \mathbf{b})}{\widetilde{x}_{b_m}(\mathbf{v}^{(m)}, \mathbf{b})} &\leq \frac{p_{n+m\cdot k}(\mathbf{v}^{(m)}, \mathbf{b}_{-m}, \gamma)}{x_{n+m\cdot k}(\mathbf{v}^{(m)}, \mathbf{b}_{-m}, \gamma)} && \text{by Lemma 4.6} \\
&\leq \gamma - \delta \\
&= \left(\gamma - \frac{1}{2}\delta\right) - \frac{1}{2}\delta \\
&\geq \left(\gamma - \frac{1}{2}\min\left(\delta, \epsilon^{(m-1)}\right)\right) - \frac{1}{2}\min\left(\delta, \epsilon^{(m-1)}\right) \\
&= \min S_m^{(m)} - \epsilon^{(m)}.
\end{aligned}$$

  Thus, (Prop*-5)$^{(m)}$ is satisfied.

Corollary 4.24. *Only trivial TFMs, where no bids are ever confirmed, satisfy UIC, MIC, MRHP, and weak* 1*-CRHP in the plain model.*

Proof. By Theorem 4.19, a TFM that satisfies UIC, MIC, and 1-CRHP in the plain model must have zero miner revenue. Suppose for the sake of contradiction that for some bid vector $\mathbf{b}$, $x_i(\mathbf{b}) > 0$. By Lemma 4.5, this implies a bid vector $\mathbf{b}'$ exists where $\mathsf{util}_i(\mathbf{b}') > \mathsf{util}_i(\mathbf{b}) \geq 0$. Since the miner revenue is always 0, in a world where $\mathbf{b}'$ is the honest bid vector, the miner can ignore all bids to harm user i at no cost, contradicting MRHP. □

5 MPC-assisted Model

In this section, we present the feasibility and impossibility results in the MPC-assisted model. Recall that in the MPC-assisted model, a committee of M miners jointly run a multi-party computation (MPC) to implement the TFM. Cryptography guarantees that as long as there is one honest miner, either the protocol succeeds and returns the correct outcome based on all received bids, or the protocol aborts.

5.1 Feasibility in the MPC-assisted Model

In this section, we present a mechanism that achieves all desired properties, UIC, MIC, all RHP variants, and positive revenue in the MPC-assisted model.

MPC-assisted, posted price with random selection //Reserve price $\mathsf{res} > 0$.
//Revenue parameter $0 < \epsilon \leq \mathsf{res}$.

- **Inclusion & Confirmation**: Let S be the set of bids strictly above the reserve price res. If $|S| \leq k$, include and confirm all bids in S. Otherwise, randomly choose k bids in S to include and confirm.
- **Payment & Revenue**: Each confirmed bid pays the reserve price res, and the miners jointly receive a revenue of ϵ per confirmed bid.

Theorem 5.1. *The above mechanism satisfies UIC, MIC, URHP, MRHP, and d-CRHP for any $d \geq 1$ in the MPC-assisted model. Additionally, the miner revenue is positive.*

Proof. We have previously shown in the plain model by Lemma 4.3 that the mechanisms satisfies UIC, MIC, positive miner revenue, and URHP. These results still hold in the MPC-assisted model. The rest of the proof focuses on MRHP and CRHP.

d-CRHP: To harm an honest user $i \notin \mathcal{C}$, a coalition $\mathcal{C}$ of at most $M-1$ miners and d users can only fail the auction or reduce the probability of user i getting confirmed by injecting fake bids. Let m denote the number of miners in the coalition. Since otherwise, the auction is guaranteed to be honestly implemented based on the submitted bids, as guaranteed in the MPC-assisted model. If a user $i \notin \mathcal{C}$ is harmed, it implies that in the honest case, user i must get confirmed and pays res. The coalition's joint utility is at least the miner revenue from user i, which is $\epsilon \cdot m/M$ in the honest case. Thus, failing the auction harm the coalition themselves.

