What Can Cryptography Do For Decentralized Mechanism Design?

Elaine Shi, Hao Chung, and Ke Wu*

Carnegie Mellon University
{runting@cs, haochung@andrew, kew2@andrew}.cmu.edu

Abstract

Recent works of Roughgarden (EC'21) and Chung and Shi (SODA'23) initiate the study of a new decentralized mechanism design problem called transaction fee mechanism design (TFM). Unlike the classical mechanism design literature, in the decentralized environment, even the auctioneer (i.e., the miner) can be a strategic player, and it can even collude with a subset of the users facilitated by binding side contracts. Chung and Shi showed two main impossibility results that rule out the existence of a *dream* TFM. First, any TFM that provides incentive compatibility for individual users and miner-user coalitions must always have zero miner revenue, no matter whether the block size is finite or infinite. Second, assuming finite block size, no non-trivial TFM can simultaneously provide incentive compatibility for any individual user and for any miner-user coalition.

In this work, we explore what new models and meaningful relaxations can allow us to circumvent the impossibility results of Chung and Shi. Besides today's model that does not employ cryptography, we introduce a new MPC-assisted model where the TFM is implemented by a joint multi-party computation (MPC) protocol among the miners. We prove several feasibility and infeasibility results for achieving *strict* and *approximate* incentive compatibility, respectively, in the plain model as well as the MPC-assisted model. We show that while cryptography is not a panacea, it indeed allows us to overcome some impossibility results pertaining to the plain model, leading to non-trivial mechanisms with useful guarantees that are otherwise impossible in the plain model. Our work is also the first to characterize the mathematical landscape of transaction fee mechanism design under approximate incentive compatibility, as well as in a cryptography-assisted model.

*Author order is randomized.

Contents

1 Introduction

The widespread adoption of blockchains and cryptocurrencies spurred a new class of *decentralized* mechanism design problems. The recent works of Roughgarden [Rou20, Rou21] as well as Chung and Shi [CS21] considered a particularly important decentralized mechanism design problem, that is, *transaction fee mechanism* (TFM) design. In a transaction fee mechanism (TFM), we are auctioning space in the block to users who want their transactions included and confirmed in the block. If the block can contain up to k transactions, one can equivalently think of selling k identical products to the bidders.

Prior works [LSZ19, Yao, BEOS19, BCD+, Rou20, Rou21, FMPS21] observed that transaction fee mechanism design departs significantly from classical mechanism design [NRTV07]. The vast majority of classical auctions assume that the auctioneer honestly implements the prescribed mechanism. In comparison, in a blockchain environment, the auctioneer (i.e., the miner of the block), can be a strategic player in itself: it can deviate from the prescribed mechanism if it increases its expected gain; or it can collude with a subset of the users, and play strategically to improve the coalition's joint utility. As earlier works pointed out [LSZ19, Yao, BEOS19, BCD+, Rou20, Rou21], the existence of decentralized smart contracts in blockchain environments make it easy for the miner and users to rendezvous and engage in *binding* side contracts. Such side contracts allow the coalition to split their gains off the table in a binding fashion.

Observing the new challenges that arise in a decentralized environment, earlier works [LSZ19, Yao, BEOS19, BCD+, Rou20, Rou21] formulated a set of desiderata for a "dream" TFM:

- *User incentive compatibility (UIC)*: a user's best strategy is to bid truthfully, even when the user has observed others' bids.
- *Miner incentive compatibility (MIC)*: the miner's best strategy is to implement the honest mechanism, even when the miner has observed all users' bids.
- *c-side-contract-proofness (c-SCP)*: playing honestly maximizes the joint utility of a coalition consisting of the miner and at most c users, even after having observed all others' bids.

A line of works explored how to get a dream TFM. However, assuming that the block size is *finite*, i.e., there can be more bids than the block size, all known works fall short of achieving all three properties at the same time. The closest we have come to in terms of achieving a dream TFM is in fact Etherem's EIP-1559. At a very high-level, when there is congestion, EIP-1559 behaves like a first-price auction which is not UIC. When the block size is infinite (i.e., no congestion), EIP-1559 approximates the following "burning posted price" auction: there is a fixed reserve price r, every bid that is at least r gets included and confirmed, and pays the price of r. All users' payment is burnt and the miner gets nothing[1]. Roughgarden [Rou20, Rou21] proved that when the block size is *infinite*, indeed, the burning posted price auction achieves all three properties at the same time!

Subsequently, Chung and Shi [CS21] further explored the landscape of TFM. They proved two interesting impossibility results:

1. *Zero miner revenue.* Any (possibly randomized) TFM that satisfies both UIC and SCP must always have 0 miner revenue, even when the miner colludes with at most one user, and no matter whether the block size is finite or infinite. This shows that the total burning in EIP-1559 is no accident: it is necessary to achieve all three properties under infinite block size.

[1] In practice, the miner gets a fixed block reward that is irrelevant to our game-theoretic analysis, so we ignore the fixed block reward in our modeling.

2. *Finite-block impossibility.* Suppose that block size is finite, then no non-trivial (possibly randomized) TFM can achieve UIC and SCP at the same time, even when the miner colludes with at most one user. This shows that it is no accident that all prior works fail to achieve the dream TFM for finite block sizes — indeed, there is a mathematical impossibility!

Given the status quo of our understanding, we ask the following natural question:

Are there meaningful new models or relaxations that allow us to circumvent the impossibility results of Chung and Shi?

