Foundations of Transaction Fee Mechanism Design

Hao Chung
CMU
haochung@andrew.cmu.edu

Elaine Shi
CMU
runting@cs.cmu.edu

Abstract

In blockchains such as Bitcoin and Ethereum, users compete in a transaction fee auction to get their transactions confirmed in the next block. A line of recent works set forth the desiderata for a "dream" transaction fee mechanism (TFM), and explored whether such a mechanism existed. A dream TFM should satisfy 1) *user incentive compatibility* (UIC), i.e., truthful bidding should be a user's dominant strategy; 2) *miner incentive compatibility* (MIC), i.e., the miner's dominant strategy is to faithfully implement the prescribed mechanism; and 3) *miner-user side contract proofness* (SCP), i.e., no coalition of the miner and one or more user(s) can increase their joint utility by deviating from the honest behavior. The weakest form of SCP is called 1-SCP, where we only aim to provide resilience against the collusion of the miner and *a single* user. Sadly, despite the various attempts, to the best of knowledge, no existing mechanism can satisfy all three properties in all situations.

Since the TFM departs from classical mechanism design in modeling and assumptions, to date, our understanding of the design space is relatively little. In this paper, we further unravel the mathematical structure of transaction fee mechanism design by proving the following results:

- *Can we have a dream TFM?* We prove a new impossibility result: *assuming finite block size*, no single-parameter, non-trivial, possibly randomized TFM can simultaneously satisfy UIC and 1-SCP. Consequently, no non-trivial TFM can satisfy all three desired properties simultaneously. This answers an important open question raised by Roughgarden in his recent work.
- *Rethinking the incentive compatibility notions.* We observe that the prevalently adopted incentive compatibility notions may be too draconian and somewhat flawed. We rectify the existing modeling techniques, and suggest a relaxed incentive compatibility notion that captures additional hidden costs of strategic deviation. We construct a new mechanism called the "burning second-price auction", and show that it indeed satisfies the new incentive compatibility notions. We additionally prove that the use of randomness is necessary under the new incentive compatibility notions for "useful" mechanisms that resist the coalitions of the miner and at least 2 users.
- *Do the new design elements make a difference?* Unlike classical mechanisms, TFMs may employ a couple new design elements that are idiosyncratic to blockchains. For example, a burn rule (employed by Ethereum's EIP-1559) allows part to all of the payment from the users to be burnt rather than paid to the miner. Some mechanisms also allow unconfirmed transactions to be included in the block, to set the price for others. Our work unveils how these new design elements actually make a difference in TFM design, allowing us to achieve incentive compatible properties that would otherwise be impossible.

Contents

1 Introduction

In decentralized blockchains such as Bitcoin and Ethereum, miners are incentivized to participate in and collectively maintain the public ledger, since they can collect block rewards and transaction fees. Today, a simple "pay your bid" auction is implemented by major blockchains. In a "pay your bid" auction, the miners' dominant strategy is to take the highest bids. However, users may be incentivized to bid strategically, e.g., bid close to 0 when there is no congestion, or bid the minimum possible to get selected when there is congestion. Earlier works [LSZ19, Yao18, BEOS19] pointed out such strategic bidding is indeed happening in real life, and is generally considered undesirable. Consequently, several works [LSZ19, Yao18, BEOS19, Mon20, Rou20, Rou21b, FMPS21] call out to the community to rethink the design of transaction fee mechanisms (TFMs). These works raise the following important question: *what is the ideal transaction fee mechanism*?

Desiderata of a dream TFM. Partly due to its decentralized nature, transaction fee mechanism (TFM) design departs from classical mechanism design [Mye81, NRTV07] in modeling and assumptions. First, we are confronted with several new challenges that arise from the strategic behavior of the miner and of miner-user coalitions:

- *Challenge 1: strategic behavior of the miner.* In the classical setting, we typically assume that the auctioneer is trusted and implements the prescribed mechanism honestly — therefore, we mainly care about how to design mechanisms such that the users are incentivized to bid truthfully. In a decentralized environment, the auctioneer is no longer fully trusted. In a blockchain transaction fee mechanism, the miners and the logic of the blockchain jointly serve as the "auctioneer". Although the logic of the blockchain is hard-coded and unalterable, miners can deviate from the prescribed mechanism, and behave strategically to increase its financial gains. As a simple example, consider a classical Vickrey auction [Vic61]. Suppose that each block has size B. We can then include the top B bids into the block, among which the first $B-1$ are considered *confirmed* and they pay the B-th price. If there are strictly fewer than B bids, everyone gets confirmed and they all pay a price of 0. All users' payment goes to the miner that mines the block. Classical algorithmic game theory [Vic61, NRTV07] tells us that such a Vickrey auction is dominant strategy incentive compatible (DSIC) for the users, assuming that the miner indeed behaves honestly. Unfortunately, several prior works [BEOS19, Rou20] pointed out that the Vickrey auction is not incentive compatible for the miner, since the miner may want to inject a fake transaction whose price is between the $(B-1)$-th and B-th price to increase its revenue.

- *Challenge 2: miner-user collusion.* In a decentralized blockchain, it is easy for two or more parties to form binding side contracts through smart contracts. A miner could collude with a user to increase the joint utility of the coalition, and the two can then split the gains with a binding side contract. In the aforementioned Vickrey auction example, the miner could alternatively ask the B-th bidder to raise its bid to be infinitesimally smaller than the $(B-1)$-th bid, and then split its gains with the B-th bidder in a side contract.

  Most prior works [LSZ19, BEOS19, Rou20, Rou21b] focused on miner-user collusion rather than user-user collusion, likely for the following reason: it is much easier to facilitate miner-user rendezvous since the big miners are well-known. In comparison, users are ephemeral and thus user-user rendezvous is much more costly to facilitate.

With these challenges in mind, prior works [LSZ19, Rou20] have suggested the following desiderata for a "dream" transaction fee mechanism:

1. *User incentive compatibility (UIC).* Assuming that the miner implements the mechanism honestly, then following the honest bidding strategy or truthful bidding should be a dominant strategy for the users.

2. *Miner incentive compatibility (MIC).* Assuming that users follow the honest bidding strategy or bid truthfully, a miner's dominant strategy should be to implement the prescribed mechanism faithfully.

3. *Miner-user side contract proofness (c-SCP).* No coalition of the miner and up to c users can increase their joint utility through any deviation. In the above, c is a parameter that specifies an upper bound on the coalition's size. The larger the c, the more side contract resilient. Note that it is generally harder for a miner and a large number of users to engage in a side contract, than, say, a miner and a single user.

To the best of our knowledge, all prior works [LSZ19, Yao18, BEOS19, Rou20] fall short of achieving all three properties at the same time — see Section 2.5 for more detailed discussions on these prior works. The closest we have come to achieving all three properties is Ethereum's recent EIP-1559 [Mon20] proposal. The very recent work of Roughgarden [Rou20] showed that (a close variant of) EIP-1559 can achieve all three properties *assuming that the block size is infinite* (or more precisely, assuming that the base fee is set high enough such that the number of transactions willing to pay the base fee is upper bounded by the block size). However, when there is congestion, EIP-1559 acts like a first-price auction and therefore fails to satisfy UIC, i.e., strategic bidding could improve an individual user's utility.

Open question 1: With all these failed attempts, it is natural to ask: *is it actually feasible to have a "dream" transaction fee mechanism that satisfies all three properties simultaneously?* Is the community's lack of success so far due to a more fundamental mathematical impossibility? Roughgarden also raised this as a major open question in his recent work [Rou20, Rou21b].

Open question 2: If there is indeed a mathematical impossibility, then the natural next question to ask is: are the current incentive compatibility notions overly stringent? If so, can we relax the incentive compatibility notion to circumvent the impossibilities?

TFM design space enriched by new elements. In comparison with classical mechanism design, transaction fee mechanisms may employ a couple novel features that are idiosyncratic to blockchains. For example, Ethereum's EIP-1559 [Mon20, Rou20] suggested the novel usage of a *burn rule*, where part to all of the fees collected from the confirmed transactions may be "burnt" rather than paid to the miner. Another design consideration that is being debated in the community is whether we should allow blocks to contain unconfirmed transactions that are just there to "set the price". Although this approach has been employed by some suggested mechanisms [LSZ19, Yao18], an argument against it is that real estate on a blockchain is scarce — therefore, we ideally do not waste space including unconfirmed transactions [Rou21a].

An intriguing question is the following:

Open question 3: Do these novel elements actually make a difference in the design of transaction fee mechanisms? Can they help achieve incentive compatible mechanism designs that would otherwise be impossible?

1.1 Our Results and Contributions

1.1.1 Impossibility of a Having a "Dream" Transaction Fee Mechanism

We prove an impossibility result (Theorem 1.1) showing that assuming finite block size, there is no non-trivial transaction fee mechanism (TFM) that satisfies UIC and 1-SCP, where 1-SCP means resilience against side contracts between the miner and *a single* user. Consequently, there is also no non-trivial TFM that satisfies all three desired properties.

Theorem 1.1 (Impossibility of a "dream" transaction fee mechanism (informal))**.** *Suppose that the block size is finite. There does not exist a non-trivial, single-parameter transaction fee mechanism (TFM) that simultaneously satisfies UIC and 1-SCP. Moreover, this impossibility holds for both deterministic and randomized mechanisms.*

Another way to understand Theorem 1.1 is the following: the only TFM that satisfies UIC and 1-SCP simultaneously is the trivial mechanism that always confirms nothing and pays the miner nothing.

1.1.2 Definitional Contribution: Incentive Compatibility under γ-Strict Utility

While our aforementioned impossibility result paints a pessimistic outlook, we observe that the previously formulated incentive compatibility notions appear too draconian and somewhat flawed. So far, almost all prior works [LSZ19, Yao18, BEOS19, Rou20, Rou21b] model the TFM in a standalone setting, where the players are myopic and care only about their gain or loss in the current auction instance. In this setting, if a strategic player (which is either a user, a miner, or a miner-user coalition) injects a fake transaction whose true value is 0, or if it overbids (i.e., bids more than the transaction's true value), we assume that the offending transaction is free of charge as long as it is not confirmed in the present block — since an unconfirmed transaction need not pay any fees.

In practice, however, the TFM is executed in a repeated setting as blocks get confirmed. Any transaction that has been posted to the network cannot be retracted even if it is not confirmed in the present block. In particular, a fake or overbid transaction could be confirmed in a future block, and thus the strategic player would end up paying fees to the future block, potentially mined by a different miner. For example, consider the Vickrey auction example again, where we include the B highest bids in the block, among which the top $B-1$ are confirmed and pay the B-th price. Suppose that all payment goes to the miner. In this case, the miner may want to inject a fake transaction whose bid is in between the $(B-1)$-th and the B-th price, to increase its revenue. In prior works as well as our aforementioned impossibility result, we assume that injecting this fake transaction is free because it is unconfirmed. However, in practice, the injected transaction may be confirmed and paying fees in a future block.

A natural question is whether we can capture this cost of cheating in our model, and thus circumvent the aforementioned impossibility. In our new approach, we still model the TFM as a single-shot auction, but we want to more accurately charge the cost of cheating in the utility model. Unfortunately, we are faced with a notable challenge: accurately predicting the cost of cheating is difficult, since what the offending transaction actually pays in the future depends on the environment, e.g., what other users are bidding, as well as the mechanism itself.

Defining γ-strict utility. To make progress, we take the following approach. We first ask what is the worst-case cost of cheating. This is when the overbidding or fake transaction that is unconfirmed in the present ends up paying its full bid in the future, thus incurring a cost as high as the difference between the bid and the true value of the transaction. This is the worst case for the

cheater and thus the best case for the mechanism designer. Asking whether there is a mechanism that satisfies incentive compatibility under the most strict cost model is equivalent to asking: can we at least design mechanisms that defend against *paranoid* strategic players who only want to deviate if there is a sure chance of gain and no chance of losing. Understanding the feasibility of mechanism design under the worst-case cost can provide shed light on whether this is a worthwhile direction. Further, it is also useful to adopt the worst-case cost model in proving lower bounds, since that makes the lower bounds stronger.

Next, we generalize the cost model and imagine that in reality, the offender only needs to pay γ fraction of the worst-case cost, where $\gamma \in [0, 1]$ is also called the *discount* factor. This generalization is useful because in practice, we can often estimate the cost of cheating from past data. A mechanism that satisfies UIC (or MIC, c-SCP, resp.) under this cost model is also said to satisfy UIC (or MIC, c-SCP, resp.) under *γ-strict utility*. Specifically, when $\gamma = 0$, there is no cost of cheating — in this case, our new incentive compatibility notions would then degenerate to the previous notions. When $\gamma = 1$, this is when we are charging the worst-case cost for cheating. Since we are often particularly interested in the case of $\gamma = 1$ (e.g., when proving lower bounds), for convenience, a mechanism that satisfies UIC (or MIC, c-SCP, resp.) under 1-strict utility is also said to satisfy *weak UIC (or weak MIC, c-weak-SCP, resp.)*.

We present our new incentive compatibility notions more formally in Section 5.

1.1.3 The Mathematical Structure of Incentive Compatibility under γ-Strict Utility

The burning second-price auction. Using our new γ-strict utility notion, we can circumvent the aforementioned impossibility (Theorem 1.1). Specifically, we describe a new mechanism called the burning second-price auction (see Section 2.3) that achieves UIC, MIC, and c-SCP under γ-strict utility for any $\gamma \in (0, 1]$, and any choice of coalition resilience parameter $c \geq 1$. The mechanism is randomized, and one can view the parameters c and γ that allow us to tradeoff the degree of incentive compatibility and the efficiency of the mechanisms (in terms the expected number of bids confirmed).

Formally, we prove the following theorem:

Theorem 1.2 (Burning second price auction)**.** *For any $\gamma \in (0, 1]$ and any $c \geq 1$, there exists a TFM that satisfies UIC, MIC, and c-SCP under γ-strict utility. Further, the TFM can support any finite block size, and except for the case when $c = 1$ and $\gamma = 1$, the TFM is randomized.*

Necessity of randomness. As mentioned in Theorem 1.2, except for the special case $c = 1$ and $\gamma = 1$, our burning second-price auction is randomized. In particular, the mechanism employs trusted on-chain to pick a random subset of eligible, included transactions to confirm. Although unbiased and unpredictable on-chain randomness can be generated using standard cryptographic techniques [CKS00, BSKN21, DKIR21], such coin toss protocols introduce some extra overhead, and ideally we would like to avoid them. Unfortunately, we prove a lower bound that for $c \geq 2$, randomness is necessary to achieve UIC and c-SCP for any $\gamma \in [0, 1]$, as long as the mechanism is "useful" in the sense that it sometimes confirms at least 2 bids.

Theorem 1.3 (Necessity of randomness for weak incentive compatibility)**.** *Consider an arbitrary deterministic TFM and assume finite block size. Suppose that there exists a bid vector such that the TFM confirms at least two bids. Then, the TFM cannot satisfy both weak UIC and 2-weak-SCP simultaneously.*

In the above lower bound, the restriction that the mechanism must sometimes confirm 2 bids is necessary. Specifically, we construct a deterministic mechanism called the solitary mechanism

(Appendix B.1) that always confirms a single bid, and achieves weak UIC, weak MIC, and c-weak-SCP for any $c \geq 1$.

1.1.4 Understanding New Design Elements for TFM

Finally, as mentioned, unlike classical auctions, TFMs can employ a couple new design elements. First, the mechanism can employ a "burn rule", which allows part to all of the users' payment to be "burnt" on the blockchain, and not paid to the miner of the present block. For example, Ethereum's EIP-1559 makes critical use of such a burn rule [Mon20, Rou20, Rou21b]. Second, some prior works [LSZ19, Yao18] have suggested including transactions in a block that are not confirmed eventually, but serve the role of setting the price for others. For example, even though we know that the Vickrey auction is not an awesome auction in a decentralized environment, hypothetically, imagine we want to implement the Vickrey auction on a blockchain. This would require that the block to include B bids, among which only the top $B-1$ are eventually confirmed, whereas the B-th bid is included only to set the price. Moreover, our own burning second-price auction (Theorem 1.2) also include some transactions in the block that have no chance of being confirmed, but are just there to set the price.

Do these new design elements make a difference in TFM design, and can they help us achieve incentive compatibility designs that would otherwise be impossible? We give a nuanced answer to this question. First, we point our that our earlier impossibility results (Theorems 1.1 and 1.3) hold even when the TFM is allowed to employ both of these design elements.

On the other hand, we also show scenarios in which these new design elements do make a difference. Specifically, we prove the following results.

The burn rule is critical to Ethereum's EIP-1559. Recall that Roughgarden [Rou20, Rou21b] argued that *assuming infinite block size*, Ethereum's EIP-1559 approximates a simple "posted price, burn all" auction: there is an a-priori fixed price tag r, and anyone who bids at least r would get their transaction confirmed, paying only r. All users' payment is burnt, and the miner gets nothing. Roughgarden [Rou20, Rou21b] also proved that this simple "posted price, burn all" auction would indeed satisfy UIC, MIC, and c-SCP for any $c \geq 1$, assuming infinite block size. We show that without the burn rule, even under infinite block size, the only way for a TFM to satisfy both UIC and 1-SCP is for users to always pay nothing. More generally, we prove that even when assuming infinite block size and whether we allow a burn rule or not, any (randomized) TFM that satisfies both UIC and 1-SCP must always pay the miner nothing:

Theorem 1.4 (The burn rule makes a difference assuming infinite block size)**.** *Any (randomized) TFM that satisfies both UIC and 1-SCP must always pay the miner nothing. This impossibility holds regardless of whether the block size is infinite or finite, and regardless of whether the TFM has a burn rule or not*[1]*.*

As a direct implication, if we want a mechanism like EIP-1559 that has non-trivial user payment and satisfies all three properties in the infinite block size regime, a burn rule is necessary. Therefore, Theorem 1.4 and Roughgarden's result [Rou20, Rou21b] together show that having a burn rule does make a difference assuming infinite block size.

[1] In fact, in our actual proof, we prove Theorem 1.4 first, which is then used as a stepping stone towards proving Theorem 1.1.

Necessity for blocks to contain unconfirmed transactions. In the cryptocurrency community, there is an ongoing debate whether it is a waste of space for blocks to contain unconfirmed transactions. We argue that the ability for a block to contain unconfirmed transactions could indeed make a difference for the mechanism designer. Specifically, we prove a corollary showing that if one insist on confirming all transactions included in a block, then even with the weak incentive compatibility notion, it is still impossible to construct any non-trivial (possibly randomized) TFM that satisfies weak UIC and 1-weak-SCP.

Corollary 1.5 (Allowing unconfirmed transactions in a block can make a difference)**.** *Suppose that all transactions in a block must be confirmed. Then, there is no non-trivial (possibly randomized) TFM that satisfies weak UIC and 1-weak-SCP simultaneously.*

Additional results. In the appendices, we additionally show a variant of Theorem 1.3 that says if the TFM is not allowed to have a burning rule, then Theorem 1.3 holds even when the block size is allowed to be infinite.

2 Technical Roadmap

2.1 Transaction Fee Mechanism and Incentive Compatibility

In a transaction fee mechanism (TFM), we are selling slots in a block to bidders who want to get their transactions included and confirmed in the block. For simplicity, we assume that all slots are identical commodities, and we often use the terms "transaction" and "bid" interchangeably. For convenience, we assume that each bid comes from a different user.

Transaction fee mechanism. A transaction fee mechanism (TFM) includes the following rules:

- An *inclusion rule* executed by the miner. Given a bid vector $\mathbf{b} = (b_1, b_2, \ldots, b_m)$, the inclusion rule decides which of the bids to include in the block;
- A *confirmation rule* executed by the blockchain. The confirmation rule chooses a subset of the included bids to confirmed. In the most general form, not all transactions included in the block are necessarily confirmed, and only confirmed transactions are considered final, i.e., the money has been transferred to the merchant's account and the merchant can now provide the promised service.
- A *payment rule* and a *miner revenue* rule executed by the blockchain, which decides (using only information recorded in the block) how much each confirmed bid pays, and how much revenue the miner gets. Any (possibly included) transaction that is not confirmed pays nothing. Furthermore, we assume that the miner's revenue is upper bounded by the total payment collected from all confirmed bids. In particular, if the miner's revenue is strictly smaller than the total payment of all bids, then we often say that part of the payment is *burnt*.

In our model, a strategic miner (possibly colluding with some users) may not implement the honest inclusion rule, if deviating can benefit the miner (or coalition). However, the blockchain is trusted to implement the confirmation, payment, and miner-revenue rules honestly.

In comparison with Roughgarden's model [Rou20, Rou21b], we explicitly distinguish the inclusion rule from the confirmation rule in our modeling. By contrast, Roughgarden's model calls the union of the inclusion rule and the confirmation rule the *allocation rule*. Making the distinction between the inclusion and confirmation rules explicit is useful for us since we want to tease out the

fine boundaries between feasibility and infeasibility, depending on whether the block size is finite or infinite.

Strategy space and incentive compatibility. A *strategic player* can be a user, a miner, or the coalition of the miner and up to c users. The strategic player can deviate in the following ways: 1) if one or more users are involved, then some of the users can decide to bid untruthfully *after* examining all other bids; 2) the strategic player can inject fake bids after examining all other bids; and 3) if the miner is involved, then the miner may not implement the inclusion rule honestly.

Every user has a true value for its transaction to be confirmed. If a user is confirmed, its utility is its true value minus its payment. An unconfirmed user has utility 0. The miner's utility is its revenue. If the miner colludes with some users, the coalition's joint utility is the sum of the utilities of all coalition members.

Incentive compatibility. The honest strategy for a user is to bid its true value. The honest strategy for a miner is to implement the correct inclusion rule. A TFM is incentive compatible for a strategic player iff deviating from the honest strategy cannot increase the strategic player's expected utility; i.e., playing honestly is the strategic player's best strategy (or one of the best strategies). A TFM is said to be user incentive compatible (UIC), if it is incentive compatible for any individual user. A TFM is said to be miner incentive compatible (MIC), if it is incentive compatible for the miner. Finally, a TFM is said to be c-side-contract-proof, if it is incentive compatible for any coalition consisting of the miner and at least 1 and at most c users. The notions UIC, MIC, and c-SCP are incomparable as shown in Appendix A.

Note that in a blockchain environment, user-user coalitions are much harder to form: since users are ephemeral, rendezvous between them is challenging. By contrast, there are typically a stable set of big miners which makes miner-user rendezvous easy. For this reason, most works in this space are more interested in defending against miner-user rather than user-user coalitions.

2.2 Impossibility of a "Dream" TFM under Finite Block Size

We now sketch how to prove Theorem 1.1, that is, assuming finite block size, no non-trivial TFM can achieve UIC and 1-SCP at the same time. We shall first sketch how the proof works for deterministic TFMs, then we explain how to generalize the proof to the randomized case.