Now suppose that the coalition injects fake bids to harm honest user i. Let n_H and n_C be the number of honest users and colluding users, respectively, whose true values are above res, and let V be the sum of all the true values of colluding users in $\mathcal{C}$ that are above res. Then in the honest case, the coalition's joint utility is $u_H := \min\{k/(n), 1\}\,(V - n_C \cdot \mathsf{res}) + \min\{k, n\}\epsilon \cdot m/M$. Let f be the number of fake bids that are above res. If $f = 0$, this harms no honest player. Therefore, we assume that $f > 0$. Now each bid gets confirmed with a smaller probability of $\frac{k}{n+f}$, and the coalition needs to pay res for the fake bid when the fake bids gets confirmed. Thus, now the coalition's expected joint utility becomes $\min\{k/(n+f), 1\}\,(V - n_C \cdot \mathsf{res}) - \min\{k/(n+f), 1\} \cdot f \cdot \mathsf{res} + \min\{k, n+f\}\epsilon \cdot m/M$, which is strictly smaller than the honest expected utility u_H.

The proof for MRHP is the same as the above argument by replacing V and n_C to be zero. □

Not 1-SCP: By the same reasoning in Section 4.1.3. Imagine that there are less than k users with true value above res in the honest case and the colluding user's true value is $\mathsf{res} - \epsilon \cdot m/2M$, where m is the number of colluding miners in the coalition. The colluding user can just raise its bid to $\mathsf{res} + \delta$ for some $\delta > 0$. This increase the coalition's joint utility by $\epsilon \cdot m/2M > 0$.

5.2 Characterization of Deterministic Mechanisms

5.2.1 Deterministic Feasibility in the MPC-Assisted Model

While the previous mechanism achieves all desired properties, it relies on randomness to choose the set of k bids to confirm. In this section, we show that the all-or-nothing posted price is a deterministic mechanism satisfies all desired properties in the MPC-assisted model.

MPC-assisted, all or nothing posted price auction. //Reserve price $\mathsf{res} > 0$.
//Revenue parameter $0 < \epsilon \leq \mathsf{res}$.

Mechanism:

- **Inclusion Rule & Confirmation Rule:** If there are $\leq k$ bids and each bid is strictly greater than res, include and confirm all bids. Otherwise, include and confirm no bids.
- **Payment Rule & Miner Revenue Rule:** Each confirmed bid pays the reserve price res, and the miner receives a revenue share of ϵ per confirmed bid.

Lemma 5.2. *The above TFM satisfies UIC, MIC, URHP, MRHP d-CRHP for any $d \geq 1$, and has positive miner revenue in the MPC-assisted model.*

Proof. By Lemma 4.2, this TFM satisfies UIC, d-CRHP for all $d \geq 1$, URHP, and MRHP in the plain model. These results translate to the MPC-assisted model. The rest of proof focuses on MIC.

MIC: Since this is a deterministic mechanism, the miners can only inject fake bids to increase their revenue. Failing the auction cannot benefit any coalition of miners. Since each confirmed bid pays the same amount, the only way to increase miner revenue is to increase the number of users that are confirmed.

However, in the honest case, if any user gets confirmed, all users get confirmed. There is no method for the miner coalition to increase the number of users to confirm by injecting fake bids. □

Not 1-SCP: Imagine a world with k users, where $k-1$ of them have true value above res, but one has a true value $(\mathsf{res} - \frac{\epsilon}{2M})$. Consider the coalition consisting of one miner and this user with true value below res. In the honest case, no one gets confirmed, so this coalition gets a joint utility of 0. However, if the user raises their bid to $\mathsf{res} + 1$, the joint utility now becomes $k \cdot \frac{\epsilon}{M} + (\mathsf{res} - \frac{\epsilon}{2M}) - \mathsf{res} > 0$.

5.3 Deterministic Impossibility in the MPC-Assisted Model

The above all-or-nothing posted price is degenerate in the sense that it confirms no bid when there are more than k users. Unfortunately, this is necessary for deterministic mechanisms, even in the MPC-assisted model.

Lemma 5.3. *Suppose a deterministic TFM $(\mathbf{I}, \mathbf{C}, \mathbf{P}, \mathbf{R})$ satisfies UIC and URHP in the MPC-assisted model. For any bid vector $\mathbf{b} = (b_1, \ldots, b_n)$, if there exists some $b_i \in \mathbf{b}$ such that $\mathsf{util}_i(\mathbf{b}) > 0$, then for every j where b_j is unique in $\mathbf{b}$ and $b_j > b_i$, it must hold that $x_j(\mathbf{b}) = 1$.*

Proof. For the sake of contradiction, suppose that there exists $\mathbf{b}$ where $\mathsf{util}_i(\mathbf{b}) > 0$ and $x_j(\mathbf{b}) = 0$ for some unique $b_j > b_i$. Consider the bid vector $\mathbf{b}' = (b_1, \ldots, b_{i-1}, b_j, b_i + 1, \ldots, b_{j-1}, b_i, b_j + 1, \ldots, b_n)$ where the user i and user j swap bid values. By weak symmetry, the bid at b_j must now be unconfirmed, i.e. $x_i(\mathbf{b}') = 0$, since $\mathbf{b}'$ and $\mathbf{b}$ only differ by metadata and the bid at b_j is unique.

However, consider the world where $\mathbf{b}$ is the honest bid vector and the intermediate bid vector $\mathbf{b}^* = (\mathbf{b}_{-j}, b_i)$ is achieved by the strategy where user j underbids to b_i. By Lemma 4.5, $\mathsf{util}_j(\mathbf{b}^*) \leq \mathsf{util}_j(\mathbf{b}) = 0$. Consequently, by URHP, $\mathsf{util}_i(\mathbf{b}^*) \geq \mathsf{util}_i(\mathbf{b}) > 0$ which implies $x_i(\mathbf{b}^*) = 1$ in a deterministic TFM. Now, by Myerson's Lemma (Lemma 3.5), when user i raises their bid to b_j to transform $\mathbf{b}^*$ into $\mathbf{b}'$, $x_i(\mathbf{b}') = 1$, which is a contradiction. □

Lemma 5.4. *Suppose a deterministic TFM $(\mathbf{I}, \mathbf{C}, \mathbf{P}, \mathbf{R})$ satisfies UIC and URHP in the MPC-assisted model. For any bid vector $\mathbf{b} = (b_1, \ldots, b_n)$, if there exists some $b_i \in \mathbf{b}$ such that $\mathsf{util}_i(\mathbf{b}) > 0$, then for every j where b_j is unique and $b_j > p_i(\mathbf{b})$, it must hold that $x_j(\mathbf{b}) = 1$.*

Proof. For the sake of contradiction, suppose that there exists $\mathbf{b}$ where $\mathsf{util}_i(\mathbf{b}) > 0$ and $x_j(\mathbf{b}) = 0$ for some unique $b_j > p_i(\mathbf{b})$. We then define $\mathbf{b}' = (\mathbf{b}_{-j}, b'_j)$ where $b'_j \in (p_i(\mathbf{b}), b_j) \setminus \{b_z : z \in [n]\}$. Note that b'_j is unique in $\mathbf{b}'$. By Lemma 4.5, $\mathsf{util}_j(\mathbf{b}') \leq \mathsf{util}_j(\mathbf{b}) = 0$. In a world where $\mathbf{b}$ is the honest bid vector, user j can adopt the strategy to underbid to b'_j to achieve the realized bid vector $\mathbf{b}'$. Thus, by URHP, $\mathsf{util}_i(\mathbf{b}') \geq \mathsf{util}_i(\mathbf{b}) > 0$ which implies $p_i(\mathbf{b}') \leq p_i(\mathbf{b})$ in a deterministic TFM.

Claim 5.5. *For $v \in (p_i(\mathbf{b}), b'_j)$, $\mathsf{util}_j(\mathbf{b}_{-i}, v) > 0$.*

Proof. It is sufficient to prove that $x_j(\mathbf{b}'_{-i}, v) = 1$ as *Lemma 4.5* would imply that $\mathsf{util}_j(\mathbf{b}_{-i}, v) > \mathsf{util}_j(\mathbf{b}'_{-i}, v) \geq 0$. By Myerson's Lemma (Lemma 3.5), $x_i(\mathbf{b}'_{-i}, v) = 1$ since $v > p_i(\mathbf{b}) \geq p_i(\mathbf{b}')$. Furthermore, $v < b'_j$, so by Lemma 5.3, $x_j(\mathbf{b}'_{-i}, v) = 1$. □

Consider the world where $(\mathbf{b}_{-i}, v)$ is the honest bid vector where $v \in (p_i(\mathbf{b}), b'_j)$. By the above claim, we know that $\mathsf{util}_j(\mathbf{b}_{-i}, v) > 0$. Now consider the strategy where user i overbids to b_i to achieve the realized bid vector $\mathbf{b}$. Since the mechanism is deterministic and $v > p_i(\mathbf{b})$, Myerson's Lemma (Lemma 3.5) implies $\mathsf{util}_i(\mathbf{b}_{-i}, v) = \mathsf{util}_i(\mathbf{b})$. However, $\mathsf{util}_j(\mathbf{b}) = 0$ by assumption, contradicting URHP since user j is hurt by the deviation. □