Deterministic case: miner has 0 revenue. Recall that if a TFM satisfies UIC, it must respect the constraints imposed by the famous Myerson's Lemma [Mye81]. For deterministic mechanisms, this means that the confirmation decision is *monotone*, and moreover, every confirmed bid pays *the minimum price it could have bid and still remained confirmed*, assuming everyone else's bids remain the same.

To prove Theorem 1.1, we go through an intermediate stepping stone: we shall actually prove Theorem 1.4 first, that is, any TFM that satisfies both UIC and 1-SCP *must always pay the miner nothing*, regardless whether the block size is finite or infinite. Henceforth, let $\mu(\mathbf{b})$ denote the miner revenue under the bid vector $\mathbf{b}$. We use $p_i(\mathbf{b})$ to denote user i's payment under $\mathbf{b}$, and if user i is unconfirmed, $p_i(\mathbf{b}) = 0$.

Consider an arbitrary deterministic TFM that is UIC and 1-SCP. Consider an arbitrary bid vector $\mathbf{b} = (b_1, \ldots, b_m)$ and we want to argue that the miner has 0 revenue under $\mathbf{b}$. To do this, we want to lower each user's bid to 0 one by one, and argue that the miner revenue is unaffected in this process. If this is the case, we can show that the miner revenue is 0 under $\mathbf{b}$, since at the very end of this process, when we have lowered everyone's bid to 0, the miner revenue must be 0.

It suffices to prove the following. Let $\mathbf{b} = (b_1, \ldots, b_m)$ be an arbitrary bid vector and $i \in [m]$ be an arbitrary user. We want to show that $\mu(\mathbf{b}) = \mu(\mathbf{b}_{-i}, 0)$. First, we show that if a user changes its bid such that its confirmation status remains unaffected, then the miner revenue should stay the same (Claim 4.4). If this is not true, then the miner and the user can collude, and there is a way for the user to bid untruthfully without affecting its confirmation status and thus its utility, but increasing the miner revenue. Overall, the coalition strictly gains and this violates 1-SCP. Suppose that p_i is the minimum price that some user i could bid to let it be confirmed, assuming that everyone else is bidding $\mathbf{b}_{-i}$. The above means that if user i bids anywhere between $[p_i, \infty]$ such that it remains confirmed, then the miner revenue is unaffected. Similarly, if user i bids anywhere between $(p_i, 0]$ such that it is unconfirmed, then the miner revenue is unaffected too.

It remains to rule out the possibility that there is a sudden jump in miner revenue, when user i lowers its bid from p_i to $p_i - \epsilon$ for an aribitrarily small ϵ. Suppose for the sake of contradiction that there is a sudden $\Delta > 0$ increase in the miner revenue when user i lowers its bid from p_i to $p_i - \epsilon$ (and the proof for the other direction is similar). From what we proved earlier, the entire jump of Δ must occur within an arbitrarily small interval p_i and $p_i - \epsilon$, and in particular, we may assume that $\epsilon < \Delta$. In this case, if the miner colludes with user i whose true value is actually $p_i - \epsilon$, the user should bid p_i instead. This way, the miner's gain Δ outweighs the user's loss ϵ, and the coalition strictly gains. This violates 1-SCP. A formal presentation of this proof can be found in Section 4.3.

Theorem 1.4 + finite block size $\Longrightarrow$ Theorem 1.1. Once we have proven Theorem 1.4, i.e., the miner always has 0 revenue, we can now throw in the finite block size assumption, to prove Theorem 1.1. We show it for the deterministic case below. Specifically, suppose there is a bid vector $\mathbf{b} = (b_1, \ldots, b_m)$ under which some bid b_i is confirmed where $i \in [m]$. Now, imagine that the real world actually consists of the bids $\mathbf{b}$ plus sufficiently many users bidding $b_i + \epsilon$, such that the number of users bidding $b_i + \epsilon$ exceeds the block size. We know that one of the users bidding $b_i + \epsilon$ must be unconfirmed — let us call this user u. The miner can now collude with u, and ask u to bid b_i instead. The miner can now pretend that the world consists of the bid vector $\mathbf{b}$ where b_i is replaced with u's bid, and run the honest mechanism. This helps the user u get confirmed and gain a positive utility, and meanwhile, the miner itself always gets 0 revenue no matter what it does. Thus, overall, the coalition strictly gains, which violates 1-SCP.

Generalizing to randomized TFMs. At a high level, our earlier impossibility proof for deterministic TFMs use Myerson's Lemma as a blackbox. Since the TFM is UIC, we argue that the mechanism must fall within the solution space characterized by Myerson's Lemma. Our proof then shows that the constraints imposed by Myerson conflict with the requirements of 1-SCP. For the randomized case, instead of following the same blueprint, we present an alternative proof that uses Myerson's Lemma (the randomized case) in a slightly non-blackbox manner — we review Myerson's Lemma generalized to the randomized case in Section 4.2. Below, keep in mind that the notations $p_i(\mathbf{b})$ and $\mu(\mathbf{b})$ can be random variables.

We first give a slightly incorrect intuition. As a thought experiment, imagine that the coalition of the miner and user i forms a "virtual-user" i. Virtual-user i 's true value is v_i, i.e., same as user i's true value. Virtual-user i's payment is $p_i(\mathbf{b}) - \mu(\mathbf{b})$. Observe that virtual-user i's true value minus its payment is exactly the coalition's utility in the original TFM. Now, imagine a "virtual auction" among a set of virtual users, where each virtual user i is the coalition of the miner and the user i. Each virtual user's strategy space is either overbidding or underbidding. Since the original TFM satisfies 1-SCP, it must be that each virtual user does not want to overbid or underbid, i.e., the

virtual auction is dominant strategy incentive compatible for each virtual user. Now, we can apply Myerson's Lemma to this virtual auction, and argue that each virtual user's payment $p_i(\mathbf{b}) - \mu(\mathbf{b})$ must satisfy the unique payment rule stipulated by Myerson's Lemma. However, since the original TFM is UIC, it must be that each user's payment $p_i(\mathbf{b})$ also satisfies the unique payment rule stipulated by Myerson's Lemma. This gives us $p_i(\mathbf{b}) - \mu(\mathbf{b}) = p_i(\mathbf{b})$, i.e., $\mu(\mathbf{b}) = 0$.

The above argument is slightly incorrect, though, since the unique payment rule of Myerson's Lemma relies on the border condition that if a user bids 0, it pays 0. When we consider the virtual auction, a virtual user's payment is of the form $p_i(\mathbf{b}) - \mu(\mathbf{b})$ — and it is not immediately clear that this quantity is 0 (even though at the end of the proof, we can see that it is indeed 0). It takes a little more work to make this intuition correct, and we give a formal proof below that makes slightly non-blackbox usage of the proof of the Myerson's Lemma — see Section 4.4 for details.

The above proves Theorem 1.4 for the randomized case. Similarly, we can now rely on Theorem 1.4 and additionally throw in the finite block size assumption to get Theorem 1.1 for the randomized case. The proof of this is a little more complicated than the deterministic case, and we defer the formal details to Section 4.5.

2.3 Incentive Compatibility under γ-Strict Utility

γ-strict utility. As observed earlier in Section 1.1.2, the current modeling approach does not charge for certain costs of cheating. Specifically, an overbid or fake transaction that is not confirmed in the present is incorrectly assumed to be free of cost. We therefore refine the model by changing the utility definition to account for this cost. As mentioned, since the exact cost is hard to predict, we define a parametrizeable utility notion called γ-strict utility, where the discount factor $\gamma \in [0, 1]$ can potentially be measured from historical data. In comparison with the utility notion introduced in Section 2.1, the only difference here is that for any overbid or fake transaction that is not confirmed in the present, we charge the strategic player γ times the worst-case cost, where the worst-case cost is the difference between the bid amount and the true value, since the strategic player may end up paying the full bid amount in a future block (of which it may not be the miner). We may assume that any fake transaction has a true value of zero.

We can define UIC, MIC, and c-SCP just like before but now using the γ-strict utility notion. The notions UIC, MIC, and c-SCP under γ-strict utility are incomparable for any $\gamma \in [0, 1]$ as shown in Appendix A.

Burning second-price mechanism. For any $\gamma \in (0, 1]$ and any $c \geq 1$, we present a TFM that achieves UIC, MIC, and c-SCP under γ-strict utility.

The burning second-price auction

Parameters:

- the block size B,
- the maximum coalition size $c \in \mathbb{N}$,
- the discount factor $\gamma \in [0, 1]$,
- $k, k' \in \mathbb{N}$ such that $k + k' = B$ and $1 \leq k' \leq \lfloor \frac{\gamma k}{c} \rfloor$, where k denotes the number of included bids that are eligible and might be confirmed with some probability, and k' is the number of included bids that are not eligible for confirmation, but are used to set the price.

Mechanism:

- *Inclusion rule.* Choose the B highest bids to include in the block, breaking ties arbitrarily. Let $(b_1, \ldots, b_B)$ denote the included bids where $b_1 \geq \cdots \geq b_B$. If the block is not fully filled, any remaining empty slot is treated as a bid of 0.
- *Confirmation rule.* Select a random subset $S \subseteq \{b_1, \ldots, b_k\}$ of size exactly $\lfloor \frac{\gamma k}{c} \rfloor$ using (trusted) on-chain randomness. The set S is confirmed and all other bids $\{b_1, \ldots, b_B\} \setminus S$ are unconfirmed.
- *Payment rule.* Any confirmed bid pays b_{k+1}. All unconfirmed bids pay nothing.
- *Miner revenue rule.* The miner is paid $\gamma \cdot (b_{k+1} + \cdots + b_{k+k'})$. Burn any remaining payment collected from the confirmed bids.

Regarding on-chain randomness. In the above burning second-price auction, the inclusion rule executed by the miner is deterministic, and only the confirmation rule that is executed by the blockchain is randomized. To implement such a mechanism in practice, we will need trusted on-chain randomness. How to sgenerate unbiased and unpredictable random coins in distributed environment has been extensively studied [CKS00,BSKN21,DKIR21]. Since such "trusted" random coins could be expensive to generate in a decentralized environment, we would ideally like to avoid them. Unfortunately, we will show later that randomness is actually necessary to get weak incentive compatibility when $c \geq 2$.

Some interesting observations. We can make a few intereseting observations about this mechanism:

1. First, the larger the coalition resistance parameter c, the smaller the number of confirmed bids $\left\lfloor \frac{\gamma k}{c} \right\rfloor$. Similarly, when γ is larger, i.e., when we are charging harsher costs for cheating, the mechanism can confirm more bids.

   In other words, both c and γ can be viewed as knobs that allow us to smoothly tradeoff the strength of incentive compatibility achieved and the efficiency of the mechanism.

2. Second, when $\gamma = 0$, i.e., when there is no cost for overbid/fake unconfirmed bids, the number of confirmed bids $\left\lfloor \frac{\gamma k}{c} \right\rfloor = 0$ — in other words, the mechanism becomes degenerate. This is consistent with the impossibility result we have shown for (strong) incentive compatibility (see Corollary 4.9).

3. Third, when $\gamma = 1$ and $c = 1$, the mechanism acutally becomes *deterministic*, since the number of confirmed bids $\left\lfloor \frac{\gamma k}{c} \right\rfloor = k$. In other words, the top k included bids are surely confirmed. We give a full description of the mechanism for this particularly interesting special case below.

   On the other hand, if $c > 1$, then the mechanism is randomized even for $\gamma = 1$. This is no co-incidence, since later, we will prove that randomness is actually necessary for $c > 1$ for any "interesting" mechanism.

The burning second-price auction: special case when $c = 1$, $\gamma = 1$

Parameters: the block size B, and $0 < k' \leq k < B$ such that $k + k' = B$, where k denotes the number of confirmed transactions per block, and k' denotes the number of unconfirmed

transactions in a block that are used to set the price and miner revenue.

Mechanism:

- Choose the B highest bids to include in the block. The highest k bids are considered confirmed, and they each pay the $(k+1)$-th price. Unconfirmed transactions, included or not, pay nothing.
- The miner is paid the sum of the $(k+1)$-th to the B-th prices (which cannot exceed the total payment by construction). All remaining payment collected from the confirmed transactions is burnt.
- If the block is not fully filled, any remaining empty slot is treated as a bid of 0.

Interestingly, for the special case $\gamma = 1$ and $c = 1$, the mechanism behaves like an ordinary second-price auction from a user's perspective. However, the miner does not collect all payment from the users. Part of the payment is burnt, and the mechanism may employ multiple "included but unconfirmed" bids to set the miner's revenue.

Theorem 2.1 (Burning second-price auction, restatement of Theorem 1.2)**.** *For any $c \geq 1$ and $\gamma \in (0, 1]$, the burning second-price auction satisfies UIC, MIC, and c-SCP under γ-strict utility.*

The proof of Theorem 2.1 is provided in Section 5.2.

2.4 Necessity of Randomness for Weak Incentive Compatibility

We present an informal roadmap for the proof of Theorem 1.3, that is, any deterministic and 2-user-friendly TFM cannot satisfy weak UIC and 2-weak-SCP simultaneously. Recall that weak incentive compatibility corresponds to the case when $\gamma = 1$. In other words, we are charging the worst-case cost for cheating, and this makes our lower bounds stronger. Henceforth, if there exists a bid vector such that the TFM confirms at least two bids, we say that the TFM is *2-user-friendly*.

Myerson's lemma holds for deterministic and weak UIC mechanisms. Recall that Myerson's Lemma holds for any UIC mechanism. Since we now are considering a more relaxed notion, namely, weak UIC, it may not be immediately clear that Myerson's Lemma still holds. Fortunately, we can prove that assuming *deterministic* and no random coins, then even weak UIC mechanisms must satisfy the requirements imposed by Myerson's Lemma (Fact 6.2). We stress that this observation is actually somewhat subtle, since it is not too clear whether Myerson's Lemma holds for *randomized* mechanisms that satisfy weak UIC.

Weak UIC + 2-weak-SCP + 2-user-friendly $\Longrightarrow$ several natural properties. Next, we establish a few natural structural properties for any deterministic, 2-user-friendly TFM that is both weak UIC and 2-weak-SCP.

1. All confirmed bids must pay the same, and thus there is a universal payment (Lemma 6.4);
2. The mechanism must confirm the highest bids where the number of confirmed bids may depend on the bid vector (Lemma 6.5); and
3. The universal payment must be at least as high as the top unconfirmed bid (Lemma 6.6). In other words, anyone bidding strictly higher than the universal payment must be confirmed.

Influence of an individual bidder. Earlier in Section 4.3 when we proved the impossibility for (strong) incentive compatibility, we used the fact that when an individual user moves its bid up or down, as long as its confirmation decision is unaffected, the user's own utility does not change. Now, due to 1-SCP, the miner's revenue should be unaffected too. This statement is not entirely true any more now that we have changed our utility definition. In particular, if an unconfirmed user increases its bid while still remaining unconfirmed, there is now an extra cost to the user. The key to proving Theorem 6.1 is to understand how fast the universal payment and miner revenue can change as we change a single user's bid. There are a few cases (stated informally below):

- **Lemma 6.3[2].** If a confirmed user changes its bid such that it is still confirmed, then the miner revenue is unaffected. This can be shown using the same argument as in Section 4.3 relying on weak UIC and 1-weak-SCP, since for a confirmed bid, the new and old utility notions coincide.

  Additionally, using 2-weak-SCP, we can show something even stronger: if there are two confirmed bids b_1 and b_2 such that $b_1 > p$ where p is the universal payment, then, b_1's confirmation status and the universal payment amount are also unaffected when b_2 changes its bid as long as it remains confirmed.

- **Lemma 6.7.** If an individual user changes its bid by Δ, then the miner utility cannot change by more than Δ. Roughly speaking, this is because even under our new utility notion, the extra cost to a user is at most Δ if it changes its bid by Δ. If the miner revenue changed by more than Δ, then the miner-user coalition has a deviating strategy that allows them to strictly gain.

- **Lemma 6.8.** If a user k increases its bid from 0 to Δ, the universal payment cannot increase by more than $\Delta/2$. Had it not been the case, then the coalition of a miner and two confirmed users can gain in the following way: replace user k's bid $b_k > 0$ with a 0-bid instead. In this way, the two colluding users each pay a lot less, and due to the earlier Lemma 6.7, the miner's revenue does not change that much. So overall, the coalition can strictly gain.

- **Lemma 6.9.** If a user k drops its bid from b_k to 0, then the universal payment cannot increase by more than b_k. Otherwise, the miner can collude with one paying user, and suppose user k's actual bid is 0, but the miner changes it to a fake bid of b_k. In this case, the paying user would pay a lot less which outweighs the cost to the miner is only b_k.

Demonstrating the contradiction (Figure 1). With the above key observations, we can finally demonstrate a contradiction, assuming that there indeed exists a deterministic, 2-user-friendly mechanism that is weak UIC and 2-weak-SCP.

1. First, we show that there exists a bid vector $\mathbf{b} = (b_1, b_2, \ldots, b_m)$ such that there are two (or more) users confirmed, and both users bid strictly higher than the payment. Note that 2-user-friendliness guarantees the existence of a vector $\mathbf{b}$ such that two users are confirmed, but does not directly guarantee that both of them bid strictly above the payment — it actually requires a bit of work to show this (which we defer to the subsequent formal presentation). Henceforth, without loss of generality, we may assume that b_1 and b_2 are the two confirmed bids, and let p be the payment. We know that $b_1 > p$, and $b_2 > p$.

2. Next, using Lemma 6.3, we can increase both b_1 and b_2 to some sufficiently large number Γ, without affecting the payment p or the miner revenue, and the resulting bid vector is $(\Gamma, \Gamma, b_3, \ldots, b_m)$.

[2] We in fact need to use this lemma to prove the aforementioned natural properties.

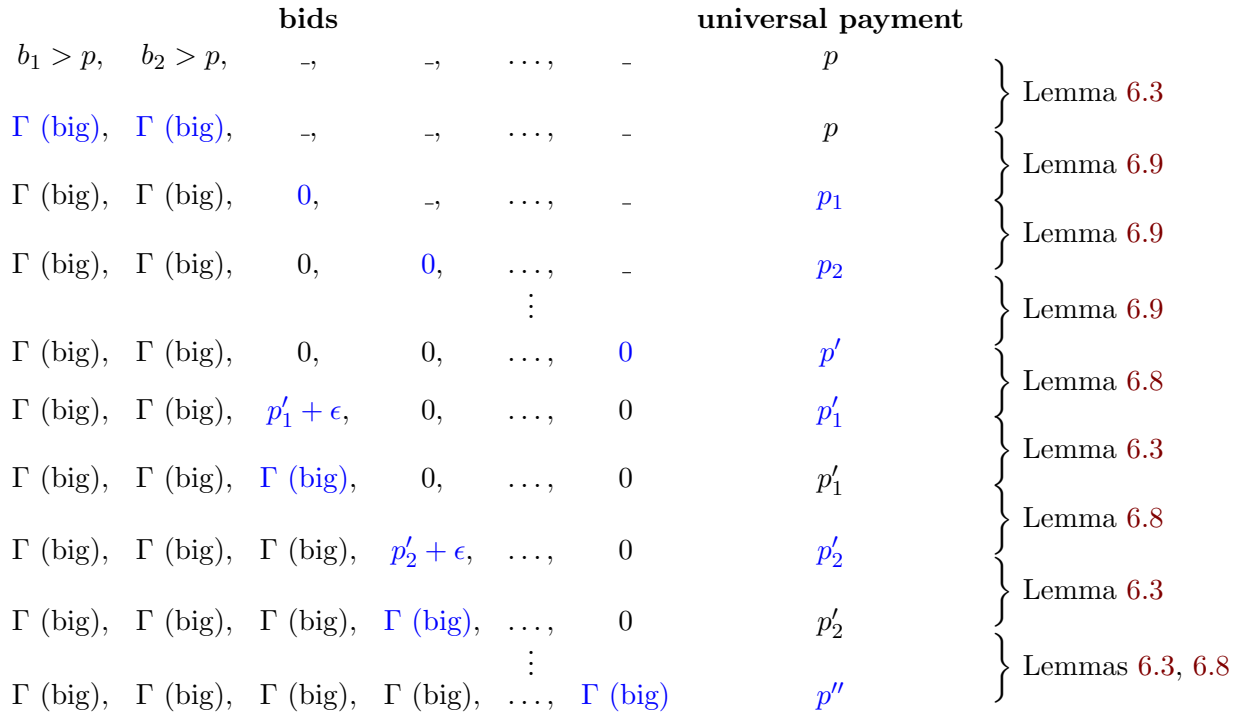

Figure 1: Proof roadmap for Theorem 6.1. We construct a sequence of bid vectors, and show that if the mechanism satisfies the desired properties, then, in the last configuration, every bid must be confirmed. Since there are more bids than the block size, we reach a contradiction. The notation "_" denotes a bid whose value we do not care about (as long as Γ is big enough w.r.t. all these values).

3. Next, we can lower $b_3, \ldots, b_m$ all to 0. Due to Lemma 6.9 and the sufficiently large choice of Γ, the increase in the payment is relatively small in comparison with Γ. This means that at the end, the first two users' bid amount Γ is still much greater than the universal payment, despite the possible increase in the universal payment. Therefore, the first two users must be still confirmed at the end (formally showing this requires a bit extra work). At this moment, we have a bid vector $(\Gamma, \Gamma, 0, 0, \ldots, 0)$, where the first two users are confirmed, and there is still a sufficiently large gap between their bid Γ and the universal payment p'.

4. Next, one by one, we shall increase the bids of users 3 through m. For each user $j \in \{3, 4, \ldots, m\}$, as we increase their bid at some rate r, the universal payment increases at rate at most $r/2$ due to Lemma 6.8. At some point, j's bid will surpass the universal payment, and at this point, due to the third natural property mentioned earlier, user j must become confirmed. Note that during this entire process, users 1 and 2 remain confirmed since their bids Γ is sufficiently large.

5. Repeating the above process, we will eventually obtain a bid vector such that all m users are confirmed. Now, as long as m is strictly greater than the block size B, we reach a contradiction — note that this is the only place where we use the finite block size assumption in the entire proof. It turns out that we can safely assume $m > B$, since if the initial vector $\mathbf{b}$ has fewer than B users, we can always append 0 bids to $\mathbf{b}$ "for free" (and showing this requires a little extra work which we defer to the subsequent formal exposition).

2.5 Additional Related Work

Transaction fee mechanism. We now review some additional related work besides the most closely reladed work EIP-1559 [Mon20] and that of Roughgarden [Rou20, Rou21b]. Specifically, we will review the transaction fee mechanisms that have been proposed, and explain which of the three properties they each fail to satisfy.