Theorem 5.6. *Let k denote the block size. If a deterministic TFM $(\mathbf{I}, \mathbf{C}, \mathbf{P}, \mathbf{R})$ satisfies UIC and URHP in the MPC-assisted model, for every bid vector $\mathbf{v}$ where $|\mathbf{v}| \geq k+1$, no bids can be confirmed.*

Proof. Suppose, for the sake of contradiction, that there exist bid vector $\mathbf{v} = (v_1, \ldots, v_{k+1})$ where $x_i(\mathbf{v}) = 1$ for some $i^* \in [k+1]$. We define the sequence of bid vectors $\mathbf{v}^{(m)}$ below.

$$\begin{aligned} b_1 &= \max(\mathbf{v}) + 1, & \mathbf{v}^{(1)} &= (\mathbf{v}_{-i^*}, b_1), \\ b_m &= \frac{1}{2}(v_{m-1} + \max(\mathbf{v})), & \mathbf{v}^{(m)} &= \left(\mathbf{v}^{(m-1)}, b_m\right) \quad \text{for } m > 1. \end{aligned}$$

Note that b_z is unique in $\mathbf{v}^{(m)}$ for all $z \in [m]$. The proof follows from the following key technical lemma. The proof of Lemma 5.7 will be given right afterwards.

Lemma 5.7. *For all $m \in \mathbb{N}$, $\widetilde{x}_{b_m}(\mathbf{v}^{(m)}) = 1$ and $\widetilde{p}_{b_m}(\mathbf{v}^{(m)}) \leq \max(\mathbf{v})$.*

Assuming the lemma holds, consider $\mathbf{v}^{(k+1)}$. We know by Lemma 5.7, $\widetilde{x}_{b_{k+1}}(\mathbf{v}^{(k+1)}) = 1$ and that $\widetilde{p}_{b_{k+1}}(\mathbf{v}^{(k+1)}) \leq \max(\mathbf{v})$. By definition,

$$\max(\mathbf{v}) < b_{k+1} < \cdots < b_1.$$

Meaning that $\mathsf{util}_{b_{k+1}}(\mathbf{v}^{(k+1)}) > 0$. Thus, by Lemma 5.3, $\widetilde{x}_{b_z}(\mathbf{v}^{(k+1)}) = 1$ for all $z \in [k+1]$. This contradicts the block size limit k. □

Proof of Lemma 5.7. We continue via induction on m.

Base Case: $m = 1$. For $\mathbf{v}^{(1)} = (\mathbf{v}_{-i^*}, b_1)$. The conclusion for base case directly follow from UIC. Specifically, by Myerson's Lemma (Lemma 3.5), we have $\widetilde{x}_{b_1}(\mathbf{v}^{(1)}) = 1$, since $b_1 > v_{i^*}$. Furthermore, we have $\widetilde{p}_{b_1}(\mathbf{v}^{(1)}) \leq v_{i^*} \leq \max(\mathbf{v})$.

Inductive Step: $m > 1$. It suffices to prove that $\widetilde{x}_{b^*}(\mathbf{v}^{(m-1)}, b^*) = 1$ for all $b^* \in (\max(\mathbf{v}), b_{m-1})$ since $b_m \in (\max(\mathbf{v}), b_{m-1})$ and Myerson's Lemma (Lemma 3.5) would then imply $\widetilde{p}_{b_1}(\mathbf{v}^{(1)}) \leq \max(\mathbf{v})$. Note that b^* is unique in $(\mathbf{v}^{(m-1)}, b_{m-1})$.

Claim 5.8. $\widetilde{x}_{b^*}(\mathbf{v}^{(m-1)}, b^*) = 1$ *for all* $b^* \in (\max(\mathbf{v}), b_{m-1})$.

Proof. Suppose that there exists $b' \in (\max(\mathbf{v}), b_{m-1})$ where $\widetilde{x}_{b'}(\mathbf{v}^{(m-1)}, b') = 0$. By the contrapositive of Lemma 5.4, this would imply $\widetilde{p}_{b_{m-1}}(\mathbf{v}^{(m-1)}, b') > b'$ or $\widetilde{x}_{b_{m-1}}(\mathbf{v}^{(m-1)}, b') = 0$. In either case, consider the world where $(\mathbf{v}^{(m-1)})$ is the honest bid vector. By the inductive hypothesis, $\widetilde{p}_{b_{m-1}}(\mathbf{v}^{(b_{m-1})}) \leq \max(\mathbf{v}) < b'$ and $\widetilde{x}_{b_{m-1}}(\mathbf{v}^{(b_{m-1})}) = 1$. Since $|\mathbf{v}^{(m-1)}| \geq |\mathbf{v}| > k$, we know there exists some j^* where $x_{j^*}(\mathbf{v}^{(m-1)}) = 0$. Consider the deviation by user $j*$ where they inject a fake bid at b'. By assumption, $\widetilde{x}_{b'}(\mathbf{v}^{(m-1)}, b') = 0$, so $\mathsf{util}_j(\mathbf{v}^{(m-1)}, b') \geq \mathsf{util}_j(\mathbf{v}^{(m-1)})$. Conversely, now either $\widetilde{p}_{b_{m-1}}(\mathbf{v}^{(m-1)}, b') > b' > \max(\mathbf{v}) \geq \widetilde{p}_{b_{m-1}}(\mathbf{v}^{(m-1)})$ or $\widetilde{x}_{b_{m-1}}(\mathbf{v}^{(m-1)}, b') = 0$, both of which imply $\widetilde{\mathsf{util}}_{b_{m-1}}(\mathbf{v}^{(m-1)}, b') < \widetilde{\mathsf{util}}_{b_{m-1}}(\mathbf{v}^{(m-1)}, b')$. This contradicts URHP. □

Lemma 5.7 thus follows. □

6 IC and RHP are Incomparable

This section presents the mechanisms witness to show that IC and RHP are incomparable with respect to all three types of strategic players and in both the plain and the MPC-assisted models. We summarize the mechanisms in Table 2 for the plain model and Table 3 for the MPC-assisted model.

6.1 Comparison in the Plain Model

Table 2: Mechanisms for the incomparability of IC and RHP in the plain model.

Strategic Player	Mechanism	IC	RHP
User	Burning first-price auction (Section 6.1.1)	✗ UIC	✓ URHP
	Second price auction (Section 6.1.1)	✓ UIC	✗ URHP
Miner	All-or-nothing posted price (Section 4.1.1)	✗ MIC	✓ MRHP
	2-winner second-price (Section 6.1.2)	✓ MIC	✗ MRHP
Miner-user coalition	All-or-nothing posted price (Section 4.1.1)	✗ 1-SCP	✓ 1-CRHP
	2-winner second-price (Section 6.1.2)	✓ 1-SCP	✗ 1-CRHP

6.1.1 UIC vs. URHP in the Plain Model

Second-price auction

- **Inclusion & Confirmation:** Include highest two bids, and confirm the top bid. Break ties arbitrarily.
- **Payment & Revenue:** All confirmed bids pay the price of the lowest included bid and all payments go to the miner.

Lemma 6.1 (UIC does not imply URHP in the plain model)**.** *The above second-price auction satisfies UIC but not URHP in the plain model.*

Proof. **UIC:** UIC follows from the fact that for any fixed user i, for fixed other users' bids $\mathbf{b}_{-i}$, user i's allocation is monotone and the price is exactly as defined in Lemma 3.5.
Not URHP: Suppose the honest bid vector was $(0, v_2)$ where $0 < v_2$. Then, in the honest case, user 1 has utility 0 and user 2 has utility v_2. When user 1 raises their bid to $\frac{v_2}{2}$, their utility remains the same. However, user 2's payment rises, resulting in half the utility $v_2 - \frac{v_2}{2} = \frac{v_2}{2}$. □

Burning first-price auction

- **Inclusion & Confirmation:** Include and confirm the top k bids, breaking ties arbitrarily.
- **Payment & Revenue:** All confirmed bid price pay their bids, and all payments are burned.

Lemma 6.2 (URHP does not imply UIC in the plain model)**.** *The above burning first-price auction satisfies URHP but not UIC in the plain model. Additionally, it satisfies MIC, MRHP and d-CRHP for any $d \geq 1$.*

Proof. **URHP:** If everyone bids honestly, the total utility of all users and the miner is 0, so there is trivially no way to harm anyone.
Not UIC: There is an incentive for strategic underbidding, where a user can bid below their true valuation to increase utility while still being confirmed. □

The mechanism is also not 1-SCP because the miner can censor the top bids to help an unconfirmed user with positive valuation get confirmed.

6.1.2 MIC vs. MRHP in the Plain Model

2-winner second-price auction //Block size $k \geq 2$.