Lavi, Sattath, and Zohar [LSZ19] pointed out that today's "pay your bid" auction has resulted in complex strategic bidding behavior. In particular, when there is no congestion, users would bid almost 0, resulting in very little transaction fee revenue for the miners. To alleviate the problem, [LSZ19] suggests two alternative mechanisms, Monopolistic Price, and Random Sampling Optimal Price (RSOP), initially proposed in [GHK+06]. As [LSZ19] acknowledged, Monopolistic Price is not strictly user incentive compatible (by the classical DSIC notion), and is not even 1-side-contract resilient. For RSOP, [LSZ19] demonatrated an attack showing that it is not MIC. In fact, a slightly modified attack can also show that RSOP is not side contract resilient. Yao [Yao18] proved that although Monopolistic Price is not strictly UIC, it is nearly UIC assuming any i.i.d. distribution of the users' true values, and as the number of users goes to infinity. Further, Yao also proved a conjecture in [LSZ19] regarding the relative revenue of the two mechanisms.

Basu, Easley, O'Hara, and Sirer [BEOS19] suggested mechanism that involves paying the transfaction fees forward to some number of future blocks. Roughgarden [Rou20] simplified and analyzed their scheme, and argued that it does not satisfy any of the three properties, although it is approximately UIC when the number of users goes to infinity.

Ferreira, Moroz, Parkes, and Stern [FMPS21] suggest a modification to EIP-1559: whereas EIP-1559 approximates a first price auction in the congested regime and approximates a posted price auction in the infinite block size regime, [FMPS21] suggest to adopt a posted price mechanism no matter which regime one is in, by modifying the reserve price over time. [FMPS21]'s approach does not adopt a burn rule, and fails to satisfy even 1-side-contract-proofness.

More remotely related work. Several works in the mechanism design literature are related to our work. Akbarpour and Li [AL20] proposed a notion called credible auctions, i.e., auctions where the auctioneer does not have incentives to implement any "safe" deviations. In particular, safe deviations are for which there exists a plausible explanation. While their definition is somewhat similar in spirit to miner incentive compatibility (MIC), their modeling is incomptible with TFM. In TFM, all transactions included in the block must be visible to the public, whereas Akbarpour and Li [AL20] consider auctions where a bidder may not be able to see others' bids — and the cheating auctioneer could exploit this to explain his cheating behavior away. The elegant work of Ferreira and Weinberg [FW20] showed that using cryptographic commitments can help overcome some of the lower bound results shown by Akbarpour and Li [AL20].

A line of works also consider collusion among bidders in auctions [GL79, GH05, CM12, kCK09, MM12, DM17]. Traditional auctions like the Vickrey auction do not satisfy incentive compatibility if bidders can collude through binding side contracts. Therefore, this line of work explores under what modeling assumptions or incentive compatibility notions is it possible to resist bidder collusion. The transaction fee mechanism (TFM) line of work has not focused on user-user collusion — as mentioned earlier, user-user rendezvous is difficult to facilitate since users are ephemeral in decentralized blockchain settings.

3 Definitions

In this section, we define a transaction fee mechanism (TFM) formally, as well as incentive compatibility notions. Our modeling choice can be viewed as a generalization of that of Roughgarden's [Rou20, Rou21b]. Specifically, Roughgarden's model only cares about which transactions are eventually confirmed, but does not care about which ones are included in the block. By contrast, our modeling explicitly separates the "inclusion rule" from the "confirmation rule". Both our lower bound and upper bound will demonstrate that explicitly separating the "inclusion rule" and the "confirmation rule" is important for understanding the feasibilities and infeasibilities of transaction fee mechanism design. See also Remarks 1 and 2 for additional philosophical discussions about the modeling.

3.1 Transaction Fee Mechanism

We consider a single auction instance corresponding to the action of mining the next block. Suppose that there is a mempool containing the list of pending transactions submitted by users. We may assume that each transaction is submitted by a distinct user. We consider a single parameter environment, i.e., each user i has a true value $v_i \in \mathbb{R}$ for getting its transaction confirmed in the next block; moreover, its bid contains only a single value $b_i \in \mathbb{R}$ as well. Henceforth, we use $\mathbf{b} := (b_1, b_2, \ldots, b_m)$ to denote the vector of all bids; we also use the same notation $\mathbf{b}$ to denote the current mempool. For convenience, we often use the terms *bid* and *transaction* interchangeably, e.g., b_i can be called a bid or a transaction.

A Transaction Fee Mechanism (TFM) consists of the following (possibly randomized) algorithms:

- **Inclusion rule $\mathbf{I}(\cdot)$**: given a bid vector $\mathbf{b}$, $\mathbf{I}(\mathbf{b})$ outputs a subset of the vector $\mathbf{b}$, denoting the bids to be included in the block.
- **Confirmation rule $\mathbf{C}(\cdot)$**: given a set of bids $\mathbf{b}'$ included in the block, the confirmation rule $\mathbf{C}(\mathbf{b}')$ outputs which of these bids are confirmed.
- **Payment rule $\mathbf{P}(\cdot)$**: given a set of bids $\mathbf{b}'$ included in the block, the payment rule $\mathbf{P}(\mathbf{b}')$ outputs how much each confirmed bid pays. We assume that 1) any bid that is not confirmed pays a price of 0; and 2) each transaction pays at most what it bids.
- **Miner-revenue rule $\mathbf{M}(\cdot)$**: given a set of bids $\mathbf{b}'$ included in the block, the miner-revenue rule $\mathbf{M}(\mathbf{b}')$ outputs the total payment received by the miner for mining this block. We assume that the miner revenue does not exceed the total payment of all confirmed bids.

The inclusion rule is implemented by the miner, possibly subject to certain validity constraints enforced by the blockchain (e.g., block size limit). The other rules, including confirmation, payment, and miner-revenue rules are enforced by the blockchain itself; and they use only on-chain information.

There are a few important things to note about this definition:

1. *Included vs confirmed:* In the most general form, not all transactions included in the block must be confirmed. It could be that some transactions are included in the block to set the price, but they are not considered confirmed. For example, consider a Vickrey auction where the k highest bids are included in the block, among which the $k-1$ highest are considered confirmed, paying the k-th price. In this case, the k-th transaction is included just to set the price.

2. *Encoding the burn rule.* Not all the payment from the users will necessarily go to the miner of the block. It was pointed out earlier, e.g., in Ethereum's EIP-1559 [Mon20, Rou20, Rou21b] that in a blockchain, part to all of the payment can be burnt. In our definition, we require that the miner revenue $\mathbf{M}$ be upper bounded by the total payment from all confirmed transactions. In case the miner's revenue is strictly less than the total user payment, the difference is essentially "burnt".

In some cases, the TFM may need to perform tie breaking. For example, if there are more bids bidding the same price than the block can contain, only a subset of them will be included. Our formulation implicitly implies that the TFM is *identity agnostic*, i.e., the TFM does not use the bidders' identities for tie-breaking. In other words, if we swap two users' actions, their outcomes would be swapped too. More formally, given a bid vector $\mathbf{b} := (b_1, \ldots, b_m)$ and two different users i and j, let $x_i, x_j \in \{0, 1\}$ denote whether each user is confirmed, and let p_i, p_j denote their respective payments. Now, imagine that we swap users i and j's roles as follows. We make i bid b_j and make j bid b_i instead, and we swap i and j's positions in the bid vector. In other words, we still have the same bid vector $\mathbf{b}$ as before. However, now, the i-th coordinate now contains the bid from user j and the j-th coordinate now contains the bid from user i. In this case, the outcomes for i and j would be swapped too, that is, user i's outcome becomes x_j, p_j and user j's outcome becomes x_i, p_i.

Remark 1 (On separating the inclusion and confirmation rules)**.** In comparison, Roughgarden [Rou20, Rou21b] adopts a simpler notation that does not explicitly differentiate between the inclusion rule and the confirmation rule. Indeed, parts of our impossibility proofs do not care about this differentiation — and in these cases, we use a simplified notation that coalesces the inclusion and confirmation rules (see Section 4.1). However, our results show that it is important to explicitly separate the inclusion rule and the confirmation rule in the modeling, to further our understanding about TFMs. For example, making the inclusion rule explicit is important for proving the impossibility under *finite* block size (see Corollary 4.9). Having this distinction is also useful in constructing our upper bounds.

3.2 Strategic Behavior and Utility

Strategic player. We will consider three types of *strategic players*, 1) an individual user; 2) the miner of the current block; and 3) the miner colluding with a single user. Henceforth, we will use the term *strategic player* to refer to either a user, the miner, or the coalition of a miner and a single user.

As mentioned earlier, user-user rendezvous is much more difficult since users are ephemeral, and this is likely why this line of works [LSZ19, BEOS19, Rou20] focused on miner-user collusion (as opposed to user-user collusion). Moreover, it is easier for the miner to form a side contract with a single user rather than more users.

Strategy space. A strategic player may rely on strategic deviations to improve its utility. We first define the strategy space in the most general form, capturing all possible deviations. Our impossibility proof will rely on a much more restricted strategy space (which makes the impossibility result stronger) — we will explicitly point out the strategy space needed by our impossibility in Section 4. On the other hand, our weakly incentive compatible upper bound in Section 5 defends against the broad strategy space defined below.

A strategic player can engage in the following types of deviations or a combination thereof:

- *Bidding untruthfully.* A user or a user-miner coalition can bid untruthfully, after examining all other users' bids.
- *Injecting fake transactions.* A user, miner, or a user-miner coalition can inject fake transactions, after examining all other users' bids. Fake transactions offer no intrinsic value to anyone, and their true value is 0.
- *Strategically choosing which transactions to include in the block.* A strategic miner or a miner-user coalition may not implement the inclusion rule faithfully. It may choose an arbitrary subset of transactions from the mempool to include in the block, as long as it satisfies any block validity rule enforced by the blockchain.

Utility. The utility of the miner or a miner-user coalition is computed as the following, where S denotes the set of all real and fake transactions submitted by the miner or the miner-user coalition:

$$\text{miner revenue} + \sum_{\forall b \in S \text{ and } b \text{ confirmed}} (\text{true value of } b - \text{payment of } b)$$

The utility of a sole user is computed as as the following, where S denotes the set of all real and fake transactions submitted by the user:

$$\sum_{\forall b \in S \text{ and } b \text{ confirmed}} (\text{true value of } b - \text{payment of } b)$$

3.3 Incentive Compatibility

We would like to have mechanisms that incentivize honest behavior, i.e., no deviation of a strategic player can increase its utility. Depending on whether the strategic player is a user, the miner, or the coalition of the miner and a single user, we can define user incentive compatibility, miner incentive compatibility, and side-contract-proofness, respectively.

Definition 1 (User incentive compatibility)**.** A TFM is said to be user incentive compatible (UIC), iff the following holds: assuming that the miner implements the mechanism honestly, an individual user's (expected) utility is always maximized if it bids truthfully, no matter what the other users' bids are.

Definition 2 (Miner incentive compatibility)**.** A TFM is said to be miner incentive compatible (MIC), iff no matter what the users' bids are, the miner's (expected) utility is always maximized if it it creates the block by honestly implementing the inclusion rule.

Definition 3 (c-side-contract-proofness)**.** For any $c \in \mathbb{N}$, a TFM is said to be c-side-contract-proof (c-SCP), iff for any coalition consisting of the miner and at least one and at most c user(s), its (expected) utility is maximized when the colluding users bid truthfully and the miner plays by the book, no matter what the other users' bids are.

Remark 2 (Comparison with Roughgarden's incentive compatibility notions)**.** Our UIC and MIC notions are equivalent to Roughgarden's notions [Rou20, Rou21b]. For the SCP notion, we modify Roughgarden's offchain-agreement-proofness notion and parametrize it with the coalition size c. Note that Roughgarden's notion wants that there is no side contract that strictly benefits *every* coalition member in comparison with the honest on-chain strategy — this is equivalent to saying that the coalition cannot deviate strategically to increase their *joint* utility. If they can increase their *joint* utility there is always a way to split it off using a binding side contract such that every coalition member strictly benefits.

4 Impossibility Results

4.1 Simplified Notation and Restricted Strategy Space for our Impossibility

To rule out the existence of a UIC and 1-SCP mechanism under finite block size, our proof takes two main steps. First, we shall prove that any TFM that satisfies UIC and 1-SCP simultaneously must always have 0 miner-revenue (Theorem 4.3 and 4.7), no matter whether the block size is infinite or finite. These theorems hold even when the strategic player is onfined to a very restricted strategy space: assuming that the miner always implements the mechanism faithfully; however, either an individual user or a user colluding with the miner may bid untruthfully. In the second part of the proof, we additionally throw in the finite block size restriction which leads to the stated impossibility result (Corollary 4.9).

Simplified notations for deterministic mechanisms. We can simplify the notation in the first part of our proof, since this part makes use of a very restricted strategy space as mentioned above. Instead of using the full tuple $(\mathbf{I}, \mathbf{C}, \mathbf{P}, \mathbf{M})$ to denote the TFM, we will use the following simplified notation:

1. **Allocation rule x**: given a bid vector $\mathbf{b} := (b_1, \ldots, b_m) \in \mathbb{R}^m$, the allocation rule $\mathbf{x}(\mathbf{b})$ outputs a vector $(x_1, x_2, \ldots, x_m) \in \{0,1\}^m$, indicating whether each transaction (i.e., bid) in $\mathbf{b}$ is *confirmed* in the next block.

2. **Payment rule p**: given a bid vector $\mathbf{b} := (b_1, \ldots, b_m) \in \mathbb{R}^m$, the payment rule $\mathbf{p}(\mathbf{b})$ outputs a vector $(p_1, p_2, \ldots, p_m) \in \mathbb{R}^m$, indicating the price paid by each transaction in $\mathbf{b}$. It is guaranteed that $p_i \leq b_i$ for $i \in [m]$, i.e., a user never pays more than its bid.

3. **Miner-revenue** rule μ: given a bid vector $\mathbf{b} := (b_1, \ldots, b_m) \in \mathbb{R}^m$, the miner-revenue rule $\mu(\mathbf{b})$ outputs a single value in $\mathbb{R}$ denoting the amount paid to the miner.

More specifically, one can view:

- $\mathbf{x}$ as the composition of the inclusion rule $\mathbf{I}$ and the blockchain-enforced confirmation rule $\mathbf{C}$;
- $\mathbf{p}$ as the composition of the inclusion rule $\mathbf{I}$ and the blockchain-enforced payment rule $\mathbf{P}$; and
- μ as the composition of the inclusion rule $\mathbf{I}$ and the blockchain-enforced miner-revenue rule $\mathbf{M}$.

Additional notations. For convenience, we often use the notation $x_i(\mathbf{b})$ and $p_i(\mathbf{b})$ to denote whether the i-th transaction in $\mathbf{b}$ is confirmed in the next mined block, and what price it actually pays. We assume that if $x_i(\mathbf{b}) = 0$, then, $p_i(\mathbf{b}) = 0$ — in other words, if the i-th transaction is not confirmed in the next block, then the i-th user pays nothing. Let $\mathbf{b} = (b_1, b_2, \ldots, b_m)$ be a bid vector. We often use the notation $\mathbf{b}_{-i} = (b_1, b_2, \ldots, b_{i-1}, b_{i+1}, \ldots, b_m)$ to denote everyone except user i's bids; and the notation $(\mathbf{b}_{-i}, b_i)$ and $\mathbf{b}$ are used interchangeably.

Notations for randomized mechanisms. We use the same notations $(\mathbf{x}, \mathbf{p}, \mu)$ to denote a randomized mechanism but their meaning is modified as follows. The allocation rule now outputs the probability that each bid is confirmed, that is, $x_i(\mathbf{b}) \in [0,1]$ is the probability that user i's bid is confirmed given the included bids are $\mathbf{b}$. Also, we view $p_i(\mathbf{b})$ as the expected payment of user i and $\mu(\mathbf{b})$ as the expected miner-revenue.

We say that a TFM enjoys *non-trivial miner revenue* iff $\mu(\cdot)$ is not the constant 0 function, i.e., the miner sometimes can receive positive revenue.

4.2 Preliminary: Myerson's Lemma

If a single-parameter TFM satisfies UIC (even when the user's strategy space is restricted only to untruthful bidding), the mechanism's allocation rule $\mathbf{x}$ and payment rule $\mathbf{p}$ must satisfy the famous Myerson's Lemma [Mye81]. Specifically, we only need a special case of Myerson's Lemma: the mechanism can be randomized, and each user's bid is either confirmed or unconfirmed. In this case, the allocation rule x_i returns a real number in $[0, 1]$, which is the probability that user i's bid is confirmed. Additionally, p_i is the expected payment of user i. Myerson's Lemma implies the following:

Lemma 4.1 (Myerson's Lemma)**.** *Let $(\mathbf{x}, \mathbf{p}, \mu)$ be a single-parameter TFM that is UIC. Then, it must be that*

1. *The allocation rule $\mathbf{x}$ is monotone, where monotone is defined as follows. Consider $\mathbf{b} := (b_1, \ldots, b_m)$, and let $\mathbf{b}_{-i}$ be the vector obtained when we remove b_i from $\mathbf{b}$. An allocation rule $\mathbf{x}$ is said to be monotone iff for any $\mathbf{b} := (b_1, \ldots, b_m)$, and any $b_i' > b_i$, it must be that $x_i(\mathbf{b}_{-i}, b_i') \geq x_i(\mathbf{b}_{-i}, b_i)$.*

2. *The payment rule $\mathbf{p}$ is defined as follows. For any user i, bids $\mathbf{b}_{-i}$ from other users, and bid b_i from user i, it must be*

$$p_i(\mathbf{b}_{-i}, b_i) = b_i \cdot x_i(\mathbf{b}_{-i}, b_i) - \int_0^{b_i} x_i(\mathbf{b}_{-i}, t)dt. \tag{1}$$

Deterministic special case. When the mechanism is deterministic, the allocation rule x_i returns either 0 or 1. In this case, the unique payment rule can be simplified as

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

Conceptually, user i only needs to pay the minimal price which makes its bid confirmed.

To prove our impossibility for randomized mechanisms, we need to open up Myerson's Lemma and use the following technical lemma that is used in the proof of Myerson's Lemma. More specifically, the proof of Myerson's Lemma showed that if a mechanism is UIC, then a user i's payment must satisfy the following inequality (also called a "payment sandwich") where the allocation rule $\mathbf{x}$ is monotone:

$$r \cdot \big(x_i(\mathbf{b}_{-i}, r') - x_i(\mathbf{b}_{-i}, r)\big) \leq p(\mathbf{b}_{-i}, r') - p(\mathbf{b}_{-i}, r) \leq r' \cdot \big(x_i(\mathbf{b}_{-i}, r') - x_i(\mathbf{b}_{-i}, r)\big)$$

Assume that the above payment sandwich holds for a non-decreasing function $x_i(\mathbf{b}_{-i}, \cdot)$, and moreover, $p(\mathbf{b}_{-i}, 0) = 0$, then Myerson showed that the payment rule is of a unique form as shown in Equation (1). To prove this, Myerson essentially proved the following technical lemma.

Lemma 4.2 (Technical lemma implied by the proof of Myerson's Lemma [Mye81, Har])**.** *Let $f(z)$ be a non-decreasing function. Suppose that $z \cdot (f(z') - f(z)) \leq g(z') - g(z) \leq z' \cdot (f(z') - f(z))$ for any $z' \geq z \geq 0$, and moreover, $g(0) = 0$. Then, it must be that*

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

4.3 Deterministic Mechanisms: UIC + 1-SCP $\Longrightarrow$ Zero Miner Revenue

As a warmup, we first prove a lower bound for deterministic mechanisms. Then, in Section 4.4, we generalize the proof to randomized mechanisms. The following theorem states that no deterministic TFM with non-trivial miner revenue can achieve UIC and 1-SCP simultaneously, no matter whether the block size is finite or infinite.

Theorem 4.3 (Deterministic TFM: UIC + 1-SCP $\Longrightarrow$ 0 miner revenue)**.** *There is no deterministic TFM with non-trivial miner revenue that achieves UIC and 1-SCP at the same time. Moreover, the theorem holds no matter whether the block size is finite or infinite.*

The rest of this section will be dedicated to proving the theorem. The following claim states that if an individual user changes its bid in a way that does not affect whether it is confirmed, then the miner's revenue should not change.

Claim 4.4. *Suppose that a TFM $(\mathbf{x}, \mathbf{p}, \mu)$ satisfies UIC and 1-SCP. Suppose that $x_i(\mathbf{b}_{-i}, b_i) = x_i(\mathbf{b}_{-i}, b_i')$. Then, it must be that $\mu(\mathbf{b}_{-i}, b_i) = \mu(\mathbf{b}_{-i}, b_i')$.*

Proof. Since the TFM satisfies UIC, the tuple $(\mathbf{x}, \mathbf{p})$ satisfies Myerson's Lemma. We know that $x_i(\mathbf{b}_{-i}, b_i) = x_i(\mathbf{b}_{-i}, b_i') = 0$ or $x_i(\mathbf{b}_{-i}, b_i) = x_i(\mathbf{b}_{-i}, b_i') = 1$. In the former case, $p_i(\mathbf{b}_{-i}, b_i) = p_i(\mathbf{b}_{-i}, b_i') = 0$. In the latter case, by Myerson's Lemma, no matter whether user i's bid is b_i or b_i', its payment equals the minimal amount it bids that still allows the transaction to be confirmed. Therefore, in either case, we have that $p_i(\mathbf{b}_{-i}, b_i) = p_i(\mathbf{b}_{-i}, b_i')$.

Suppose that $\mu_i(\mathbf{b}_{-i}, b_i) \neq \mu_i(\mathbf{b}_{-i}, b_i')$. Without loss of generality, we may assume that $\mu_i(\mathbf{b}_{-i}, b_i') > \mu_i(\mathbf{b}_{-i}, b_i)$. In this case, imagine that all users' true values are represented by the vector $(\mathbf{b}_{-i}, b_i)$. Now, consider the coalition of the miner and user i. If user i bids b_i truthfully, the coalition's joint utility is $U := \mu(\mathbf{b}_{-i}, b_i) + b_i - p_i(\mathbf{b}_{-i}, b_i)$. However, if user i' strategically bids b_i' instead, the coalition's joint utility is $U' := \mu(\mathbf{b}_{-i}, b_i') + b_i - p_i(\mathbf{b}_{-i}, b_i')$. Since $p_i(\mathbf{b}_{-i}, b_i) = p_i(\mathbf{b}_{-i}, b_i')$, $U' - U = \mu(\mathbf{b}_{-i}, b_i') - \mu(\mathbf{b}_{-i}, b_i) > 0$. This shows that the coalition can gain if user i bids untruthfully, thus violating 1-SCP. □

Lemma 4.5. *Let $(\mathbf{x}, \mathbf{p}, \mu)$ be any TFM with non-trivial miner revenue. Then, there exists a bid vector $\mathbf{b} = (b_1, \ldots, b_m)$ and a user i such that $\mu(\mathbf{b}_{-i}, 0) < \mu(\mathbf{b})$.*

Proof. Since the mechanism enjoys non-trivial miner revenue, there exists a bid vector $\mathbf{b}^{(0)} = (b_1, \ldots, b_m)$ such that $\mu(\mathbf{b}^{(0)}) > 0$. Now, consider the following sequence of bid vectors: for $i \in [m]$, let $\mathbf{b}^{(i)}$ be obtained by setting the first i coordinates of $\mathbf{b}^{(0)}$ to 0. Observe that $\mathbf{b}^{(m)} = \mathbf{0}$.