- **Inclusion Rule & Confirmation Rule:** Include and confirm the top 2 bids, breaking ties in some arbtrary deterministic way based on metadata.
- **Payment Rule:** All confirmed bids pay the price of the lowest confirmed bid.
- **Miner Revenue Rule:** The miner receives the price of the lowest confirmed bid as revenue while the rest of the payments are burned.

Lemma 6.3 (MIC does not imply MRHP in the plain model)**.** *The above 2-winner second price auction is MIC but not MRHP.*

Proof. **MIC:** The miner can only improve their utility by receiving more revenue from real users. Raising revenue can only happen through injecting a confirmed bid, but this cannot help the miners because the payment associated with the bid will cancel out all revenue by construction. The rest of the proof focuses on 1-SCP and 1-CRHP and MRHP.
Not MRHP: Consider the honest bid vector $(1, 1, 3)$. In the honest case, the miner revenue is 2 and user 3 has 2 utility. The miner can then censor 3. Under the deviation, the miner revenue is still 2 while user 3 now has utility 0. □

Lemma 6.4 (MRHP does not imply MIC in the plain model)**.** *All-or-nothing posted-price auction (Section 4.1.1) with $\epsilon > 0$ is MRHP but not MIC.*

Proof. By Lemma 4.2. □

6.1.3 1-SCP vs. 1-CRHP in the Plain Model

Lemma 6.5 (1-SCP does not imply 1-CRHP in the plain model)**.** *The 2-winner second-price auction is 1-SCP but not 1-CRHP in the plain model.*

Proof. **1-SCP:** Suppose, for the sake of contradiction, that $\mathbf{v}$ is the true value vector and there exists a coalition $\mathcal{C}$ consisting of the miner and a user i with valuation v_i, and a joint strategy $S_{\mathcal{C}}$ such that $\mathsf{util}_{\mathcal{C}}(\mathbf{v}; S_{\mathcal{C}}) > \mathsf{util}_{\mathcal{C}}(\mathbf{v}; H_{\mathcal{C}})$. Let p denote the payment associated with each confirmed bid under $S_{\mathcal{C}}$ and let p_H denote the payment associated with each confirmed bid in the honest outcome.

Case 1: i is confirmed under $S_{\mathcal{C}}$. Here, $\mathcal{C}$'s joint utility is at most the user's valuation minus the payment, plus the payment received by the miner:

$$v_i - p + p = v_i.$$

The coalition's utility could be strictly less than v_i if the deviation involves fake bids that incur additional payments. If i were also confirmed under the honest strategy, the coalition's total utility would again be v_i; thus the deviation yields no gain. Therefore, we can assume that i is unconfirmed under the honest strategy. The joint utility in the honest case comes solely from the miner revenue which equals p_H. Since i is unconfirmed, we know $v_i \leq p_H$. This shows S_C does not benefit the coalition

Case 2: i is unconfirmed under $S_{\mathcal{C}}$. Here, $\mathcal{C}$'s joint utility comes solely from the miner revenue, p, but could be lower due to confirmed fake bids. In the honest case, the joint utility is at least p_H, the miner revenue. If $p \leq p_H$, the deviation does not benefit $\mathcal{C}$. If $p > p_H$, then there must be at least one confirmed fake bid injected at or above p, meaning the coalition's joint utility under $S_{\mathcal{C}}$ would be at most $0 \leq p_H$.

In both cases, $S_{\mathcal{C}}$ does not produce a joint utility higher than the honest joint utility for the coalition. This is a contradiction.

Not 1-CRHP: Consider the honest value vector $(1, 2)$ and the coalition $\mathcal{C}$ consisting of the miner and user 1. Under the honest strategies, $\mathcal{C}$'s joint utility is 1 while user 2's utility is 1. However, if the miner censors user 2, $\mathcal{C}$'s joint utility is still 1 while user 2's utility is 0. Thus, this strategy harms user 2 without harming the coalition. □

Lemma 6.6 (1-CRHP does not imply 1-SCP in the plain model)**.** *All-or-nothing posted-price auction (Section 4.1.1) with $\epsilon > 0$ is d-CRHP for any $d \geq 1$ but not 1-SCP.*

Proof. By Lemma 4.2. □

6.2 Comparison in the MPC-Assisted Model

Table 3: Mechanisms for the incomparability of IC and RHP in the MPC-assisted model.

Strategic Player	**Mechanism**	**IC**	**RHP**
User	Burning first-price auction (Section 6.1.1)	✗ UIC	✓ URHP
	Second price auction (Section 6.1.1)	✓ UIC	✗ URHP
Miner	Revenue-capped first-price (Section 6.2.2)	✗ MIC	✓ MRHP
	Trigger-posted-price (Section 6.2.2)	✓ MIC	✗ MRHP
Miner-user coalition	Burning first-price (Section 6.1.1)	✗ 1-SCP	✓ 1-CRHP
	Trigger-posted-price (Section 6.1.2)	✓ 1-SCP	✗ 1-CRHP

6.2.1 UIC vs. URHP in the MPC-Assisted Model

Lemma 6.1 and Lemma 6.2 still hold in the MPC-assisted model.

6.2.2 MIC vs. MRHP in the MPC-Assisted Model

MPC-assisted, trigger-posted-price auction //Reserve price $\mathsf{res} > 0$.

- **Inclusion Rule & Confirmation Rule:** If there is one bid strictly above reserve, include and confirm that bid. Otherwise, no bids are included or confirmed.
- **Payment & Miner Revenue Rule:** All confirmed bids pay the reserve price, and all payments are burned.

Lemma 6.7 (MIC does not imply MRHP in the MPC-assisted model.)**.** *The above MPC assisted, trigger posted price auction satisfies MIC but not MRHP.*

Proof. All confirmed bids pay the same amount, so we will only be referring to the confirmation probability in the proof below.
MIC: The miner receives no revenue so MIC, is trivially satisfied.
Not MRHP: Consider when there is an unconfirmed user and a confirmed user with positive utility in the honest case i.e. $(0, \mathsf{res}+1)$. Then, the coalition containing the miner(s) and user 1 (or just user 1 or just the miner separately) have an expected joint honest utility of 0, and user 2 has positive expected honest utility. If the coalition (or just user 1 or just the miner separately) injects two bids above reserve, the expected joint utility of the miner(s) and user 1 (or just user 1 or just the miner separately) is still 0 while user 2's bid is now unconfirmed, leading to a zero utility. □

Revenue-capped first-price auction //Block size $k \geq 2$

- **Inclusion Rule & Confirmation Rule:** Include the top 2 bid but confirm only the highest bid, breaking ties randomly.
- **Payment Rule:** All confirmed bids pay their bid.
- **Miner Revenue Rule:** The miner receives the price of the highest non-confirmed bid as payment or 0 otherwise.

Lemma 6.8 (MRHP does not imply MIC in the MPC-assisted model)**.** *The above mechanism satisfies MRHP but not MIC in the MPC-assisted model.*

Proof. **MRHP:** In the honest case, all users receive 0 utility, so the mechanism trivially satisfies MRHP.
Not MIC: In the bid vector $(0, 1)$, the miners can inject a fake bid between the highest and second highest bids to receive more revenue. □

6.2.3 1-SCP vs. 1-CRHP in the MPC-Assisted Model

Lemma 6.9 (1-SCP does not imply 1-CRHP in the MPC-assisted model)**.** *The trigger-posted-price satisfies* 1*-SCP but not* 1*-CRHP in the MPC-assisted model.*

Proof. 1**-SCP:** Since the miner revenue is 0 in all cases, every coalition's utility is just the user's utility. Thus, 1-SCP follows directly from UIC in the MPC-assisted model.
Not 1**-CRHP:** By the same reason why trigger-posted-price is not MRHP. □

Lemma 6.10 (1-CRHP does not imply 1-SCP in the MPC-assisted model)**.** *The burning first-price auction d-CRHP for any $d \geq 1$ but not* 1*-SCP in the MPC-assisted model.*

Proof. d**-CRHP:** If everyone bids honestly, the total utility of all users and the miner is 0, so there is trivially no way to harm anyone.
Not 1**-SCP:** Since the miner gets 0 revenue in all cases, this follows from the mechanism not being UIC. □

Acknowledgements

This work was supported in part by University of Michigan start-up funds and by the Stellar Academic Research Grant (2024).