Since a user can pay at most its bid, we have $\mu(\mathbf{b}) \leq |\mathbf{p}(\mathbf{b})|_1 \leq |\mathbf{b}|_1$ for any bid vector $\mathbf{b}$. Therefore, $\mu(\mathbf{b}^{(m)}) \leq |\mathbf{b}^{(m)}|_1 = 0$. Since $\mu(\mathbf{b}^{(0)}) > 0$, there exists an $i \in [m-1]$ such that $0 = \mu(\mathbf{b}^{(i)}) < \mu(\mathbf{b}^{(i-1)})$. □

Lemma 4.6. *If there exists a bid vector $\mathbf{b} = (b_1, \ldots, b_m)$ and a user i such that $\mu(\mathbf{b}_{-i}, 0) < \mu(\mathbf{b})$, then the TFM $(\mathbf{x}, \mathbf{p}, \mu)$ is either not UIC or not 1-SCP.*

Proof. For the sake of reaching a contradiction, suppose that $(\mathbf{x}, \mathbf{p}, \mu)$ is both UIC and 1-SCP. By Myerson's Lemma, we have $x_i(\mathbf{b}_{-i}, 0) \leq x_i(\mathbf{b})$. Due to Claim 4.4, it must be $x_i(\mathbf{b}_{-i}, 0) = 0$ and $x_i(\mathbf{b}) = 1$.

Let $\Delta = \mu(\mathbf{b}) - \mu(\mathbf{b}_{-i}, 0) > 0$ and $\epsilon = \frac{1}{2} \cdot \min(\Delta, p_i(\mathbf{b})) > 0$. Imagine that everyone else except user i is bidding $\mathbf{b}_{-i}$, and user i's true value is $v_i = p_i(\mathbf{b}) - \epsilon > 0$. Due to Myerson's Lemma, since $v_i < p_i(\mathbf{b})$, user i's bid would be unconfirmed if it were to bid truthfully. In this case, by Claim 4.4, the miner's utility is $\mu(\mathbf{b}_{-i}, 0)$ and user i's utility is zero.

However, the miner can sign a side contract and ask user i to bid $p_i(\mathbf{b})$ instead. By Myerson's Lemma, at this moment, user i's bid will indeed be confirmed. By Claim 4.4, the miner's utility is now $\mu(\mathbf{b})$ and user i's utility is now $v_i - p_i(\mathbf{b}) = -\epsilon$. Consequently, their joint utility becomes $\mu(\mathbf{b}) - \epsilon$, which has increased by $\Delta - \epsilon > 0$. This violates 1-SCP. □

Proof of Theorem 4.3. Theorem 4.3 follows directly from the combination of Lemma 4.5 and Lemma 4.6.

4.4 Randomized Mechanisms: UIC + 1-SCP $\Longrightarrow$ Zero Miner Revenue

We now generalize Theorem 4.3 to even randomized mechanisms. In a randomized TFM, the random coins could come from either the miner or the blockchain itself. Since we are proving an impossibility, without loss of generality, we may assume that the blockchain comes with an unpredictable random source. Our impossibility result actually does not care where the random coins come from.

Earlier in Section 2.2, we presented the intuition for this impossiblity. Therefore, below, we directly jump to the formal description.

Notations for randomized mechanisms. Recall that for randomized mechanisms, the allocation rule now outputs the probability that each bid is confirmed; that is, $x_i(\mathbf{b}) \in [0,1]$ is the probability that user i's bid is confirmed given the included bids are $\mathbf{b}$. Also, $p_i(\mathbf{b})$ is now the expected payment of user i and $\mu(\mathbf{b})$ is the expected miner-revenue.

For convenience, we define the following quantity:

$$\pi_{\mathbf{b}_{-i}}(r) = p_i(\mathbf{b}_{-i}, r) - \mu(\mathbf{b}_{-i}, r)$$

One can think of $\pi_{\mathbf{b}_{-i}}(r)$ as a virtual user i's payment in the virtual auction (see Section 2.2). The following theorem is a generalization of Theorem 4.3 to even randomized mechanisms.

Theorem 4.7 (Randomized TFM: UIC + 1-SCP $\Longrightarrow$ 0 miner revenue)**.** *There is no randomized TFM with non-trivial miner revenue that achieves UIC and 1-SCP at the same time. Moreover, the theorem holds no matter whether the block size is finite or infinite.*

We will now prove this theorem. First, we introduce a useful lemma.

Lemma 4.8. *Let $(\mathbf{x}, \mathbf{p}, \mu)$ be any randomized TFM. If $(\mathbf{x}, \mathbf{p}, \mu)$ is 1-SCP, then, for any bid vector $\mathbf{b}$, user i, and r, r' such that $r < r'$, it must be*

$$r \cdot \big(x_i(\mathbf{b}_{-i}, r') - x_i(\mathbf{b}_{-i}, r)\big) \le \pi_{\mathbf{b}_{-i}}(r') - \pi_{\mathbf{b}_{-i}}(r) \le r' \cdot \big(x_i(\mathbf{b}_{-i}, r') - x_i(\mathbf{b}_{-i}, r)\big).$$

Proof. First, we prove the case of $r \cdot (x_i(\mathbf{b}_{-i}, r') - x_i(\mathbf{b}_{-i}, r)) \le \pi_{\mathbf{b}_{-i}}(r') - \pi_{\mathbf{b}_{-i}}(r)$. For the sake of reaching a contradiction, suppose there exists a vector $\mathbf{b}$, a user i and $r < r'$ such that

$$r \cdot \big(x_i(\mathbf{b}_{-i}, r') - x_i(\mathbf{b}_{-i}, r)\big) > \pi_{\mathbf{b}_{-i}}(r') - \pi_{\mathbf{b}_{-i}}(r). \tag{2}$$

Imagine that the real bid vector is $(\mathbf{b}_{-i}, r)$ and user i's true value is r. If they do not have a side contract, the miner's expected utility is $\mu(\mathbf{b}_{-i}, r)$ and user i's expected utility is $r \cdot x_i(\mathbf{b}_{-i}, r) - p_i(\mathbf{b}_{-i}, r)$. However, the miner can sign a contract with user i and ask user i to bid r' instead. In this case, the miner's expected utility becomes $\mu(\mathbf{b}_{-i}, r')$ and user i's expected utility becomes $r \cdot x_i(\mathbf{b}_{-i}, r') - p_i(\mathbf{b}_{-i}, r')$ since the user's true value is still r. By Eq.(2), their joint expected utility increases by $r \cdot (x_i(\mathbf{b}_{-i}, r') - x_i(\mathbf{b}_{-i}, r)) - (\pi_i(\mathbf{b}_{-i}, r') - \pi_i(\mathbf{b}_{-i}, r)) > 0$. This violates 1-SCP.

The other case $\pi_{\mathbf{b}_{-i}}(r') - \pi_{\mathbf{b}_{-i}}(r) \leq r' \cdot (x_i(\mathbf{b}_{-i}, r') - x_i(\mathbf{b}_{-i}, r))$ can be proven by a similar argument, so we only sketch the proof. Suppose the inequality does not hold, that is, suppose that $\pi_{\mathbf{b}_{-i}}(r') - \pi_{\mathbf{b}_{-i}}(r) > r' \cdot (x_i(\mathbf{b}_{-i}, r') - x_i(\mathbf{b}_{-i}, r))$. Imagine that the real bid vector is $(\mathbf{b}_{-i}, r')$ and user i's true value is r'. The miner can sign a contract with user i and ask user i to bid r instead. In this case, their joint expected utility increases by $\pi_{\mathbf{b}_{-i}}(r') - \pi_{\mathbf{b}_{-i}}(r) - r' \cdot (x_i(\mathbf{b}_{-i}, r') - x_i(\mathbf{b}_{-i}, r)) > 0$. This violates 1-SCP. □

Proof of Theorem 4.7 We now continue with the proof of Theorem 4.7. Consider the following quantity:

$$\widetilde{\pi}_{\mathbf{b}_{-i}}(r) = p_i(\mathbf{b}_{-i}, r) - \mu(\mathbf{b}_{-i}, r) - (p_i(\mathbf{b}_{-i}, 0) - \mu(\mathbf{b}_{-i}, 0))$$

By Lemma 4.8, and the fact that definition of $\widetilde{\pi}_{\mathbf{b}_{-i}}(r)$ and $\pi_{\mathbf{b}_{-i}}(r)$ differs by only a fixed constant, it must be that

$$r \cdot \left(x_i(\mathbf{b}_{-i}, r') - x_i(\mathbf{b}_{-i}, r)\right) \leq \widetilde{\pi}_{\mathbf{b}_{-i}}(r') - \widetilde{\pi}_{\mathbf{b}_{-i}}(r) \leq r' \cdot \left(x_i(\mathbf{b}_{-i}, r') - x_i(\mathbf{b}_{-i}, r)\right). \tag{3}$$

Now, observe that the above expression exactly agrees with the "payment sandwich" in the proof of Myerson's Lemma [Mye81, Har]. Furthermore, we have that $\widetilde{\pi}_{\mathbf{b}_{-i}}(0) = 0$ by definition; and $\mathbf{x}$ must be monotone because the TFM is UIC and satisfies Myerson's Lemma. Due to Lemma 4.2, it must be that $\widetilde{\pi}_{\mathbf{b}_{-i}}(\cdot)$ obeys the unique payment rule specified by Myerson's Lemma, that is,

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

On the other hand, since the TFM is UIC, its payment rule itself must also satisfy the same expression, that is,

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

We therefore have that

$$\widetilde{\pi}_{\mathbf{b}_{-i}}(r) = p_i(\mathbf{b}_{-i}, r) - \mu(\mathbf{b}_{-i}, r) - (p_i(\mathbf{b}_{-i}, 0) - \mu(\mathbf{b}_{-i}, 0)) = p_i(\mathbf{b}_{-i}, r)$$

In other words, $\mu(\mathbf{b}_{-i}, r) = \mu(\mathbf{b}_{-i}, 0) - p(\mathbf{b}_{-i}, 0)$, which is a constant that is independent of user i's bid r when $\mathbf{b}_{-i}$ is fixed.

We now argue that this actually implies $\mu(\mathbf{b}_{-i}, r) = 0$, i.e., a possibly randomized TFM that is UIC and 1-SCP must always have 0 miner revenue. Suppose this is not true, i.e., suppose there exists a randomized TFM with non-trivial miner revenue $(\mathbf{x}, \mathbf{p}, \mu)$ that is UIC and 1-SCP. Since it enjoys non-trivial miner revenue, there exists a bid vector $\mathbf{b}^{(0)} = (b_1, \ldots, b_m)$ such that $\mu(\mathbf{b}^{(0)}) > 0$. Now, consider the following sequence of bid vectors: for $i \in [m]$, let $\mathbf{b}^{(i)}$ be obtained by setting the first i coordinates of $\mathbf{b}^{(0)}$ to 0. Observe that $\mathbf{b}^{(m)} = \mathbf{0}$.

Recall that we have argued for a fixed $\mathbf{b}_{-i}$, the miner revenue $\mu(\mathbf{b}_{-i}, \cdot)$, is a constant function independent of user i's bid. Thus, $\mu(\mathbf{b}^{(i-1)}) = \mu(\mathbf{b}^{(i)})$ for all $i \in [m]$. Consequently, we obtain $\mu(\mathbf{b}^{(0)}) = \mu(\mathbf{b}^{(m)})$. However, users can only pay their bids at most, so we have $\mu(\mathbf{b}^{(m)}) \leq |\mathbf{b}^{(m)}|_1 = 0$. This contradicts the assumption that $\mu(\mathbf{b}^{(0)}) > 0$.

4.5 UIC + 1-SCP + Finite Block Size $\Longrightarrow$ Impossibility

Theorem 4.7 holds no matter whether the block size is finite or infinite. In this section, we prove a corollary stating that if the block size is finite, then no non-trivial TFM can satisfy UIC and 1-SCP simultaneously. Particularly, assuming finite block size, the only TFM that satisfies both

UIC and 1-SCP is the one that never confirms any transaction, and always pays the miner nothing. This corollary holds assuming the following strategic behavior is possible: an individual user or a user colluding with the miner can bid untruthfully; and the miner can arbitrarily decide which transactions to include in the block (as long as it respects the block's validity constraint).

Corollary 4.9 (UIC + 1-SCP + finite block size $\Longrightarrow$ impossibility)**.** *Suppose the size of a block is finite. Then, the only randomized TFM* $(\mathbf{x}, \mathbf{p}, \mu)$ *that satisfies both UIC and 1-SCP is the trivial mechanism that never confirms any transaction no matter how users bid, and always pays the miner nothing.*

Proof. For the sake of reaching a contradiction, suppose that there is a non-trivial TFM that satisfies UIC and 1-SCP. By Theorem 4.7, any TFM that satisfies both UIC and 1-SCP must have constant zero miner revenue. Henceforth, we may assume that the miner always gets zero payment.

Let B denote an upper bound on the block size. Since the TFM is non-trivial, there exists a bid vector $\mathbf{b} = (b_1, \ldots, b_m)$ and a user i^* such that $x_{i^*}(\mathbf{b}) > 0$. Now, let ϵ be any positive number, let $n > \frac{B \cdot (b_{i^*} + \epsilon)}{x_{i^*}(\mathbf{b}) \cdot \epsilon}$ be a sufficiently large integer. Consider another bid vector $\mathbf{b}' = (b_1, \ldots, b_m, b_{m+1}, \ldots, b_{m+n})$ where $b_j = b_{i^*} + \epsilon$ for all $j \in [m+1, m+n]$. Imagine that the real bid vector is actually $\mathbf{b}'$ and each user bids truthfully, i.e., user j's true value is $v_j = b_j$ for all $j \in [m+n]$. Since the block size is at most B, there must be a user $j \in [m+1, m+n]$ who bids b_j is included with probability at most $B/n < \frac{x_{i^*}(\mathbf{b}) \cdot \epsilon}{b_{i^*} + \epsilon}$.

Consider the coalition of the miner and user j. If everyone bids truthfully and the miners runs the honest mechanism, then their joint utility is strictly less than $b_j \cdot \frac{B}{n} < (b_{i^*} + \epsilon) \cdot \frac{x_{i^*}(\mathbf{b}) \cdot \epsilon}{b_{i^*} + \epsilon} = x_{i^*}(\mathbf{b}) \cdot \epsilon$ — since the miner always gets 0 revenue and user j's utility is upper bounded by $b_j \cdot \frac{B}{n}$. However, the miner can sign a contract with user j. The contract asks user j to change the bid from b_j to b_{i^*}, and the miner pretends that the actual bid vector is $\mathbf{b}$, where the coordinate b_{i^*} actually comes from user j. In this case, the coalition's joint utility is $(v_j - b_{i^*}) \cdot x_{i^*}(\mathbf{b}) = \epsilon \cdot x_{i^*}(\mathbf{b})$. Therefore, the coalition can increase its expected utility by deviating. This violates 1-SCP. □

5 Rethinking the Incentive Compatibility Notions

So far in our impossibility results, we have assumed it is free of charge for a strategic player to inject a fake transaction or overbid (i.e., bid higher than its true value), as long as the offending transaction is not confirmed in the present block. Not only so, in fact, the same model was implicitly or explicitly adopted in earlier works on transaction fee mechanism design [Rou20, Rou21b, BEOS19], too.

Such a model, however, may be overly draconian, since there is actually some cost associated with cheating that the existing model does not charge. In reality, the TFM is not a standalone auction, it is repeatedly executed as blocks get confirmed. Although an overbid or fake transaction need not pay fees to the present miner if it is not confirmed, in real life, any transaction that has been submitted to the network cannot be retracted. Therefore, the offending transaction could be confirmed and paying fees in a future block (e.g., paid to a different miner or simply burnt). Consequently, a risk-averse miner-user coalition may be deterred from such deviations for fear of losing the offending transaction's fees to a future block.

Therefore, a natural and interesting question is:

If we fix the existing model and more carefully account for the cost of such cheating, can this help us circumvent the impossibility results?

One challenge we are faced with, however, is the difficulty of accurately characterizing the cost of such cheating. If an overbid or fake transaction is confirmed in a future block, it is hard for us to predict how much the offending transaction will end up paying, since the payment amount may not be equal to the bid, and the payment amount depends on the environment (e.g., the other bids), as well as the mechanism itself.

Despite this difficulty, we still want to understand whether this direction is worth exploring. A reasonable approach is to start by asking what is the worst-case cost. Once we understand what is the worst-case cost, we can consider how to define a more general, parametrized cost model.

1. *Worst-case cost.* A worthwhile first step is to consider the *worst-case cost* for the aforementioned deviation. Specifically, whenever a strategic player injects a fake transaction or overbids and the offending transaction cannot be confirmed in the present block, the strategic player assumes the worst case scenario, i.e., the offending transaction can end up paying fees as high as its bid in the future.

   Assuming the worst-case cost is useful in several ways. First, it is *useful for proving lower bounds*. If we can prove lower bounds even for the worst-case cost, it would directly imply lower bounds if in reality, the cost is actually smaller than the worst case. Second, assuming the worst-case cost is also equivalent to considering strategic players who are *paranoid* — they only want to deviate if they will surely benefit, and there is no possible scenario in which they will lose. In other words, we are asking whether there is a mechanism that can at least discourage such paranoid players from deviating.

2. *General, parametrized cost model.* As mentioned, it is challenging to accurately capture or predict the cost of overbid or fake transactions that are unconfirmed in the present. In practice, however, one might be able to measure the cost of such cheating from historical data. This motivates a more generalized cost model, where we assume that there is some discount factor $\gamma \in [0, 1]$, and the cost of such cheating is actually γ times the worst-case cost.

5.1 Defining γ-Strict Utility

As we argued, the utility notions in prior work ignore certain costs associated with cheating. We therefore define a more refined utility notion that charges such cost parametrized by a "strictness" parameter $\gamma \in [0, 1]$. In other words, when $\gamma = 1$, we are charging the worst-case cost, and equivalently, we are asking whether there are incetive compatible TFMs against *paranoid* players who only want to deviate if there is a sure gain and no risk of losing. We will also be using $\gamma = 1$ to prove lower bounds, and this gives stronger lower bound results. When $\gamma = 0$, we are charging no cost — in this case, our new incentive compatibility definitions would be equivalent to the old notions in Section 3.

Recall that the term "strategic player" can refer to a user, a miner, or a miner-user coalition. An *offending transaction* is one whose bid exceeds the transaction's true value: it can be an untruthful bid or an injected fake transaction, since we may assume that a fake transaction's true value is 0. In the *worst-case* scenario, an offending transaction that is not confirmed in the present block may be charged a transaction fee equal to its full bid, when it is confirmed in a future block (possibly mined by a different miner). Let v be the true value of the offending transaction (and $v = 0$ if the offending transaction is fake), and let $b \geq v$ be the bid value. Therefore, in the worst-case scenario, the offending transaction can cost $b - v$ in utility, due to losing fees to a future block.

In practice, if we can measure the actual cost from historical data, we may be able to learn a parameter $\gamma \in [0, 1]$, and model the actual cost as γ times the worst-case cost, that is, $\gamma \cdot (b - v)$.

γ-strict utility. We now formally define the utility function of a strategic player:

γ-strict utility

- If the strategic player includes the miner, then let $u \leftarrow \mu$ where μ is the miner's revenue in the present block; else let $u \leftarrow 0$.
- For any real or fake transaction the strategic player has submitted with true value v and a bid of b:
  - if the transaction is confirmed in the present block, let $u \leftarrow u + v - p$ where p denotes its payment.
  - if the transaction is *not* confirmed in the present block and moreover $b > v$, then let $u \leftarrow u - \gamma \cdot (b - v)$. See also Remark 3.
- Output the final utility u.

Definition 4 (Incentive compatibility under γ-strict utility)**.** Let $X \in \{\text{UIC, MIC, } c\text{-SCP}\}$. We can now define X under γ-strict utility just like in Definitions 1, 2, and 3, respectively, except that now we adopt the aforementioned γ-strict utility.

Definition 5 (Weak incentive compatibility)**.** For convenience, for the special case $\gamma = 1$, we also refer to our incentive compatibility notions as *weak* incentive compatibility. More specifically, we use the following aliases:

$$\text{weak UIC} = \text{UIC under 1-strict utility}$$
$$\text{weak MIC} = \text{MIC under 1-strict utility}$$
$$c\text{-weak-SCP} = c\text{-SCP under 1-strict utility}$$

Remark 3. Since the miner has the ability to include an arbitrary set of transactions in the block, without loss of generality, we may assume that the following deviations never take place since they do not help the miner or the miner-user coalition:

1. the miner or the miner-user coalition never bids untruthfully for any transaction *not* included in the block;
2. miner or the miner-user coalition never injects a fake transaction that is *not* included in the block.

Therefore, one can equivalently view our new utility definition as only charging an additional cost for overbid or fake transactions that are unconfirmed but included in the block. Note that any transaction included in the block must have been broadcast to the network and cannot be retracted.

5.2 Burning Second-Price Mechanism

Earlier in Section 2.3, we presented the burning second-price mechanism which can be parametrized with any $\gamma \in (0, 1]$ and $c \geq 1$. We now prove Theorem 2.1, that is, for any $c \geq 1$ and $\gamma \in (0, 1]$, the burning second-price auction satisfies UIC, MIC, and c-SCP under γ-strict utility.

We prove the properties one by one. Throughout this proof, we assume the γ-strict utility notion. We may assume that $\gamma \in (0, 1]$, since if $\gamma = 0$, the burning second-price auction always confirms nothing and it trivially satisfies all these properties.

UIC. According to the utility definition for the user, any injected fake transaction cannot lead to an increase in the user's utility. Therefore, we may assume that the user does not inject any fake transactions, and the only strategic behavior is bidding untruthfully.