References

[AL20] Mohammad Akbarpour and Shengwu Li. Credible auctions: A trilemma. *Econometrica, Econometric Society*, 2020. 1.3

[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`. 1, 1.2

[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. 1

[BGR24] Maryam Bahrani, Pranav Garimidi, and Tim Roughgarden. Transaction Fee Mechanism Design in a Post-MEV World. In Rainer Böhme and Lucianna Kiffer, editors, *6th Conference on Advances in Financial Technologies (AFT 2024)*, volume 316 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 29:1–29:24, Dagstuhl, Germany, 2024. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. 1.3

[Can01] R. Canetti. Universally composable security: A new paradigm for cryptographic protocols. In *FOCS*, 2001. 1.1, 3.3.2

[Car19] Gabriel Carroll. Robustness in mechanism design and contracting. *Annual Review of Economics*, 11(1):139–166, 2019. 1.3

[CRS24] Hao Chung, Tim Roughgarden, and Elaine Shi. Collusion-resilience in transaction fee mechanism design. In *Proceedings of the 25th ACM Conference on Economics and Computation*, pages 1045–1073, 2024. 1.3

[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. (document), 1, 1.1, 1.3

[CSLZZ25] Xi Chen, David Simchi-Levi, Zishuo Zhao, and Yuan Zhou. Bayesian mechanism design for blockchain transaction fee allocation. *Operations Research*, 2025. 1.3

[DPTM24] Eduardo Duque, Juan Pereyra, and Juan Pablo Torres-Martínez. *Local Non-Bossiness and Preferences Over Colleagues.* Universidad de Chile, Departamento de Economía, 2024. 1, 1.3

[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. 1.3

[FMPS21] Matheus VX Ferreira, Daniel J Moroz, David C Parkes, and Mitchell Stern. Dynamic posted-price mechanisms for the blockchain transaction-fee market. In *Proceedings of the 3rd ACM Conference on Advances in Financial Technologies*, pages 86–99, 2021. 1

[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. 1.3

[GMW87] O. Goldreich, S. Micali, and A. Wigderson. How to play any mental game. In *ACM symposium on Theory of computing (STOC)*, 1987. 1.1, 3.3.2

[GRS90] Ronald L. Graham, Bruce L. Rothschild, and Joel H. Spencer. *Ramsey Theory*. Wiley, 1990. 2.2, 4.3, 4.20

[GTW24] Aadityan Ganesh, Clayton Thomas, and S. Matthew Weinberg. Revisiting the primitives of transaction fee mechanism design. In *Proceedings of the 25th ACM Conference on Economics and Computation (EC)*, 2024. 1.3

[GY23] Yotam Gafni and Aviv Yaish. Greedy transaction fee mechanisms for (non-)myopic miners. In *The 12th Annual Conference of the Israeli Chapter of the Game Theory Society (IGTC'23)*, 2023. 1.3

[GY24] Yotam Gafni and Aviv Yaish. Discrete and bayesian transaction fee mechanisms. In *The International Conference on Mathematical Research for Blockchain Economy*, pages 145–171. Springer, 2024. 1.3

[JM05] Philippe Jehiel and Benny Moldovanu. Allocative and informational externalities in auctions and related mechanisms. 2005. 1.3

[LRS21] Giuseppe Lopomo, Luca Rigotti, and Chris Shannon. Uncertainty in mechanism design. *arXiv preprint arXiv:2108.12633*, 2021. 1.3

[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. 1

[LSZ23] Renato Paes Leme, Jon Schneider, and Hanrui Zhang. Nonbossy mechanisms: Mechanism design robust to secondary goals. *arXiv preprint arXiv:2307.11967*, 2023. 1, 1.3

[Mye81] Roger B. Myerson. Optimal auction design. *Math. Oper. Res.*, 6(1), 1981. 3.5

[Rou21] Tim Roughgarden. Transaction fee mechanism design. *ACM SIGecom Exchanges*, 19(1):52–55, 2021. (document), 1, 1.3

[SCW23] Elaine Shi, Hao Chung, and Ke Wu. What Can Cryptography Do for Decentralized Mechanism Design? In Yael Tauman Kalai, editor, *14th Innovations in Theoretical Computer Science Conference (ITCS 2023)*, volume 251 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 97:1–97:22, Dagstuhl, Germany, 2023. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. (document), 1, 1.1, 1.3

[SS81] Mark A Satterthwaite and Hugo Sonnenschein. Strategy-proof allocation mechanisms at differentiable points. *The Review of Economic Studies*, 48(4):587–597, 1981. 1, 1.3

[Tho16] William Thomson. Non-bossiness. *Social Choice and Welfare*, 47(3):665–696, 2016. 1, 1.3

[Vic61] William Vickrey. Counterspeculation, auctions, and competitive sealed tenders. *The Journal of finance*, 16(1):8–37, 1961. 1

[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)*, volume 287 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 98:1–98:23, Dagstuhl, Germany, 2024. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. 1.3

[Yao] Andrew Chi-Chih Yao. An Incentive Analysis of Some Bitcoin Fee Designs (Invited Talk). In *ICALP 2020*. 1