Let $\mathbf{b} = (b_1, \cdots, b_m)$ be an arbitrary bid vector, where $b_1 \geq \cdots \geq b_m$ and user i bids truthfully (i.e. $b_i = v_i$). Suppose $i \geq k+1$ and thus $b_i \leq b_{k+1}$. If user i bids honestly, its utility is 0 since it is unconfirmed. Imagine that user i changes its bids to b'_i. There are two cases. First, user i's new bid b'_i is still not ranked among the top k (possibly after the tie-breaking). In this case, its utility is either zero if $b'_i < v_i$ or negative if $b'_i > v_i$. Second, the new bid b'_i is now ranked among the top k. We have $b'_i \geq b_{k+1}$. Further, user i's utility becomes

$$(1 - \frac{\gamma}{c}) \cdot \gamma(b'_i - b_i) + \frac{\gamma}{c} \cdot (b_i - b_k),$$

where the first term captures the cost if b'_i is not confirmed, and the second term captures the cost if b'_i is confirmed. Since $b'_i - b_i < 0$ and $b_i - b_k \leq 0$, its utility only decreases.

The case of $i \leq k$ can be shown by a similar argument. In this case, if user i bids honestly, it is among the top k, and its utility is at least 0. Now, imagine user i changes its bid to b'_i. There are two cases. First, if b'_i cause user i to be no longer among the top k, then its utility is 0. Second, with the new bid b'_i, user i is still among the top k. If it underbids its utility is the same as bidding honestly. If it overbids, its utility is the same as bidding honestly conditioned on it is confirmed, and its utility decreases by $\gamma(b'_i - b_i)$ conditioned on it is not confirmed.

MIC. A miner has two kinds of strategies to deviate from honest behavior: not to choose the highest bids and to inject fake bids. Without loss of generality, we assume that the miner chooses the included bids first, and then replaces some of the real bids with fake bids. We may also assume that all injected fake bids are included in the block. We will show that both steps would not increase the miner's utility. Let $(c_1, \ldots, c_B)$ be the highest B bids in the bid vector, where $c_1 \geq \cdots \geq c_B$. The miner's revenue is $\gamma(c_{k+1} + \cdots + c_B)$. Now, suppose the miner does not choose the highest bids. Let $(d_1, \ldots, d_B)$ be the resulting bids, where $d_1 \geq \cdots \geq d_B$ — we may assume that there are always infinitely many 0-bids that are "for free", and the miner can choose these 0-bids too. Then, miner's revenue becomes $\gamma(d_{k+1} + \cdots + d_B)$. Since $(c_1, \ldots, c_B)$ are the highest B bids, we have $c_j \geq d_j$ for all $j \in [B]$. Thus, miner's revenue does not increase.

We will next show that whenever the miner replaces an included real bid with a fake bid, the miner's utility does not increase. Notice that if the fake bid is confirmed, it costs the $(k+1)$-th price among the included bids. If the fake bid b is unconfirmed, it costs $\gamma \cdot b$, since its true value is zero. Let $\mathbf{e} = (e_1, \ldots, e_B)$ be an arbitrary bid vector where $e_1 \geq \cdots \geq e_B$. The bids $\mathbf{e}$ may or may not be the highest bids and some of them may be fake. Suppose the miner replaces e_i with the fake bid f. There are four possible cases.

1. $i \leq k$ and f is among the top k
2. $i \leq k$ and f is not among the top k
3. $i > k$ and f is among the top k
4. $i > k$ and f is not among the top k

Henceforth, no matter which case, let $e'_1 \geq \cdots \geq e'_B$ denote the included bids after replacing e_i with f. Let $\mu := \gamma(e_{k+1} + \cdots + e_B)$ be the miner's revenue before replacing e_i with the fake bid f, and let $\mu' := \gamma(e'_{k+1} + \cdots + e'_B)$ be the miner's revenue after replacing e_i with the fake bid f.

In the first case, for each bid among the top k, the probability that it is confirmed is γ/c, so the extra cost for the miner is

$$(1 - \frac{\gamma}{c}) \cdot \gamma \cdot f + \frac{\gamma}{c} \cdot e_{k+1} \geq 0$$

where the first term captures the expected cost if f is not confirmed, and the second term captures the expected cost if f is confirmed. In this case, it is easy to see that $e'_{k+j} = e_{k+j}$ for any $j > 0$, and thus $\mu' = \mu$. Therefore, miner's expected utility does not increase.

In the second case, f must be unconfirmed, so it costs the miner $\gamma \cdot f$ additionally to inject f. Also, since f is not among the top k, it must be $f \leq e_{k+1}$. Because e_{k+1} (or another bid equal to e_{k+1}) is among the new top k, the miner's revenue becomes $\gamma(e_{k+2} + \cdots e_B + f)$. Thus, the miner revenue decreases by $\gamma(e_{k+1} - f)$. Including the extra cost $\gamma \cdot f$, miner's utility actually decreases by

$$\gamma(f + e_{k+1} - f) \geq \gamma \cdot e_{k+1}.$$

In the third case, it must be $f \geq e_k$. Moreover, e_k becomes the largest definitely unconfirmed bid, so miner's extra cost is $(1 - \frac{\gamma}{c}) \cdot \gamma \cdot f + \frac{\gamma}{c} \cdot e_k \geq \gamma \cdot e_k$. However, it is not hard to see that $\mu' - \mu \leq \gamma \cdot e_k$. Therefore, overall, the miner's expected utility does not increase.

In the fourth case, f is unconfirmed, so it costs the miner $\gamma \cdot f$ additionally. If $f \leq e_i$, then it must be $e'_{k+j} \leq e_{k+j}$ for any $j > 0$. Therefore, we have $\mu' \leq \mu$, and the miner's revenue does not increase. Otherwise, if $f > e_i$, we have $(e'_{k+1} + \cdots + e'_B) - (e_{k+1} + \cdots + e_B) = f - e_i$. Thus, the increase in miner revenue is $\gamma(e'_{k+1} + \cdots + e'_B) - \gamma(e_{k+1} + \cdots + e_B) \leq \gamma(f - e_i) \leq \gamma \cdot f$, which is strictly smaller than the extra cost. Thus the miner's expected utility does not increase.

Finally, because $\mathbf{e}$ is an arbitrary vector which may include fake bids already, we conclude that the miner's expected utility does not increase even if there are multiple fake bids.

c-SCP. A coalition of a miner and up to c user(s) has three kinds of strategies to deviate from the honest behavior: the miner may not include the highest bids, the miner can inject fake bids, and some of the user(s) can bid untruthfully. Let C be the set of colluding users, where $|C| \leq c$. Without loss of generality, we assume the coalition prepares the block in the following order.

1. The miner chooses the included bids arbitrarily. We may imagine that there are infinitely many 0-bids that are "for free" and the miner can choose from these as well.

2. The miner replaces some of the included real bids (not including the users in C) with fake bids. Without loss of generality, we may assume that all injected fake bids are included in the block.

3. A subset of users in C change their bids and bid untruthfully.

We now show that the joint utility of the coalition does not increase after each step.

The first step of c-SCP. We may imagine that the miner deletes the real bids one by one, and then includes the highest among the remaining bids. We argue that after deleting each bid, the coalition's expected utility does not increase. Let $\mathbf{e} = (e_1, \ldots, e_m)$ be the current bid vector where $e_1 \geq \cdots \geq e_m$, which may already have some bids deleted from the real bid vector. Suppose the miner deletes a bid from $\mathbf{e}$. If the deleted bid is not among the top B, then it does not affect the coalition's utility. If the deleted bid is ranked between $[k + 2, B]$, then the miner's utility cannot increase and no user's utility increases.

If the deleted bid is among the top k, then the miner's revenue decreases by at least $\gamma \cdot (e_{k+1} - e_{k+2})$. Every user who was among the top k before and after this deletion has $\frac{\gamma}{c} \cdot (e_{k+1} - e_{k+2})$

increase in expected utility. The bid e_{k+1} (or another bid of equal value) now becomes among the top k, and its increase in expected utility is also $\frac{\gamma}{c} \cdot (e_{k+1} - e_{k+2})$. The utility of the user who got deleted decreases. All other users' utilities are unaffected. Thus, as long as the number of colluding users $|C| \leq c$, the increase in utility for users in C is upper bounded by $\gamma \cdot (e_{k+1} - e_{k+2})$. Overall, the coalition does not gain in expected utility.

If the deleted bid is ranked $k+1$ in $\mathbf{e}$, the miner's decrease in revenue is at least $\gamma \cdot (e_{k+1} - e_{k+2})$. For each user among the top k in $\mathbf{e}$, its increase in utility is $\frac{\gamma}{c} \cdot (e_{k+1} - e_{k+2})$. The utility of all other users are unaffected. Thus, as long as $|C| \leq c$, the increase in utility for users in C is upper bounded by $\gamma \cdot (e_{k+1} - e_{k+2})$. Overall, the coalition does not gain in expected utility.

The second step of c-SCP. Let $\mathbf{e}$ be an initial bid vector which may already have some bids deleted, and some real bids replaced with fake bids. Suppose the miner replaces some e_i where $i \in [B]$ with a fake bid f. Due to the proof of MIC, the miner's utility does not increase after the second step. If no user's expected utility increases after replacing a real bid with a fake one, then the coalition's expected utility cannot increase. Therefore, we only need to consider the cases in which there exists some user whose expected utility increases after replacement. Recall that in the proof of MIC, we divided into four possible cases. In cases 1 and 3, no user's expected utility would increase. Below, we focus on cases 2 and 4.

In case 2, $i \leq k$ and f is not among the top k. In this case, for every user $j \in [k]$ and $j \neq i$, its expected utility increases by $\frac{\gamma}{c}(e_{k+1} - \max(f, e_{k+2}))$. The bid e_{k+1} now becomes the top k, and its utility also increases by $\frac{\gamma}{c}(e_{k+1} - \max(f, e_{k+2}))$. The bid e_i's expected utility decreases, and all other users' expected utilities are unaffected. However, the miner's utility decreases by at least $\gamma \cdot (e_{k+1} - \max(f, e_{k+2}))$. Therefore, as long as $|C| \leq c$, the coalition's expected utility does not increase.

In case 4, $i > k$ and f is not among the top k. For some user's utility to increase, it must be that $i = k+1$ and $f < e_{k+1}$, i.e., the payment price must have decreased to $\max(f, e_{k+2})$. Similarly, for every user $j \in [k]$, its increase in expected utility is $\frac{\gamma}{c}(e_{k+1} - \max(f, e_{k+2}))$, and every other user's utility is unaffected. The miner's decrease in utility is at least $\gamma \cdot (e_{k+1} - \max(f, e_{k+2}))$. Therefore, as long as $|C| \leq c$, the coalition's expected utility does not increase.

The third step of c-SCP. At this step, the colluding users change their bids one by one. Without loss of generality, we assume the colluding users change their bids in an ascending order according to their true values; that is, the users with lower true values change their bids first. Let $\mathbf{e} = (e_1, \ldots, e_B)$ be an arbitrary bid vector included in the block, where $e_1 \geq \cdots \geq e_B$. The bids $\mathbf{e}$ may or may not be the highest bids and some of them may be fake or overbidding bids. Note that if any user whose bid is not included in the block changes its bid, the coalition's joint utility cannot increase. Thus, we may assume that a colluding user i included in the block changes its bid. Since the colluding users change their bids one by one, that means user i has not changed its bid before, and e_i must be the user's true value. Henceforth, we often use e_i to refer to the user that placed this bid without risking ambiguity. We will show that the joint utility of miner and all users in C would not increase if e_i changes its bid to b_i.

When e_i is replaced with b_i, there are four possible cases.

1. $i > k$ and b_i is among the top k
2. $i > k$ and b_i is not among the top k
3. $i \leq k$ and b_i is among the top k
4. $i \leq k$ and b_i is not among the top k

Henceforth, no matter which case, let $e'_1 \geq \cdots \geq e'_B$ denote the included bids after replacing e_i with b_i.

In the first case, e_k becomes the largest unconfirmed bid. e_i's utility was zero before, and it becomes $(1-\frac{\gamma}{c})\cdot\gamma\cdot(e_i-b_i)+\frac{\gamma}{c}\cdot(e_i-e_k)$ afterwards. Since $b_i \geq e_k \geq e_i$, the decrease in utility is at least $\gamma(e_k-e_i)$. Besides e_i, all other users' utilities cannot increase. The miner's revenue increases by at most $\gamma(e_k-e_{k+1}) \leq \gamma(e_k-e_i)$. Thus, the coalition's expected joint utility does not increase.

In the second case, we have $\sum_{i=1}^{k'} e'_{k+i} - \sum_{i=1}^{k'} e_{k+i} = b_i - e_i$. Thus the miner's change in revenue is $\gamma\cdot(b_i-e_i)$. For the users who are among the top k, their payment become e'_{k+1}. There are two subcases.

- If e_i is overbidding ($b_i > e_i$), it must be $e'_{k+1} \geq e_{k+1}$, so the utilities of the users among top k do not increase. e_i's utility reduces from zero to $\gamma(e_i-b_i)<0$. Since the miner's revenue increases at most by $\gamma(b_i-e_i)$, the coalition's joint expected utility does not increase.
- If e_i is underbidding ($b_i < e_i$), it must be $e_{k+1} \geq e'_{k+1}$. In this case, it must be $e_j \geq e'_j$ for all j. Further, e_i's utility is still zero; and the miner's revenue decreases at least $\gamma(e_i-b_i) \geq \gamma\cdot(e_{k+1}-e'_{k+1})$. The utility of each user among top k increases only by $\gamma(e_{k+1}-e'_{k+1})/c$. All other users' utilities are unaffected. Thus, even if the miner colludes with c users, their joint expected utility does not increase.

In the third case, e_i's utility and miner's utility do not change individually. Moreover, all other users' utilities do not change either, because $e_{k+1} = e'_{k+1}$. Thus, the joint utility of the coalition does not change.

In the fourth case, the miner's revenue reduces from $\gamma(e_{k+1}+\cdots+e_B)$ to $\gamma(e_{k+2}+\cdots+e_B+b_i) = \gamma(e_{k+1}-b_i)$. We now consider each user's change in utility.

- Since the user e_i's bid is replaced with b_i and now becomes unconfirmed, its utility reduces from $\gamma(e_i-e_{k+1})/c$ to zero. Also, note that e_i (who now bids b_i) must belong to C.
- For anyone that was among top k before and after the replacement, its new payment is $\max(b_i, e_{k+2})$ if confirmed. Thus its expected utility increases by $\frac{\gamma}{c}\cdot(e_{k+1}-\max(b_i,e_{k+2}))$.
- Now consider the user that bids e_{k+1}. This is the most complicated case. Let v be this user's true value. If this user is a coalition member, and its bid e_{k+1} was previously changed, then we know that $v \leq e_i$ since we are changing the coalition users' bids in ascending order of their true value. In all other cases, $v = e_{k+1} \leq e_i$.

  After the replacement of e_i with b_i, conditioned on not being confirmed, the user's utility does not change since previously it was always unconfirmed. Conditioned on being confirmed, the user's utility increases by at most $v - \max(b_i, e_{k+2}) + \max(0, \gamma\cdot(e_{k+1}-v))$, where the part $\max(0,\gamma\cdot(e_{k+1}-v))$ is because the user might be overbidding, i.e., $v < e_{k+1}$, and before the replacement it was always unconfirmed. Therefore, the user's expected gain in utility is $\frac{\gamma}{c}(v-\max(b_i,e_{k+2})+\max(0,\gamma\cdot(e_{k+1}-v)))$.
- For every other user, its utility is unaffected.

Now, suppose that the user bidding e_{k+1} belongs to the coalition. We know that e_i, whose bid is being changed to b_i, belongs to the coalition too. The joint utility of e_{k+1} and e_i increases by $\frac{\gamma}{c}(v-\max(b_i,e_{k+2})+\max(0,\gamma\cdot(e_{k+1}-v))) - \frac{\gamma}{c}(e_i-e_{k+1})$. If $e_{k+1} \geq v$, their increase in utility is

upper bounded by

$$\begin{aligned}
&\frac{\gamma}{c}\big[v - \max(b_i, e_{k+2}) + \gamma \cdot (e_{k+1} - v)\big] - \frac{\gamma}{c}(e_i - e_{k+1}) \\
=&\frac{\gamma}{c}\big[v - \max(b_i, e_{k+2}) + \gamma \cdot (e_{k+1} - v) - e_i + e_{k+1}\big] \\
=&\frac{\gamma}{c}\big[e_{k+1} - \max(b_i, e_{k+2}) + \gamma \cdot (e_{k+1} - v) - (e_i - v)\big] \\
\le&\frac{\gamma}{c}\big[e_{k+1} - \max(b_i, e_{k+2}) + \gamma \cdot (e_{k+1} - v) - (e_{k+1} - v)\big] \\
\le&\frac{\gamma}{c}(e_{k+1} - \max(b_i, e_{k+2}))
\end{aligned}$$

If $e_{k+1} < v$, their increase in utility is upper bounded by $\frac{\gamma}{c}(v - \max(b_i, e_{k+2}) - e_i + e_{k+1}) \le \frac{\gamma}{c}(e_{k+1} - \max(b_i, e_{k+2}))$, too.

All other users' expected utilities are either unaffected or increases by at most $\frac{\gamma}{c}(e_{k+1} - \max(b_i, e_{k+2}))$. Therefore, as long as $|C| \le c$, the coalition's joint utility cannot increase. Suppose that the user bidding e_{k+1} does not belong to the coalition. This case is easier since all users' expected utilities cannot increase by more than $\frac{\gamma}{c}(e_{k+1} - \max(b_i, e_{k+2}))$. Therefore, as long as $|C| \le c$, the coalition's joint utility cannot increase.

6 Randomness is Necessary for Weak Incentive Compatibility

Recall that when $c = 1$, our burning 2nd price auction becomes deterministic; but for all $c > 1$, the mechanism is randomized. In this section, we show that the randomness is in fact necessary for $c \ge 2$. To state this impossibility result, we first need to introduce a new notion that captures "non-degenerate" mechanisms, that is, we consider mechanisms that sometimes confirm 2 or more transactions:

Definition 6 (2-user-friendly)**.** We call a mechanism is *2-user-friendly* if there exists a bid vector $\mathbf{b}$ such that $x_i(\mathbf{b}) = x_j(\mathbf{b}) = 1$ for some $i \ne j$.

We prove the following impossibility result — throughout this section, we will assume that $\gamma = 1$ (also called weak incentive compatibility), since this makes our impossibility result stronger. The same impossibility result trivially extends to the case when $\gamma < 1$ as well.

Theorem 6.1. *Suppose the block size is finite. Let $(\mathbf{x}, \mathbf{p}, \mu)$ be a deterministic mechanism. If $(\mathbf{x}, \mathbf{p}, \mu)$ is 2-user-friendly, then it cannot achieve weak UIC and 2-weak-SCP at the same time.*

We stress that the 2-user-friendly restriction is in fact necessary for the above impossibility to hold. In particular, in Appendix B.1, we give a deterministic mechanism that always confirms only one transaction, and satisfies weak UIC, weak MIC, and c-weak-SCP for any c. Moreover, in Appendix B.2, we additionally show that the finite block size requirement is also necessary for the above impossibility to hold.

We presented a roadmap of the proof of Theorem 6.1 in Section 2.4. Therefore, we now directly jump to the detailed proof. To prove Theorem 6.1, we first prove that Myerson's lemma still holds for any deterministic, weakly UIC mechanism.

Fact 6.2. *Myerson's lemma holds for any deterministic, weakly UIC mechanism.*

Proof. Recall that in the definition of UIC or weak UIC, a user's strategy space involves not only bidding untruthfully, but also injecting fake transactions. To prove that Myerson's lemma

holds for weak UIC, we only care about bidding untruthfully, and we do not care about injecting fake transactions. Henceforth, if a mechanism disincentivizes an individual user from overbidding or underbidding under the *old* utility notion, we say that it is user-DSIC (short for dominant-strategy-incentive-compatible). Similarly, if a mechanism disincentivizes an individual user from overbidding or underbidding under the *new* utility notion, we say that it is weakly user-DSIC. Clearly, UIC implies user-DSIC and weak UIC implies weakly user-DSIC. Since Myerson's lemma holds for user-DSIC, it suffices to show that any *deterministic* TFM that is weakly user-DSIC must be user-DSIC, too.

Suppose for the sake of contradiction that there is a deterministic TFM that is weakly user-DSIC but not user-DSIC. This means that there is an untruthful bidding strategy that is profitable under the old utility notion (i.e., 0-strict utility) but not profitable any more under the new utility notion (i.e., 1-strict utility). In comparison with the old utility, the only difference in the new utility is that "overbidding but unconfirmed" is charged an additional cost. Therefore, such an untruthful bidding strategy as mentioned above must be overbidding but unconfirmed. However, we know that under the old utility notion, such an untruthful bidding strategy results in utility 0 and thus is not profitable. Thus the user does not want to adopt this strategy even under the old utility. This leads to a contradiction. □

Lemma 6.3. *Let $(\mathbf{x}, \mathbf{p}, \mu)$ be a deterministic mechanism which is weak UIC and 2-weak-SCP. Let $\mathbf{b} = (b_1, \ldots, b_m)$ be an arbitrary bid vector, where there exists a user i having a confirmed bid, i.e., $x_i(\mathbf{b}) = 1$. Then, for any bid vector $\mathbf{b}' = (\mathbf{b}_{-i}, b_i')$ such that $x_i(\mathbf{b}') = 1$, the followings holds.*

1. *Miner's revenue does not change; that is, $\mu(\mathbf{b}) = \mu(\mathbf{b}')$.*
2. *For any user j, if $x_j(\mathbf{b}) = 1$ and $b_j > p_j(\mathbf{b})$, it must be $x_j(\mathbf{b}') = 1$ and $p_j(\mathbf{b}) = p_j(\mathbf{b}')$.*

Proof. Because $x_i(\mathbf{b}) = x_i(\mathbf{b}') = 1$, we know that $p_i(\mathbf{b}) = p_i(\mathbf{b}')$ by Myerson's lemma. Recall that there is no cost for overbidding as long as the bid is confirmed. Therefore, user i's utility does not change no matter it bids b_i or b_i'. However, if $\mu(\mathbf{b}) \neq \mu(\mathbf{b}')$, the miner can sign a side contract to ask user i to bid the price that makes miner's revenue higher, thus violating 2-weak-SCP. For example, suppose $\mu(\mathbf{b}) < \mu(\mathbf{b}')$, then, in case the actual bid vector is $\mathbf{b}$ (where everyone's bidding its true value), the coalition of user i and the miner can gain by having user i bid b_i' instead of its true value b_i. Similarly, if $\mu(\mathbf{b}) > \mu(\mathbf{b}')$, a symmetric argument holds. Thus, it must be $\mu(\mathbf{b}) = \mu(\mathbf{b}')$.

We next prove that $x_j(\mathbf{b}') = 1$ for any user j with $x_j(\mathbf{b}) = 1$ and $b_j > p_j(\mathbf{b})$. For the sake of reaching a contradiction, suppose that $x_j(\mathbf{b}') = 0$. We now show that the coalition of the miner, user i, and user j can gain if everyone's true value is $\mathbf{b}'$. Suppose user i were to bid its true value b_i', user j's utility would be 0 since $x_j(\mathbf{b}') = 0$. Therefore, the coalition is better off having user i bid b_i instead. In this case, user j's utility would be $b_j - p_j(\mathbf{b}) > 0$. Furthermore, as we have shown, $\mu(\mathbf{b}') = \mu(\mathbf{b})$, and moreover, by Myerson's Lemma, user i's payment and utility do not change as long as it bids high enough to be confirmed. Thus the coalition gains positively by having user i bid b_i instead of its true value b_i'. This violates 2-weak-SCP.

Finally, we prove that $p_j(\mathbf{b}) = p_j(\mathbf{b}')$ for any user j with $x_j(\mathbf{b}) = 1$ and $b_j > p_j(\mathbf{b})$. Because we have shown $x_j(\mathbf{b}) = x_j(\mathbf{b}') = 1$, user j's utility is $v_j - p_j(\mathbf{b})$ if user i bids b_i, and $v_j - p_j(\mathbf{b}')$ if user i bids b_i'. Suppose for the sake of contradiction that $p_j(\mathbf{b}) \neq p_j(\mathbf{b}')$. There are two cases. First, suppose that $p_j(\mathbf{b}) > p_j(\mathbf{b}')$. Imagine now that everyone's true value is $\mathbf{b}'$. In this case, the miner can collude with both user i and user j, and have user i bid b_i rather than its true value to increase the coalition's joint utility. This violates 2-weak-SCP. Similarly, we can rule out the casw where $p_j(\mathbf{b}) < p_j(\mathbf{b}')$ due to a symmetric argument. □

Lemma 6.4. *Let $(\mathbf{x}, \mathbf{p}, \mu)$ be a deterministic mechanism that achieves weak UIC and 1-weak-SCP. Then, for all users i, j, if $x_i(\mathbf{b}) = x_j(\mathbf{b}) = 1$, it must be $p_i(\mathbf{b}) = p_j(\mathbf{b})$. In other words, all confirmed users must pay the same price.*

Proof. Suppose i and j are two confirmed users; that is $x_i(\mathbf{b}) = x_j(\mathbf{b}) = 1$. For the sake of reaching a contradiction, we assume that $p_i(\mathbf{b}) \neq p_j(\mathbf{b})$. There are two possible cases:

1. At least one user's bid is higher than its payment; that is, either $b_i > p_i(\mathbf{b})$ or $b_j > p_j(\mathbf{b})$ (or both).

2. Both users pay their bids; that is, $b_i = p_i(\mathbf{b})$ and $b_j = p_j(\mathbf{b})$.

We start from the first case. Without loss of generality, assume $b_i > p_i(\mathbf{b})$. According to Lemma 6.3, user j can increase its bid without changing user i's confirmation and payment. Furthermore, by Myerson's lemma, user j's payment should not change. Thus, we have another bid vector $\mathbf{b}' = (\mathbf{b}_{-j}, b'_j)$ such that $b_i > p_i(\mathbf{b}')$ and $b'_j > p_j(\mathbf{b}')$. Using Lemma 6.3 again, we can increase user i's and user j's bid arbitrarily while remaining their confirmation and payment. Consequently, we have a bid vector $\mathbf{c} = (c_1, \ldots, c_m)$ such that user i and user j have the same bid. Formally, $x_i(\mathbf{c}) = x_j(\mathbf{c}) = 1$, $c_i > p_i(\mathbf{c}) = p_i(\mathbf{b})$, $c_j > p_j(\mathbf{c}) = p_j(\mathbf{b})$ and $c_i = c_j$. Without loss of generality, we assume $p_i(\mathbf{c}) > p_j(\mathbf{c})$. Imagine that the real bid vector is $\mathbf{c}$. In this case, miner's utility is $\mu(\mathbf{c})$, and user i's utility is $c_i - p_i(\mathbf{c})$. The miner can sign a contract with user i, and switch user i's and user j's positions in the bid vector. Since users i and j are bidding the same, the miner's revenue is unaffected if their positions are switched. On the other hand, user i and user j's payments will be switched as a result. Thus, user i's utility has increased to $c_i - p_j(\mathbf{b})$. This violates 1-weak-SCP.

Next, we analyze the second case. Without loss of generality, we assume $b_i = p_i(\mathbf{b}) > b_j = p_j(\mathbf{b})$. By Myerson's lemma, user j can increase its bid without changing its payment. Thus, user j can increase its bid to b_i, and we have a bid vector $\mathbf{b}' = (\mathbf{b}_{-j}, b_i)$. By Lemma 6.3, miner's revenue should not change, so we have $\mu(\mathbf{b}) = \mu(\mathbf{b}')$. If $x_i(\mathbf{b}') = 1$, it goes back to the first case, so we assume $x_i(\mathbf{b}') = 0$. Now, imagine that the real bid vector is $\mathbf{b}'$. In this case, miner's utility is $\mu(\mathbf{b}') = \mu(\mathbf{b})$, and user i's utility is zero. However, the miner can sign a contract with user i, and ask it to bid b_j instead. Consequently, the miner prepares a bid vector $\mathbf{b}$, where b_j comes from user i. In this case, miner's utility is still $\mu(\mathbf{b}') = \mu(\mathbf{b})$, while user i's utility becomes $b_i - p_j(\mathbf{b}) > 0$. This violates 1-weak-SCP. □

Lemma 6.5. *Let $(\mathbf{x}, \mathbf{p}, \mu)$ be a deterministic mechanism that achieves weak UIC and 2-weak-SCP. Then, for all users i, j, if $x_i(\mathbf{b}) = 1$ and $x_j(\mathbf{b}) = 0$, it must be $b_i \geq b_j$. In other words, the confirmed bids must be the highest k bids for some $k \in \mathbb{N}$ where k may be a function of the bid vector.*

Proof. For the sake of reaching a contradiction, suppose that there exist two users i, j such that $x_i(\mathbf{b}) = 1$ and $x_j(\mathbf{b}) = 0$, while $b_i < b_j$. By Myerson's lemma, user i can increase its bid to b_j without changing its confirmation and payment. Thus, we have a bid vector $\mathbf{b}' = (\mathbf{b}_{-i}, b_i = b_j)$ such that $x_i(\mathbf{b}') = 1$ and $p_i(\mathbf{b}') = p_i(\mathbf{b})$. There are two possible cases: either $x_j(\mathbf{b}') = 1$ or $x_j(\mathbf{b}') = 0$.

First, we assume $x_j(\mathbf{b}') = 1$. Imagine the real bid vector is $\mathbf{b}$ which also represents everyone's true value. In this case, miner's utility is $\mu(\mathbf{b})$, user i's utility is $b_i - p_i(\mathbf{b})$, and user j's utility is zero. However, the miner can sign a contract with user i and user j, and ask user i to bid b_j instead. By Lemma 6.4, we have $p_i(\mathbf{b}') = p_j(\mathbf{b}')$. Because $p_i(\mathbf{b}') = p_i(\mathbf{b})$ and $p_i(\mathbf{b}) \leq b_i < b_j$, we have $p_j(\mathbf{b}') < b_j$. Besides, by Lemma 6.3, $\mu(\mathbf{b}) = \mu(\mathbf{b}')$. Therefore, after signing the contract, miner's utility is still $\mu(\mathbf{b})$, user i's utility is still $b_i - p_i(\mathbf{b})$, while user j's utility becomes $b_j - p_j(\mathbf{b}') > 0$. This violates 2-weak-SCP.

Next, we assume $x_j(\mathbf{b}') = 0$. Imagine the real bid vector is $\mathbf{b}'$ which also represents everyone's true value. Recall that in $\mathbf{b}'$, user i and user j are bidding the same; however, user i is confirmed but user j is not. Furthermore, user i is bidding strictly higher than its payment as we have shown above. The coalition of the miner and user j can strictly benefit, if the miner switched user i and user j's positions in the bid vector; since this does not affect the miner's utility, but user j's utility now becomes positive. This violates 1-weak-SCP.

□

Notation for the universal payment. According to Lemma 6.4, all confirmed users must pay the same price. Thus, we may simplify the notation, and define $p(\mathbf{b})$ to be the universal payment price for all confirmed users under the bid vector $\mathbf{b}$. If no one is confirmed under $\mathbf{b}$, then we define $p(\mathbf{b}) = 0$.

Lemma 6.6. *Let $(\mathbf{x}, \mathbf{p}, \mu)$ be a deterministic mechanism that achieves weak UIC and 2-weak-SCP. Let $\mathbf{b}$ be a bid vector such that at least one user is confirmed. Then, for any unconfirmed user i, it must be $b_i \le p(\mathbf{b})$.*

Proof. For the sake of reaching a contradiction, suppose there exists an unconfirmed user i such that $b_i - p(\mathbf{b}) = \Delta$ for some $\Delta > 0$. Let j be a confirmed user in $\mathbf{b}$. By Lemma 6.5, we know that $b_j \ge b_i$. Now, consider another bid vector $\mathbf{b}' = (\mathbf{b}_{-j}, p(\mathbf{b}) + \Delta/2)$. By Myerson's Lemma, user j is still confirmed and is still paying $p(\mathbf{b})$. However, notice that user j's new bid $p(\mathbf{b}) + \Delta/2 < b_i$. By Lemma 6.5, user i must be confirmed too under the bid vector $\mathbf{b}'$, and by Lemma 6.4, it would be paying the same as user j, which is $p(\mathbf{b})$. Now, imagine that everyone's true value is the vector $\mathbf{b}$. If everyone bids honestly, the miner's utility is $\mu(\mathbf{b})$, user j's utility is $b_j - p(\mathbf{b})$, and user i's utility is zero. If the miner colludes with users i, j, the coalition can benefit by having user j bid $p(\mathbf{b}) + \Delta/2$ instead. In this case, miner's utility is still $\mu(\mathbf{b})$ due to Lemma 6.3, user j's utility is still $b_j - p(\mathbf{b})$, while user i's utility increases to $\Delta/2$. This violates 2-weak-SCP. □

Lemma 6.7. *Let $(\mathbf{x}, \mathbf{p}, \mu)$ be a deterministic, weak UIC, and 1-weak-SCP mechanism. Then, for any bid vector $\mathbf{b}$, any user k, and any $0 < \Delta \le b_k$, it must be that $\mu(\mathbf{b}) - \Delta \le \mu(\mathbf{b}_{-k}, b_k - \Delta) \le \mu(\mathbf{b})$.*

Proof. We first prove the direction $\mu(\mathbf{b}_{-k}, b_k - \Delta) \le \mu(\mathbf{b})$. We want to show that if the users' true value is $\mathbf{b}$, but now user k bids $b_k - \Delta$ instead of its true value b_k, then the miner revenue should not increase. There are two cases.

- First, if user k is unconfirmed under $\mathbf{b}$ or confirmed but paying its full bid b_k, then its utility is 0 under $\mathbf{b}$. In this case, obviously decreasing user k's bid should not make the miner benefit; since otherwise the coalition of user k and the miner can benefit by having user k bid $b_k - \Delta$ instead, thus violating 1-weak-SCP.
- Second, suppose that user k is initially confirmed under $\mathbf{b}$ and moreover, $b_k > p(\mathbf{b})$. We can first decrease user k's bid to exactly $\max(b_k - \Delta, p(\mathbf{b}))$, and let $\mathbf{b}' := (\mathbf{b}_{-k}, \max(b_k - \Delta, p(\mathbf{b})))$ be the resulting new bid vector. Due to Myerson's Lemma, Lemma 6.3, and Lemma 6.4, $x_k(\mathbf{b}') = 1$, $\mu(\mathbf{b}') = \mu(\mathbf{b})$, and $p(\mathbf{b}') = p(\mathbf{b})$. Then, we can decrease user k's bid from $\max(b_k - \Delta, p(\mathbf{b}))$ to $b_k - \Delta$, and due to the same argument as the first case, the miner's revenue should not increase.

We next prove the other direction, that is, $\mu(\mathbf{b}) - \Delta \le \mu(\mathbf{b}_{-k}, b_k - \Delta)$. If no one is confirmed under $\mathbf{b}$, the statement trivially holds. Henceforth, we assume that at least one user is confirmed under $\mathbf{b}$. Again, there are two cases.

- First, suppose that user k is not confirmed under $\mathbf{b}$ or confirmed but paying its full bid b_k. Due to Lemma 6.6, we know that $b_k \leq p(\mathbf{b})$. By Myerson's Lemma, user k should be unconfirmed or confirmed but paying full bid under $\mathbf{b}' := (\mathbf{b}_{-k}, b_k - \Delta)$. Now, suppose everyone's true value is actually $\mathbf{b}'$, notice that user k's utility is 0. We argue that if user k bids b_k instead of its true value $b_k - \Delta$, it should not make the miner revenue increase by more than Δ. If so, the coalition of user k and the miner can strictly benefit by having user k bid b_k instead of its true value $b_k - \Delta$, since the cost of such overbidding is at most Δ under the new utility notion. This violates 1-weak-SCP.
- Second, suppose that user k is confirmed under $\mathbf{b}$ and moreover $b_k > p(\mathbf{b})$. In this case, due to Myerson's Lemma, Lemma 6.3, and Lemma 6.4, the miner's revenue should not be affected when we reduce user k's bid to $\max(p(\mathbf{b}), b_k - \Delta)$. We now further decrease user k's bid from $\max(p(\mathbf{b}), b_k - \Delta)$ to $b_k - \Delta$ — due to the same analysis as the first case, the miner's revenue should not increase by more than Δ in this process.

□

Lemma 6.8. *Let $(\mathbf{x}, \mathbf{p}, \mu)$ be a deterministic mechanism which is weak UIC and 2-weak-SCP. Let $\mathbf{b} = (b_1, \ldots, b_m)$ be an arbitrary bid vector. Suppose that there exist three different users i, j, k such that $x_i(\mathbf{b}_{-k}, 0) = x_j(\mathbf{b}_{-k}, 0) = 1$, and moreover, $b_i - p(\mathbf{b}_{-k}, 0) > b_k$ and $b_j - p(\mathbf{b}_{-k}, 0) > b_k$. Then, it must be that $x_i(\mathbf{b}) = x_j(\mathbf{b}) = 1$ and $p(\mathbf{b}) \leq p(\mathbf{b}_{-k}, 0) + b_k/2$.*

Proof. We first prove that $x_i(\mathbf{b}) = x_j(\mathbf{b}) = 1$. Suppose not, without loss of generality, let us suppose $x_i(\mathbf{b}) = 0$ since the case $x_j(\mathbf{b}) = 0$ has a symmetric proof. Imagine that the real bid vector is $\mathbf{b}$ which also represents everyone's true value. Suppose the miner and user i form a coalition, and they replace user k's bid with an injected 0-bid. Due to Lemma 6.7, the miner's utility decreases by at most b_k as a result. However, user i now becomes confirmed and its utility is $b_i - p(\mathbf{b}_{-k}, 0) > b_k$. Therefore, the coalition strictly gains which violates 1-weak-SCP.

We next prove that $p(\mathbf{b}) \leq p(\mathbf{b}_{-k}, 0) + b_k/2$. For the sake of reaching a contradiction, suppose $p(\mathbf{b}) > p(\mathbf{b}') + b_i/2$. Imagine the real bid vector is $\mathbf{b}$ which is also everyone's true value. In this case, miner's utility is $\mu(\mathbf{b})$, user i's utility is $b_i - p(\mathbf{b})$, and user j's utility is $b_j - p(\mathbf{b})$. However, the miner can collude with user i and user j, and miner replaces b_k with an injected 0. Notice that injecting a 0-bid costs nothing. In this case, miner's utility becomes $\mu(\mathbf{b}') \geq \mu(\mathbf{b}) - b_k$, where the inequality follows from Lemma 6.7. On the other hand, user i's utility becomes $b_i - p(\mathbf{b}')$, and user j's utility becomes $b_j - p(\mathbf{b}')$. Because $p(\mathbf{b}) > p(\mathbf{b}') + b_k/2$, user i's and user j's utilities each increases more than $b_k/2$. Consequently, the coalition's joint utility increases, which violates 2-weak-SCP. □

Lemma 6.9. *Let $(\mathbf{x}, \mathbf{p}, \mu)$ be a deterministic mechanism which is weak UIC and 1-weak-SCP. Let $\mathbf{b} = (b_1, \ldots, b_m)$ be an arbitrary bid vector. Let i and k be two different users, and suppose that $x_i(\mathbf{b}) = 1$ and $b_i - p(\mathbf{b}) > b_k$. Then, $x_i(\mathbf{b}_{-k}, 0) = 1$ and $p(\mathbf{b}_{-k}, 0) \leq p(\mathbf{b}) + b_k$.*

Proof. For the sake of contradiction, suppose either $x_i(\mathbf{b}_{-k}, 0) = 0$ or $p(\mathbf{b}_{-k}, 0) > p(\mathbf{b}) + b_k$. Imagine that the real bid vector is $\mathbf{b}' := (\mathbf{b}_{-k}, 0)$, which also represents everyone's true value. Now, the miner replaces user k's 0-bid in $\mathbf{b}'$ with an injected bid b_k. Injecting this bid costs at most b_k. Due to Lemma 6.7, the miner's utility cannot decrease, i.e., $\mu(\mathbf{b}) \geq \mu(\mathbf{b}')$. However, consider user i's utility. If $x_i(\mathbf{b}') = 0$ but $x_i(\mathbf{b}) = 1$, user i's utility has increased from 0 to $b_i - p(\mathbf{b}) > b_k$. Else, if $x_i(\mathbf{b}') = x_i(\mathbf{b}) = 1$, but $p(\mathbf{b}') > p(\mathbf{b}) + b_k$, then user i's utility increases by strictly more than b_k too. In either case, the coalition of the miner and user i can strictly increase their joint utility by replacing user k's bid with the injected b_k bid, which violates 1-weak-SCP. □

Lemma 6.10. *Let $(\mathbf{x}, \mathbf{p}, \mu)$ be a deterministic mechanism which is 2-user-friendly, weak-UIC and 2-weak-SCP. Then, there exists a bid vector $\mathbf{b} = (b_1, \ldots, b_m)$ where two different users i, j are confirmed, and moreover, $b_i > p_i(\mathbf{b})$ and $b_j > p_j(\mathbf{b})$.*

Proof. Since $(\mathbf{x}, \mathbf{p}, \mu)$ is 2-user-friendly, there exists a bid vector $\mathbf{b}$ such that at least two users' bids are confirmed. There are three possible cases:

1. There are two users i, j such that $b_i > p_i(\mathbf{b})$ and $b_j > p_j(\mathbf{b})$.
2. Only a single confirmed user bids strictly above its payment. Without loss of generality, we may assume user j is confirmed and $b_j > p_j(\mathbf{b})$; however, for any confirmed user $i \neq j$, $b_i = p_i(\mathbf{b})$, and there exists at least one such i.
3. For any confirmed bid b_i, it holds that $b_i = p_i(\mathbf{b})$.

The first case is exactly what we want. For the second case, we can raise i's bid by an arbitrary amount $\Delta > 0$, and the new bid vector is $(\mathbf{b}_{-i}, b_i + \Delta)$. By Myerson's Lemma, i should still be confirmed and paying the same price. Due to Lemma 6.3, j should still be confirmed and paying the same price, too. Therefore, the bid vector $(\mathbf{b}_{-i}, b_i + \Delta)$ satisfies the claim we want to prove. Moreover, due to Lemma 6.3, it must be $\mu(\mathbf{b}) = \mu(\mathbf{b}_{-i}, b_i + \Delta)$.[3]

We now focus on the third case which is the trickiest. Due to Lemma 6.4, it must be that everyone confirmed has the same bid, and thus $b_i = b_j$. Fix an arbitrary $\Delta > 0$. Consider the bid vector $\mathbf{b}^*$ which is the same as $\mathbf{b}$ except that user i and user j's bids are replaced with $b_i + \Delta$. We claim that the bid vector $\mathbf{b}^*$ satisfies the claim we want to prove. In other words, $x_i(\mathbf{b}^*) = x_j(\mathbf{b}^*) = 1$ and $b_i + \Delta > p(\mathbf{b}^*)$. Suppose this is not the case. There are three cases:

1. both i and j are not confirmed under $\mathbf{b}^*$;
2. exactly one of them is not confirmed under $\mathbf{b}^*$ — without loss of generality, we may assume that j is not confirmed under $\mathbf{b}^*$. In this case, by Lemma 6.6, it must be that $p_i(\mathbf{b}^*) = b_i + \Delta$;
3. both i and j are confirmed under $\mathbf{b}^*$ but $b_i + \Delta = p(\mathbf{b}^*)$.

In all of these cases, user i and j both have utility 0 if the true values are $\mathbf{b}^*$.

Let $\mathbf{b}' := (\mathbf{b}_{-i}, b_i + \Delta)$. By Lemma 6.7, $\mu(\mathbf{b}^*) \geq \mu(\mathbf{b}') \geq \mu(\mathbf{b}^*) - \Delta$. By Lemma 6.3, $\mu(\mathbf{b}') = \mu(\mathbf{b})$. Thus, $\mu(\mathbf{b}^*) \geq \mu(\mathbf{b}) \geq \mu(\mathbf{b}^*) - \Delta$.[4] Now, suppose the true values are $\mathbf{b}^*$. The miner can collude with users i and j, and have them bid $b_i = b_j$ instead. Both users i and j are paying b_i in this case. Therefore, each of them has utilty Δ now. On the other hand, the miner's utility decreases by at most Δ, and therefore the coalition's utility increases. This violates 2-weak-SCP. □

Proof of Theorem 6.1. By Lemmas 6.4 and 6.10, there exists a bid vector $\mathbf{b} = (b_1, \ldots, b_m)$ such that $x_1(\mathbf{b}) = x_2(\mathbf{b}) = 1$, $b_1 > p(\mathbf{b})$ and $b_2 > p(\mathbf{b})$ — we can always relabel the bids to make any two confirmed users with positive utility labeled as users 1 and 2.

Let B denote the block size, and we define $\Gamma = 2^{B+8} \cdot |\mathbf{b}|_1 \cdot \max(m, B+1)$ to be a sufficiently large number. Now, we consider a bid vector $\mathbf{c} = (c_1, c_2, \ldots, c_m)$, where $c_1 = c_2 = \Gamma$ and $c_i = 0$ for all $i \geq 3$. We are going to show that $x_1(\mathbf{c}) = x_2(\mathbf{c}) = 1$. By the Myerson's Lemma and Lemma 6.3, from $\mathbf{b}$, we can increase b_1 without changing the confirmation and the payment of b_1 and b_2, so $x_1(\Gamma, b_2, \ldots, b_m) = 1$ and $x_2(\Gamma, b_2, \ldots, b_m) = 1$. Similarly, we can then increase b_2 and we obtain

[3] $\mu(\mathbf{b}) = \mu(\mathbf{b}_{-i}, b_i + \Delta)$ is not important for this proof, while this fact will be useful in the proof of Lemma B.5.
[4] $\mu(\mathbf{b}^*) \geq \mu(\mathbf{b})$ is not important for this proof, while this fact will be useful in the proof of Lemma B.5.

$x_1(\Gamma, \Gamma, b_3, \ldots, b_m) = 1$, $x_2(\Gamma, \Gamma, b_3, \ldots, b_m) = 1$ and $p(\Gamma, \Gamma, b_3, \ldots, b_m) = p(\mathbf{b})$. Next, we reduce all remaining bids to zero one by one. Repeatedly applying Lemma 6.9 and observing that Γ is sufficiently large, we have that $x_1(\mathbf{c}) = x_2(\mathbf{c}) = 1$, and the payment increases by $\sum_{i=3}^{m} b_i$ at most, so $p(\mathbf{c}) \leq p(\mathbf{b}) + \sum_{i=3}^{m} b_i < |\mathbf{b}|_1$.

Without loss of generality, we may henceforth assume that $m > B$. If not, that is, if $m < B$, we can always add 0 bids one by one until there are at least $B + 1$ bids — we claim that this does not change the miner's utility nor user 1 or 2's confirmation status and payment. To see this, notice that adding or removing a 0-bid is free of charge for the miner or a miner-user coalition. Therefore, adding or removing a 0-bid should not change the joint utility of the miner and user 1 by 1-weak-SCP. This implies that user 2's confirmation status and payment should not change, since otherwise, user 2's utility would change, and thus the joint utility of the miner and users 1 and 2 would change. This means that the coalition of the miner and users 1 and 2 can cheat by adding or removing a 0-bid to increase their joint utility, thus violating 2-weak-SCP. By a symmetric argument, user 1's confirmation status or payment should not change either. Now, since the joint utility of the miner and user 1 should not change due to adding or removing a 0-bid, the miner's utility should be unaffected too.

Now, we consider another bid vector $\mathbf{d} = (d_1, d_2, \ldots, d_m)$, where $d_i = \Gamma$ for all $i \in [B + 1]$ and $d_i = 0$ for $i > B + 1$. We are going to show that $x_i(\mathbf{d}) = 1$ for all $i \in [B + 1]$ — note that this is sufficient for reaching a contradiction since the block size is only B. To see this, we start from $\mathbf{c}$, and increase the bids of each user $j \in \{3, \ldots, m\}$ one by one. Intuitively, Lemma 6.8 guarantees that if the payment grows at all during the process, it must grow slower than the increase in a user's bid, so at some point, user j's bid will catch up with the payment, as long as there is still a large enough gap left between Γ and the payment. Formally, since $\Gamma - p(\mathbf{c}) > 2|\mathbf{b}|_1$, by Lemma 6.8, it must be that $p(\Gamma, \Gamma, 2|\mathbf{b}|_1, 0, \ldots, 0) \leq p(\mathbf{c}) + |\mathbf{b}|_1 < 2|\mathbf{b}|_1$, and $x_3(\Gamma, \Gamma, 2|\mathbf{b}|_1, 0, \ldots, 0) = 1$. We now further increase user 3's bid to Γ, and by Lemma 6.3, users 1 to 3 remain confirmed and $p(\Gamma, \Gamma, \Gamma, 0, \ldots, 0) < 2|\mathbf{b}|_1$. By the same reasoning, since $\Gamma - p(\Gamma, \Gamma, \Gamma, 0, \ldots, 0) > 4|\mathbf{b}|_1$, by Lemma 6.8, it must be that $p(\Gamma, \Gamma, \Gamma, 4|\mathbf{b}|_1, \ldots, 0) \leq p(\Gamma, \Gamma, \Gamma, 0, \ldots, 0) + 2|\mathbf{b}|_1 < 4|\mathbf{b}|_1$, and $x_4(\Gamma, \Gamma, \Gamma, 4|\mathbf{b}|_1, 0, \ldots, 0) = 1$. Therefore, $p(\Gamma, \Gamma, \Gamma, \Gamma, 0, \ldots, 0) < 4|\mathbf{b}|_1$. We can now repeat this process and raise the bid of each user $i \in [B + 1]$ to Γ. It is not hard to check that our choice of Γ is sufficiently large for the reasoning to go through in all steps.

7 Necessity for Blocks to Contain Unconfirmed Transactions

Observe that in our buring second-price auction, not all transactions in the block are confirmed. In particular, only the top k have a chance of being confirmed, and remaining $B - k$ included bids are not confirmed. Instead, they serve the role of setting the price, i.e., they are used by the blockchain to compute the payment for each confirmed bid and the miner revenue. In the cryptocurrency community, there is an ongoing debate whether including unconfirmed transactions in a block is a good idea. The argument against this approach is that "real estate" on the blockchain is a scarce resource, so we ideally do not want to waste space including unconfirmed transactions in the block.

We argue that having unconfirmed transactions in the block indeed can lead to more versatile mechanisms. To show this, we argue that if "included" must be equal to "confirmed", then, even weakly incentive mechanisms are not possible. More specifically, we prove the following corollary:

Corollary 7.1 (Impossibility for "included = confirmed")**.** *Assume that all transactions included in the block must be confirmed. Then, no (possibly randomized) TFM* $(\mathbf{x}, \mathbf{p}, \mu)$ *with non-trivial miner revenue can satisfy weak UIC, weak MIC, and 1-weak-SCP at the same time — this impossibility holds no matter whether the block size is finite or infinite.*

Moreover, if the block size is finite, then the only (possibly randomized) TFM that achieves weak UIC, weak MIC, and 1-weak-SCP is the trivial mechanism that always confirms nothing and pays the miner nothing.

Myerson's lemma still holds. To prove this corollary, an important stepping stone is to prove that Myerson's lemma still holds for any weak UIC, weak MIC, and 1-weak-SCP (randomized) mechanism where "included = confirmed". It turns out that this is somewhat non-trivial to prove.

Recall that in a randomized mechanism, the random coins come from two sources: 1) the miner can flip random coins to decide which transactions to include in the block; 2) once the inclusion choices are made, the blockchain flips random coins to determine which of the included transactions are confirmed, how much each confirmed transaction pays, and how much the miner gets. In other words, the randomness in the inclusion rule is chosen by the miner, whereas the randomness in the confirmation, payment, and miner-revenue rules are chosen by the blockchain. A strategic miner may choose its random coins arbitrarily and not uniformly at random, to increase its expected gain. On the other hand, we assume that the blockchain's randomness is trusted. In other words, we assume that the blockchain can toss fresh random coins that are revealed *after* the miner commits to its inclusion decision — this makes our impossibility result stronger, since if the blockchain's randomness is revealed to the miner earlier, it makes mechanism design even harder.

Terminology and notation. Fix an arbitrary bid vector $\mathbf{b} = (b_1, \ldots, b_m)$. Let $S \subseteq \{b_1, \ldots, b_m\}$ denote a subset of these bids to include in the block. We often call S an *inclusion outcome*. Note that if "included = confirmed", the miner is essentially choosing which transactions are confirmed directly, too. Whenever the miner picks an inclusion outcome $S \subseteq \{b_1, \ldots, b_m\}$, it can calculate its expected utility denoted $\mathbb{E}(\mu|S)$ where the expectation is taken over the choice of the blockchain's random coins. We use the notation $\mathbb{E}(\mu)$ to denote the miner's expected utility under $\mathbf{b}$, had it executed the TFM honestly.

Fix an arbitrary bid vector $\mathbf{b} = (b_1, \ldots, b_m)$. We say that an inclusion outcome $S \subseteq \{b_1, \ldots, b_m\}$ is *possible* (w.r.t. $\mathbf{b}$), if it is encountered with non-zero probability in an honest execution of the TFM over $\mathbf{b}$.

Lemma 7.2. *Suppose that a randomized TFM satisfies weak MIC, and moreover, any transaction included in the blockchain must be confirmed. Fix an arbitrary bid vector $\mathbf{b} = (b_1, \ldots, b_m)$. For any possible inclusion outcome $S \subseteq \{b_1, \ldots, b_m\}$ it must be that $\mathbb{E}(\mu|S) = \mathbb{E}(\mu)$.*

As a direct corollary, for any possible inclusion outcomes $S, S' \subseteq \{b_1, \ldots, b_m\}$ it must be that $\mathbb{E}(\mu|S) = \mathbb{E}(\mu|S')$.

Proof. Suppose that there is a possible inclusion outcome S where $\mathbb{E}(\mu|S) \neq \mathbb{E}(\mu)$. It must be that there is a possible inclusion outcome S^* where $\mathbb{E}(\mu|S^*) > \mathbb{E}(\mu)$. In this case, instead of choosing the miner coins at random as prescribed by the mechanism, it strictly benefits the miner to choose the specific inclusion outcome S^*. This violates weak MIC. □

Lemma 7.3. *Suppose that a randomized TFM satisfies weak MIC and 1-weak-SCP, and moreover, any transaction included in the blockchain must be confirmed. Suppose that under some bid vector $\mathbf{b} = (b_1, \ldots, b_m)$, there is at least one possible inclusion outcome that includes b_i, and at least one possible inclusion outcome that does not include b_i. Then, consider any possible inclusion outcome S that includes b_i, it must be that conditioned on S, user i pays its full bid b_i with probability 1.*

Proof. Suppose that the claim does not hold, i.e., there is a possible inclusion outcome that includes b_i, but user i pays $p_i < b_i$; and moreover, there is at least one possible inclusion outcome that does

not include b_i. Let S^* be a possible inclusion outcome that includes b_i that minimizes the payment of user i. In this case, the miner can form a coalition with user i, and the miner can choose the inclusion outcome S^* with probability 1. Due to weak MIC and Lemma 7.2, the miner's utility is still $\mathbb{E}(\mu)$ when it adopts this strategy, i.e., the same as playing honestly. However, user i's utility is positive and is maximized under S^*. Furthermore, since there is at least one possible inclusion outcome that does not include b_i where user i's utility is 0, it must be that user i's expected utility is strictly greater under this strategy than playing honestly. Therefore, the coalition strictly benefits under this strategy, which violates 1-weak-SCP. □

Lemma 7.4. *Suppose that a randomized TFM satisfies weak UIC, weak MIC, and 1-weak-SCP, and moreover, any included transaction must be confirmed. Then, the TFM must satisfy the constraints imposed by the Myerson's Lemma.*

Proof. Recall that a mechanism disincentivizes an individual user from overbidding or underbidding under the old utility notion, we say that it is user-DSIC (short for dominant- strategy-incentive-compatible). Similarly, if a mechanism disincentivizes an individual user from overbidding or underbidding under the new utility notion, we say that it is weakly user-DSIC. Clearly, UIC implies user-DSIC and weak UIC implies weakly user-DSIC, since in our definitions of (weak) UIC, the user can misbehave in more ways besides over- or under-bidding. Since Myerson's lemma holds for user-DSIC, it suffices to show that any (randomized) TFM where "included = confirmed" and satisfying weak user-DSIC, weak MIC, and 1-weak-SCP must also satisfy user-DSIC.

Suppose that this is not true, i.e., there is some TFM where "included = confirmed" and satisfying weak user-DSIC, weak MIC, and 1-weak-SCP, however, the TFM does not satisfy user-DSIC. Notice that if a user underbids, its utility is the same under the old and new utility notions. Therefore, there must exist a bid vector $\mathbf{b} = (b_1, \ldots, b_m)$ some user $j \in [m]$, and a bid $b'_j > b_j$, such that the user j is incentivized to overbid under the old utility notion, but not incentivized to overbid under the new utility notion. There are the following cases, and we rule each one out, which allows us to reach a contradiction. Below we use the terms "included" and "confirmed" interchangeably, and we define $\mathbf{b}' := (\mathbf{b}_{-j}, b'_j)$.

- *Case 1: user j is confirmed with probability* 1 *under* $\mathbf{b}'$. In this case, user j's utility is the same under the old and new utility definitions, and therefore, it is not possible that user j wants to deviate under the old utility but does not want to under the new utility notion.

- *Case 2: user j is unconfirmed with probability* 1 *under* $\mathbf{b}'$. In this case, under the old utility, user j's utility is 0 even when it bids b'_j. Therefore, user j does not want to deviate under the old utility notion which contradicts our assumption.

- *Case 3: user j sometimes confirmed and sometimes unconfirmed under* $\mathbf{b}'$. Since the TFM is weak MIC and 1-weak-SCP, and satisfies "included = confirmed", by Lemma 7.3, whenever the user j is confirmed, it must pay its full bid. Therefore, under the old utility notion, if user j bids b'_j instead, its utility is always 0. This means that the user does not want to deviate under the old utility notion, which contradicts our assumption.

□

Proof of Corollary 7.1. We now continue with the proof of Corollary 7.1. The proof of Lemma 4.8 also makes use of the strategic deviation where a user colluding with the miner overbids relative to its true value. Specifically, the proof of Lemma 4.8 relies on the fact that such overbidding comes for free if the offending transaction is not confirmed. In general, this is not true under

the new utility function associated with weak incentive compatibility. However, we now argue that if "included" must be equal to "confirmed", then, the effect of this deviation (where the overbid transaction is not confirmed) can alternatively be realized in a way that is free of charge.

More concretely, instead of having the colluding user actually carry out the overbidding, we instead exploit the miner's ability to include an arbitrary subset of the mempool in the block. Let $\mathbf{b}$ be the current mempool, which includes the colluding user's bid b. If in the proof of Lemma 4.8, the miner wants the user to overbid $b' > b$ instead, it can simply pretend that the colluding user's bid is b'. In other words, the miner can simulate running the mechanism on $\mathbf{b} \backslash \{b\} \cup \{b'\}$. As a result, the transaction b' would not be confirmed, and thus b' would not be included in the block, either. Therefore, the fact that the colluding user has not authorized/signed the transaction b' does not matter in carrying out this deviation[5].

It is easy to see that as long as Myerson's Lemma holds and any overbid transaction that is not confirmed in the present block comes for free, Lemma 4.8 still holds. Therefore, we conclude that Lemma 4.8 still holds even under weak UIC and 1-weak-SCP, if we insist that "included" be equal to "confirmed". Now, as long as Lemma 4.8 and Myerson's Lemma still hold, the proof of Theorem 4.7 follows in the same way as before, and so does the proof of Corollary 4.9.

If we restrict ourselves to *deterministic* mechanisms, we can actually prove a counterpart of Corollary 7.1 without having to even rely on weak MIC. This is formally stated in the following corollary:

Corollary 7.5. *Assume that all transactions included in the block must be confirmed. Then, no deterministic TFM* $(\mathbf{x}, \mathbf{p}, \mu)$ *with non-trivial miner revenue can satisfy weak UIC and 1-weak-SCP at the same time — this impossibility holds no matter whether the block size is finite or infinite. Moreover, if the block size is finite, then the only deterministic TFM that achieves weak UIC and 1-weak-SCP is the trivial mechanism that always confirms nothing and pays the miner nothing.*

Proof. Almost the same as the proof of Corollary 7.1, except that now, since the TFM is promised to be deterministic, we can use Fact 6.2 instead of Lemma 7.4 to establish the fact that Myerson's lemma still holds. □

8 Conclusion and Open Questions

Mechanism design in decentralized settings (e.g., cryptocurrencies) departs significantly from the classical literature in terms of modeling and assumptions, and thus is relatively little understood. For example, our work shows that even how to formally define incentive compatibility is subtle and requires careful thought. Our work helps to unravel the mathematical structures of incentive compatible TFMs, we hope that our definitional contributions can serve as a basis for future work in this space. Our work also raises more open questions than the ones we can answer. For example, are there other reasonable relaxations in the modeling and in incentive compatible notions that allow us to circumvent the impossibility results we showed? Can we formally model and reason about the repeated nature of the TFM, and reason about potential strategic behavior over a longer time scale? Can cryptography help in the design of transaction fee mechanisms? For example, the elegant work of Ferreira and Weinberg [FW20] showed that using cryptographic commitments can help overcome some of the lower bound results shown by Akbarpour and Li [AL20]. However, as mentioned in Section 2.5, the modeling approach there is fundamentally incompatible with the setting for transaction fee mechanisms. Besides these, Roughgarden [Rou20,Rou21b] also presented a comprehensive list of open questions, most of which remain unanswered.

[5] In other words, b' exists only in the simulation in the miner's head, but is not released to the public network.

Acknowledgments

We gratefully acknowledge helpful technical discussions with Kai-Min Chung during an early phase of the project. We also thank T-H. Hubert Chan for insightful technical discussions.

References

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

[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.

[BSKN21] Adithya Bhat, Nibesh Shrestha, Aniket Kate, and Kartik Nayak. Randpiper - reconfiguration-friendly random beacons with quadratic communication. In *ACM CCS*, 2021.

[CKS00] Christian Cachin, Klaus Kursawe, and Victor Shoup. Random oracles in constantinople: Practical asynchronous byzantine agreement using cryptography. In *in Proc. 19th ACM Symposium on Principles of Distributed Computing (PODC*, pages 123–132, 2000.

[CM12] Jing Chen and Silvio Micali. Collusive dominant-strategy truthfulness. *J. Econ. Theory*, 147(3):1300–1312, 2012.

[DKIR21] Sourav Das, Vinith Krishnan, Irene Miriam Isaac, and Ling Ren. Spurt: Scalable distributed randomness beacon with transparent setup. Cryptology ePrint Archive, Report 2021/100, 2021. https://ia.cr/2021/100.

[DM17] Alan Deckelbaum and Silvio Micali. Collusion, efficiency, and dominant strategies. *Games Econ. Behav.*, 103:83–93, 2017.

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

[FW20] Matheus V. X. Ferreira and S. Matthew Weinberg. Credible, truthful, and two-round (optimal) auctions via cryptographic commitments. In *Proceedings of the 21st ACM Conference on Economics and Computation (EC)*, page 683712, 2020.

[GH05] Andrew V. Goldberg and Jason D. Hartline. Collusion-resistant mechanisms for single-parameter agents. In *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2005, Vancouver, British Columbia, Canada, January 23-25, 2005*, pages 620–629. SIAM, 2005.

[GHK+06] Andrew V. Goldberg, Jason D. Hartline, Anna R. Karlin, Michael E. Saks, and Andrew Wright. Competitive auctions. *Games Econ. Behav.*, 55(2):242–269, 2006.

[GL79] Jerry Green and Jean-Jacques Laffont. On coalition incentive compatibility. *The Review of Economic Studies*, 46(2):243–254, 04 1979.

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

[kCK09] Yeon koo Che and Jinwoo Kim. Optimal collusion-proof auctions. *Journal of Economic Theory*, pages 565–603, 2009.

[LSZ19] Ron Lavi, Or Sattath, and Aviv Zohar. Redesigning bitcoin's fee market. In *The World Wide Web Conference, WWW 2019, San Francisco, CA, USA, May 13-17, 2019*, pages 2950–2956. ACM, 2019.

[MM12] Robert C. Marshall and Leslie M. Marx. *The Economics of Collusion: Cartels and Bidding Rings*. The MIT Press, 2012.

[Mon20] Barnabe Monnot. A transaction fee market proposal. https://github.com/ethereum/rig/blob/master/eip1559/eip1559.ipynb, 2020.

[Mye81] Roger B. Myerson. Optimal auction design. *Math. Oper. Res.*, 6(1):5873, February 1981.

[NRTV07] Noam Nisan, Tim Roughgarden, Eva Tardos, and Vijay V. Vazirani. *Algorithmic Game Theory*. Cambridge University Press, USA, 2007.

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

[Rou21a] Private communication with Tim Roughgarden and Matt Weinberg, 2021.

[Rou21b] Tim Roughgarden. Transaction fee mechanism design. In *EC*, 2021.

[Vic61] William Vickrey. Counterspeculation, auctions, and competitive sealed tenders. *Journal of finance*, 1961.

[Yao18] Andrew Chi-Chih Yao. An incentive analysis of some bitcoin fee designs. *CoRR*, abs/1811.02351, 2018.

A Relations Between Incentive Compability Notions

The notions UIC, MIC, and 1-SCP are incomparable as depicted in Figure 2.

We explain Figure 2 in more detail below:

- UIC $\not\Rightarrow$ MIC, UIC $\not\Rightarrow$ 1-SCP: the second-price auction satisfies UIC, but does not satisfy MIC or 1-SCP. This was pointed out in several earlier works [BEOS19, Rou20, Rou21b]. Recall that in the second-price auction, the highest B bids are included in the block, the top $B-1$ are confirmed and they pay the B-th price, where B is the block size. The miner gets all payment.

- MIC $\not\Rightarrow$ UIC, 1-SCP $\not\Rightarrow$ UIC: the first-price auction satisfies MIC and c-SCP for any $c \geq 1$, but is not UIC. This was also pointed out in earlier works [BEOS19, Rou20, Rou21b]. Recall that in the first-price auction, the top B bids are included and confirmed, they each pay their bid, and the miner gets all payment.

- MIC $\not\Rightarrow$ 1-SCP: the posted price auction satisfies MIC but not 1-SCP. Recall that in the posted-price auction, there is a fixed reserve price r. Everyone bidding at least r is included and confirmed and pays exactly r. The miner gets all payment. It is easy to check that the mechanism is indeed MIC. However, it is not 1-SCP, since if a user's true value is $0 < r' < r$, the miner can

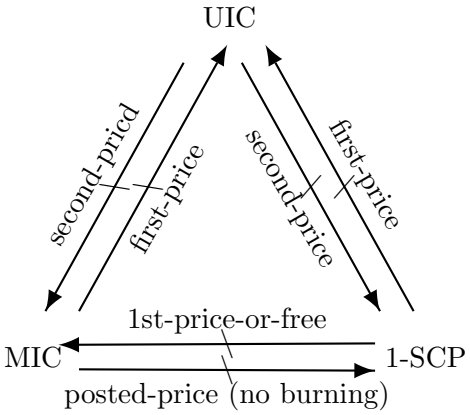

Figure 2: Relationship among incentive compatibility notions. The same chart holds for UIC, MIC, and 1-SCP under γ-strict-utility for any $\gamma \in [0, 1]$.

collude with the user, have the user bid r instead, and the joint utility of the coalition strictly increases.

- 1-SCP $\not\Rightarrow$ MIC: this is the most subtle to see. We construct the following "first-price-or-free" mechanism which is c-SCP for any $c \geq 1$, but not MIC. The mechanism is not MIC since if there is only one bid, it makes sense for the miner to inject a fake bid to increase its utility. We show that the mechanism satisfies c-SCP for any $c \geq 1$ below.

First-price-or-free mechanism

- Choose all bids in the current bid vector to include in the block. Let $\mathbf{b} = (b_1, \ldots, b_m)$ be the included bids of the block, where $b_1 \geq \cdots \geq b_m$.
- Only the highest bid (b_1) is confirmed. Every other bid is unconfirmed.
- If there is only one bid in the block ($m = 1$), the only confirmed user pays nothing. Otherwise, if $m \geq 2$, the only confirmed user pays b_1.
- The miner gets all the payment.

Theorem A.1. *The first-price-or-free mechanism is c-SCP for all $c \geq 1$.*

Proof. Suppose there is only one user with true value v^*. The miner and that user is the only possible coalition. If they play honestly, that user's bid is the only bid in the block, so it must be confirmed. Thus, in the honest case, the joint utility is v^*. If the coalition now deviates, then the confirmed bid is either a fake bid or the colluding user's bid. In either case, the coalition's utility cannot exceed v^*.

Suppose the number of users is $m \geq 2$. There are two cases. First, suppose the highest bid b_1 does not belong to the coalition if all colluding users are bidding truthfully. In this case, the coalition's utility is b_1 when it behaves honestly, and $b_1 \geq v^*$ where v^* denotes the highest true value of any colluding user. Second, the highest bid b_1 belongs to the coalition if colluding users are bidding truthfully. In this case, the coalition's utility is $b_1 = v^*$ if it behaves honestly. In either case, we show that if the coalition deviates, it cannot gain. Suppose b_1' is the new highest bid after deviating. If b_1' belongs to the coalition, then the miner revenue offsets the coalition's payment,

and thus the coalition's utility cannot exceed the highest true value of any colluding user. If b'_1 does not belong to the coalition, it must be that $b'_1 \le b_1$, and the coalition's utility is $b'_1 \le b_1$. □

Relationship for incentive compatibility notions under γ-strict-utility. Note that in Figure 2, for each arrow $X \not\Rightarrow Y$ shown by some example mechanism, it is easy to check that the same mechanism also shows that $X \not\Rightarrow$ weak Y. Thus, Figure 2 in fact also holds for UIC, MIC, and 1-SCP under γ-strict-utility, for any choice of $\gamma \in [0, 1]$.

B Additional Results for Weak Incentive Compatibility

In this section, we present some additional results that further unfold the mathematical landscape of weakly incentive compatible mechanisms.

B.1 The Solitary Mechanism

Earlier in Section 6, we ruled out the existence of a deterministic, 2-user-friendly mechanism that satisfies weak UIC and 2-weak-SCP simultaneously, assuming finite block size. In this section, we show that the 2-user-friendly restriction is necessary for this theorem to hold. In particular, we describe a mechanism called the solitary mechanism, which always confirms a single transaction, and satisfies weak UIC, weak MIC, and c-weak-SCP for all $c \in \mathbb{N}$.

The solitary mechanism

- Choose the highest two bids to include in the block.
- Only the highest bid is confirmed. Other bids are all unconfirmed. The highest bid pays the second highest bid, and the miner is paid the second highest bid.

Theorem B.1 (The solitary mechanism)**.** *The solitary mechanism satisfies weak UIC, weak MIC, and c-weak-SCP for all $c > 0$. The theorem holds no matter the block size is infinite or finite.*

Proof. We prove the three properties one by one.

Weak UIC. A user has two kinds of strategies to deviate from the honest behavior: to bid strategically or to inject fake transactions. Since the payment is decided by the second highest bid, injecting fake transactions can only increase the payment, no matter whether the user is bidding truthfully or not. Moreover, since this is exactly a classical second-price auction, bidding truthfully is known to be DSIC for an individual user, even under the old utility notion where overbidding is never penalized. Therefore, bidding untruthfully is not incentive compatible under the new utility notion as well.

Weak MIC. The miner has two kinds of strategies to deviate from honest behavior: not to choose the highest two bids and to inject fake bids. Without loss of generality, we assume that the miner chooses the included bids first, and replaces some of them with fake bids then. We will show that both steps would not increase the miner's utility.

The miner's revenue is decided by the second highest bid that is included, so if miner does not choose the highest two bids to include, its revenue can only decrease or remain the same. Next, suppose that the miner replaces one or both of the included bids with fake ones. If after the

replacement, the highest bid is a fake one, then the miner has to pay the fee for the highest bid which is equal to its revenue. Therefore, the miner's utility cannot be greater than 0. If after the replacement, the highest bid is not a fake one but the second highest bid is a fake one, and the bid amount is b. Then, the miner revenue is b. However, the cost to inject the fake bid b is also b. Thus, the miner does not gain overall.

Weak c-SCP. Consider an arbitrary coalition of the miner and a subset of the users. Suppose the coalition plays honestly:

- if the top confirmed bidder is in the coalition, the coalition's utility is top bidder's true value denoted v_1;
- if the top confirmed bidder is not in the coalition, then the coalition's utility is the true value of the 2nd bidder $v_2 \leq v_1$.

Now, consider an arbitrary strategy where two bids are included and the higher of the two gets confirmed. Each included bid can either come from some user, or is a fake bid. Without loss of generality, we may equivalently assume that a fake bid belongs to some imaginary user which belongs to the coalition, and its true value is 0. We may use the fake indices 0 and -1 to refer to the one or two imaginary users. There are the following cases:

- *Case 1: The confirmed user i belongs to the coalition.* In this case, the coalition's utility is upper bounded by (the possibly imaginary) user i's true value v_i. If i has the highest true value, it means the coalition's utility is upper bounded by v_1; else if i does not have the highest true value, it means that the coalition's utility is upper bounded by v_2. Either way, the coalition's utility cannot exceed the aforementioned honest case.
- *Case 2: The confirmed user i does not belong to the coalition.* In this case, suppose that the included but unconfirmed user is denoted j where j is possibly an imaginary user. The coalition's utility is $b_j - \max(0, b_j - v_j) \leq v_j$ where b_j is user j's bid, v_j is its true value, and the part $\max(0, b_j - v_j)$ is the penalty due to overbidding. Since the confirmed user i does not belong to the coalition, it must be a real user and it must be bidding its truthful value, i.e., $b_i = v_i$. Note that the b_j cannot be bidding higher than b_i since b_j is unconfirmed but b_i is confirmed. Therefore, $v_j \leq b_j \leq b_i = v_i$. This also implies that that $v_j \leq v_2$, and thus the coalition's utility is also upper bounded by v_2. Recall that the coalition's utility is at least v_2 had it played honestly; therefore, the coalition does not gain anything in comparison with playing honestly.

□

B.2 The Solitary-Or-Posted-Price Mechanism

Earlier in Section 6, we ruled out the existence of a deterministic, 2-user-friendly mechanism that satisfies weak UIC and 2-weak-SCP simultaneously, assuming *finite block size*. In this section, we show that the finite block size restriction is necessary for this lower bound to hold, by showing a deterministic, 2-user-friendly mechanism that satisfies weak UIC, weak MIC, and 2-weak-SCP, but only under *infinite* block size.

The solitary-or-posted price mechanism

Parameters: a reserve price r.

Mechanism:

- Choose the top two bids as well as every other bid that is at least r to include in the block.
- Every bid at least r is confirmed, and the highest bid in the block is always confirmed (even if it is smaller than r).
- Let b_2 be the second highest bid in the block. Every confirmed bid pays $\min(b_2, r)$, and miner is paid $\min(b_2, r)$. The remaining payment is burnt.

Theorem B.2 (Solitary-or-posted-price mechanism)**.** *Suppose the block size is infinite. The solitary-or-posted-price mechanism satisfies weak UIC, weak MIC, and c-weak-SCP for all $c > 0$.*

Proof. We prove the three properties one by one.

Weak UIC. A user has two kinds of strategies to deviate from the honest behavior: to bid strategically or to inject fake transactions. Since the payment is decided by $\min(b_2, r)$, injecting fake transactions can only increase the payment, no matter whether the user is bidding truthfully or not.

Suppose the real bid vector is $\mathbf{b} = (b_1, \ldots, b_m)$, where $b_1 \geq \cdots \geq b_m$. Suppose user i bids truthfully and other users may bid arbitrarily. Let v_i and b_i' be user i's true value and strategic bid, respectively.

- *Case 1:* $b_2 \geq r$. User i is facing a posted-price auction such that it if confirmed if and only if $b_i' \geq r$. In this case, it is not hard to see that no matter user i overbids or underbids, its utility does not increase.
- *Case 2:* $b_2 < r$. When $v_i = b_1$, then user i't utility is $v_i - b_2 \geq 0$ in the honest case. If user i overbids, it still pays b_2 so the utility does not change. If user i underbids, it is either confirmed with the same payment, or becomes unconfirmed. In either case, the utility does not increase.

Weak MIC. The miner has two strategies to deviate: not to choose the highest two bids and to inject fake bids. Without loss of generality, we assume that the miner chooses the included bids first, and replaces some of them with fake bids then. We will show that both steps would not increase the miner's utility. The miner's revenue is decided by the second highest bid that is included, so if miner does not choose the highest two bids to include, its revenue can only decrease or remain the same.

Now, suppose the miner replaces some of the included bids with fake ones. If any fake bid is confirmed, then the fake bid must be paying an amount equal to the miner revenue, and thus the miner's utility is at most 0. If no fake bid is confirmed, and some fake bid is unconfirmed and its bid amount is b. Then, it must be that b is the second highest bid and the highest bid is smaller than r. In this case, the miner gets revenue b; however, it costs b to inject this fake bid. Thus, the miner does not gain overall.

Weak c-SCP. Suppose there are m users, and their true values are $(v_1, \ldots, v_m)$ where $v_1 \geq \cdots \geq v_m$. Henceforth, we also call the user with the highest true value v_1 the top user. There are two possible cases.

- *Case 1:* $v_2 < r$. When everyone behaves honestly, the miner's revenue is v_2, user 1's utility is $v_1 - v_2$, and all other users are zero since they are unconfirmed. Suppose the miner colludes with

a subset of users, and they prepare a bid vector $\mathbf{e} = (e_1, \ldots, e_m)$ where $e_1 \geq \cdots \geq e_m$. Each bid e_i in $\mathbf{e}$ is either a non-colluding bid coming from a non-coalition user in which case $e_i = v_i$, or it is a colluding bid, i.e., one that comes from a colluding user or a fake bid. If only one user i is confirmed in $\mathbf{e}$, there are two possibilities.

- Suppose user i is not in the coalition. The utility of the coalition is miner's revenue (e_2) minus potential extra cost if there are overbid or fake bids that are unconfirmed. If i is not the top user, then $e_2 \leq v_i \leq v_2$. This means the utility of the coalition cannot exceed v_2, which can be achieved by playing honestly.

  If i is top user, to make the miner's revenue larger than the honest case, it must be that e_2 is a colluding bid and $e_2 > v_2$. Let v' be the true value of this colluding user or $v' = 0$ if e_2 is fake. Since e_2 is unconfirmed in $\mathbf{e}$, its utility becomes $v' - e_2$. Thus, the utility of the coalition cannot exceed $v' \leq v_2$, which can be achieved by playing honestly.
- Suppose user i is in the coalition. The utility of the coalition is v_i minus potential extra cost. However, if top user is also in the coalition, the utility of the coalition is v_1 in the honest case, and is at most $v_i \leq v_1$ in the strategic case. If the top user is not in the coalition, the utility of the coalition is v_2 in the honest case, and is at most $v_i \leq v_2$ in the strategic case.

If there are two or more confirmed bids in $\mathbf{e}$, the miner's revenue becomes r, while each confirmed user needs to pay r. In this case, the utility of the top user decreases by $r - v_2$. The utilities of all other bids (including fake ones) are non-positive, since if they are confirmed, they have to pay r which is higher than their true values. Since $v_2 < r$, there must be a colluding user (that is not the top user) bidding $b' \geq r$ or the miner injects a fake bid $b' \geq r$. Let v' be the true value of this colluding user or $v' = 0$ if the bid is fake. The utility of this bid is $v' - r$. The joint utility of this colluding or fake bid and the miner is $v' \leq v_2$, so the joint utility does not increase.

- *Case 2:* $v_2 \geq r$. When everyone behaves honestly, the miner's utility is r, user i's utility is $v_i - r$ for all confirmed user i, and all other users are zero since they are unconfirmed. In this case, miner's utility is already maximized. Suppose the miner colludes with a subset of users, and they prepare a bid vector $\mathbf{e}$. If there is only one confirmed bid in $\mathbf{e}$, then only the confirmed user can benefit by the deviation since its payment decreases. However, the amount that the confirmed user gains is exactly what the miner loses, so the joint utility of any coalition does not increase. If there are two or more confirmed bids in $\mathbf{e}$, then every user's utility is maximized when they bid truthfully, since the payment is fixed at r regardless of others' bids.

□

B.3 Necessity of Burning

Earlier in Section 6, we ruled out the existence of a deterministic, 2-user-friendly mechanism that satisfies weak UIC and 2-weak-SCP simultaneously, assuming *finite block size*. In this section, we show that if the mechanism is not allowed to use a burning mechanism, i.e., if the miner's payment must be the sum of all users' payment, then, the same lower bound would hold even under infinite block size. This lower bound also shows that the burning in the solitary-or-posted-price mechanism is necessary.

Theorem B.3. *Let $(\mathbf{x}, \mathbf{p}, \mu)$ be a deterministic mechanism without burning. If $(\mathbf{x}, \mathbf{p}, \mu)$ is 2-user-friendly, then it cannot achieve UIC and 2-weak-SCP at the same time.*

The remainder of this section will focus on proving Theorem B.3. We first prove a useful lemma that says in a mechanism satisfying the desired properties, if in some bid vector, all unconfirmed bids are bidding 0, then all confirmed bids must be paying 0.

Lemma B.4. *Let $(\mathbf{x}, \mathbf{p}, \mu)$ be a deterministic mechanism without burning that is weak UIC and 2-weak-SCP. Suppose there exists a bid vector $\mathbf{b} = (b_1, \ldots, b_m)$ that confirms at least one bid, and moreover, all unconfirmed bids are 0. Then, all confirmed bids must pay 0.*

Proof. Due to Lemma 6.4, let $p := p(\mathbf{b})$ denote the universal payment for $\mathbf{b}$. Let $\epsilon < p/2m$ be a sufficiently small positive number. By Lemma 6.3 and Myerson's Lemma, we can change all confirmed bids in $\mathbf{b}$ to $p+\epsilon$ such that all confirmed bids in $\mathbf{b}$ remain confirmed, and their payment unaffected. Let $\mathbf{b}'$ be the resulting bid vector, and let $u \in [m]$ be the number of confirmed users in $\mathbf{b}'$, and recall all unconfirmed users bid 0. Since there is no burning, $\mu(\mathbf{b}') = u \cdot p$. Let i be a confirmed user in $\mathbf{b}'$. Now, suppose that the real bid vector is actually $\mathbf{b}'' := (\mathbf{b}'_{-i}, p - \epsilon)$, and this also represents everyone's true value. User i becomes unconfirmed in $\mathbf{b}''$ by Myerson's Lemma. Thus $\mu(\mathbf{b}'') \le (p+\epsilon) \cdot (u-1)$. In this case, the miner can collude with user i and ask it to bid $p+\epsilon$ instead. In this case, user i's utility is $-\epsilon$, however, the miner's revenue is $p \cdot u > \mu(\mathbf{b}'') + \epsilon$. Thus, the coalition can strictly gain from this deviation, which violates 1-weak-SCP.

□

Lemma B.5. *Let $(\mathbf{x}, \mathbf{p}, \mu)$ be a deterministic mechanism without burning which is 2-user-friendly, weak-UIC, 2-weak-SCP, and with non-trivial miner revenue. Then, there exists a bid vector $\mathbf{b} = (b_1, \ldots, b_m)$ where two different users i, j are confirmed, and moreover, $\mu(\mathbf{b}) > 0$, $b_i > p_i(\mathbf{b})$ and $b_j > p_j(\mathbf{b})$.*

Proof. It suffices to show that there exists a bid vector $\mathbf{b}$ such that $\mu(\mathbf{b}) > 0$ and at least two users' bids are confirmed. If so, we can use the same argument in the proof of Lemma 6.10 to show that there exists a bid vector $\mathbf{b}'$ such that $\mu(\mathbf{b}') > 0$, and moreover at least two users' bids are confirmed, and they are both bidding strictly higher than their payment. In particular, cases 1 and 2 follow just like the proof of Lemma 6.10. For case 3, suppose that $\mathbf{b}$ is a bid vector such that $\mu(\mathbf{b}) > 0$ and two different users i and j are confirmed, and both bid exactly their payment. In this case, the proof of Lemma 6.10 constructed a new bid vector $\mathbf{b}'$ which is otherwise equal to $\mathbf{b}$ except that b_i and b_j now bid $b_i + \Delta$ for an arbitrary $\Delta > 0$, and showed that under $\mathbf{b}'$, both i and j are confirmed and bidding strictly above payment. Here, we only need to additionally argue that $\mu(\mathbf{b}') > 0$. This can be achieved by choosing Δ to be sufficiently small, and then applying Lemma 6.7.

Therefore, below, we focus on proving that there exists a bid vector $\mathbf{b}$ such that $\mu(\mathbf{b}) > 0$ and at least two users' bids are confirmed. Suppose this is not true. In other words, for any bid vector $\mathbf{b}$ satisfying $\mu(\mathbf{b}) > 0$, only one user is confirmed. We will show that this contradicts 2-weak-SCP.

Since the mechanism is 2-user-friendly, Lemma 6.10 guarantees that there exists a bid vector $\mathbf{c} = (c_1, \ldots, c_m)$ such that $x_1(\mathbf{c}) = x_2(\mathbf{c}) = 1$ and $c_1 > p(\mathbf{c})$ and $c_2 > p(\mathbf{c})$. By our assumption, it must be $\mu(\mathbf{c}) = 0$. Because there is no burning, we have $p(\mathbf{c}) = 0$. By Lemma 6.3, we can increase user 1's and user 2's bids arbitraily without changing their confirmation and payment. As a result, we obtain $\mathbf{c}' = (\Gamma, \Gamma, c_3, \ldots, c_m)$, where $\Gamma = |\mathbf{c}|_1$. By Lemma 6.9, we can now reduce each bid $b_3, \ldots, b_m$ down to zero one by one, without changing user 1's and user 2's confirmation. Formally, we obtain a bid vector $\mathbf{c}'' = (\Gamma, \Gamma, 0, \ldots, 0)$ such that $x_1(\mathbf{c}'') = x_2(\mathbf{c}'') = 1$. By our assumption, both confirmed users in $\mathbf{c}''$ are paying 0.

Since the mechanism has non-trivial miner revenue, there must exist a bid vector $\mathbf{b}$ where $\mu(\mathbf{b}) > 0$. By our assumption, only one user denoted i is confirmed in $\mathbf{b}$. By Lemma B.4, there

must be another user j bidding non-zero. Suppose $\mathbf{b}$ also represents everyone's true value. In this case, miner's utility is $\mu(\mathbf{b})$, user i's utility is $b_i - \mu(\mathbf{b})$, and user j's utility is zero. The miner can collude with user i and user j, and ask them to bid Γ instead. Then, the coalition prepares a bid vector $\mathbf{c}'' = (\Gamma, \Gamma, 0, \ldots, 0)$ where the first two bids are user i and user j's bids. In this case, the miner's utility is zero, while user i's utility is b_i and user j's utility is b_j. The joint utility increases by $b_j > 0$, which violates 2-weak-SCP. □

The proof of Theorem B.3. By Lemma 6.4 and Lemma B.5, there exists a bid vector $\mathbf{b}^{(0)} = (b_1, \ldots, b_m)$ such that $\mu(\mathbf{b}^{(0)}) > 0$, $x_1(\mathbf{b}^{(0)}) = x_2(\mathbf{b}^{(0)}) = 1$, $b_1 > p(\mathbf{b}^{(0)})$ and $b_2 > p(\mathbf{b}^{(0)})$ — note that we can always relabel the bids to make the first two bids represent two confirmed bids.

Now, one by one, we reduce every unconfirmed bid down to zero and increase every confirmed bid to a sufficiently large value $\Gamma > |\mathbf{b}|_1$. Formally, for $i = 1, \ldots, m$, we define

$$\mathbf{b}^{(i)} = \begin{cases} (\mathbf{b}_{-i}^{(i-1)}, 0), & \text{if } x_i(\mathbf{b}^{(i-1)}) = 0, \\ (\mathbf{b}_{-i}^{(i-1)}, \Gamma), & \text{if } x_i(\mathbf{b}^{(i-1)}) = 1. \end{cases}$$

By Lemma 6.3, when increasing user 1's and user 2's bids, their confirmation, payment, and miner revenue do not change. Later on, when increasing any confirmed user i's bid where $i > 2$, any previous user bidding Γ would still remain confirmed and pay the same. When decreasing any unconfirmed user i's bid to 0 where $i > 2$, since Γ is sufficiently large, and by Lemma 6.9, any previous user bidding Γ would remain confirmed, and although their payment may change, change in the payment is slow. Thus, at the end, users 1 and 2 are confirmed in the final vector $\mathbf{b}^{(m)}$.

By Lemma B.4 and the fact that there is no burning, it must be that $\mu(\mathbf{b}^{(m)}) = 0$. Let i^* be the smallest integer $i \in \{1, \ldots, m\}$ such that $\mu(\mathbf{b}^{(i)}) = 0$. Then, we have $\mu(\mathbf{b}^{(i^*-1)}) > 0$ and $\mu(\mathbf{b}^{(i^*)}) = 0$. By Lemma 6.3, increasing a confirmed user's bid does not change miner revenue, so user i^* must be unconfirmed in $\mathbf{b}^{(i^*-1)}$. Imagine the real bid vector is $\mathbf{b}^{(i^*-1)}$ which also represents everyone's true value. In this case, the miner's revenue is $\mu(\mathbf{b}^{(i^*-1)})$, user 1's utility is $\Gamma - p(\mathbf{b}^{(i^*-1)})$, and user i^*'s utility is zero. The miner can collude with user 1 and user i^*, and ask user i^* to bid Γ instead. The coalition now prepares a bid vector $\mu(\mathbf{b}^{(i^*)})$ where the second coordinate Γ actually comes from user i^* and $b_{i^*} = 0$ is a fake bid injected by the miner. Since there is no burning and $\mu(\mathbf{b}^{(i^*)}) = 0$, the payment must be zero. Therefore, the miner's revenue becomes zero, while user 1's utility becomes Γ, and user i^*'s utility becomes b_{i^*}. By Lemma 6.7, we have $b_{i^*} \geq \mu(\mathbf{b}^{(i^*-1)})$, and thus the coalition strictly gains from this deviation, which violates 2-weak-SCP.