Chung and Shi [CS21] made an initial exploration along this line. They show a relaxation that allows us to circumvent the impossibilities and achieve positive miner revenue under finite block size. In particular, their relaxation requires the additional assumption that offending bids (e.g., overbid or fake transactions) that have been posted to the public cannot be retracted in the future, and thus the offender may have to pay a cost when the offending transaction is confirmed in the future. While this assumption holds for some cryptocurrencies such as Bitcoin, it may not be universally true for all cryptocurrencies. Therefore, an important question is what other models or relaxations allow us to circumvent the impossibilities.

In this paper, we explore two new directions, aiming to understand whether they allow us to circumvent the impossibilities of Chung and Shi [CS21]: *i)* using an approximate notion of incentive compatibility that allows an ϵ additive slack; and *ii)* having the miners jointly run a multi-party computation (MPC) protocol to realize the TFM. Throughout the paper, we refer to the today's model, which does employ cryptography, as the *plain model*, and we refer to the case where the TFM is realized with MPC as the *MPC-assisted model*.

1.1 Our Results and Contributions

Our paper makes novel contributions at both *conceptual* and *technical* levels. From a technical perspective, prior to our work, we lacked techniques for characterizing the solution space of approximate incentive compatibility — in particular, classical tools like Myerson's Lemma [Mye81] breaks down when we allow ϵ slack in the incentive compatibility, and thus our classical insights often fail. One of our main technical contributions is to develop new techniques for mathematically reasoning about approximate incentive compatibility. On the conceptual front, while an elegant line of work has shown ways in which cryptography and game theory can help each other [HT04, KN08, ADGH06, OPRV09, AL11, ACH11, GKM+13, GKTZ15, GTZ15, Kat08, DR07, GLR10, CGL+18, WAS22, CCWS21, PS17, KMSW22, FW20, EFW22] (see Section 1.2 for more discussions), our work is of a different nature. Our results reveal exciting new connections between cryptography and mechanism design, motivated by a practical problem. The popularity of blockchains and decentralized applications poses many exciting new challenges for decentralized mechanism design, and cryptography-meets-game-theory is a natural and promising paradigm. We thus hope that our new conceptual contributions can provide fodder and inspire new works in this exciting and much explored space.

We give a summary of our main results below.

1.1.1 Characterizing Miner Revenue under Approximate Incentive Compatibility

We first focus on the plain model that was studied in earlier works [LSZ19, Yao, BEOS19, Rou20, Rou21, FMPS21, CS21]. Recall that assuming infinite block size, it is possible to achieve a dream TFM (e.g., the burning posted price auction), but the miner revenue has to be zero. We ask the following question: *suppose we are willing to relax the incentive compatibility notion and allow an*

ϵ additive slack, can we circumvent the zero miner revenue lower bound? If so, exactly how much miner revenue can we hope for?

More specifically, ϵ-incentive-compatibility (including ϵ-UIC, ϵ-MIC, and ϵ-SCP) requires that any deviation cannot increase the strategic individual or coalition's utility by more than ϵ. We show that under ϵ-incentive-compatibility, we can achieve linear (in the number of users) miner revenue assuming infinite block size. Moreover, we give matching upper- and lower-bounds that tightly characterize exactly how much miner revenue can be attained.

Infinite block size. Consider the simple posted price auction with reserve price $r \leq \frac{\epsilon}{c}$ where c is the maximum number of users controlled by the strategic coalition: all bids that bid at least r are confirmed. Each confirmed bid pays r. All payment goes to the miner. It is not hard to show that the above auction satisfies strict UIC, strict MIC (for an arbitrarily sized miner-coalition), and ϵ-SCP against c-sized coalitions. Further, the expected total miner revenue is $\Theta(n \cdot \frac{\epsilon}{c})$ when the users' true values are not too small.

Although the above posted price achieves linear in n revenue, the drawback is that the miner revenue is unscalable: even as the users' bids scale up (e.g., by some multiplicative factor), the miner revenue does not grow proportionally. We therefore ask if randomization can help achieve scalability in miner revenue. We show that indeed the following randomized TFM achieves scalability in miner revenue:

Proportional auction // Let r be a fixed reserve price.

- Every bid $b \geq r$ is confirmed with probability 1 and every candidate bid $b < r$ is confirmed with probability b/r. Each confirmed bid b pays $p = \min\{\frac{b}{2}, \frac{r}{2}\}$.
- For each confirmed bid, miner gets a pre-determined threshold $r' = \sqrt{\frac{2r\epsilon}{9c}}$ if $p \geq r'$.

For example, suppose all users' bids are sampled independently from some distribution $\mathcal{D}$, and let m be the median of the distribution such that $\Pr_{x\sim\mathcal{D}}[x \geq m] \geq 1/2$ (or any other constant). Then, if we set $r = m$, the expected miner revenue (taken over the randomness of users' bids as well as of the TFM itself) is $\Omega(n \cdot \min(m, \sqrt{\frac{m\epsilon}{c}}))$.

Combining the posted price auction and the proportional auction, we have the following theorem:

Theorem 1.1. *Consider the hybrid auction which, given some bid distribution $\mathcal{D}$ with median m, runs either the posted posted price auction with reserve price $r = \min(\frac{\epsilon}{c}, m)$ or the proportional auction with the reserve price $r = m$, depending on which one has higher expected revenue. The hybrid auction is strict UIC, strict MIC (for an arbitrarily sized miner coalition), and ϵ-SCP against any miner-user coalition with at most c users. Further, it achieves $\Omega\left(n \cdot (\min(\frac{\epsilon}{c} + \sqrt{\frac{m\epsilon}{c}}, m))\right)$ expected total miner revenue.*

Next, we prove a matching bound that shows the limitation on how much miner revenue can be attained under approximate incentive compatibility, as stated in the following theorem — this bound holds no matter whether the block size is finite or infinite.

Theorem 1.2 (Limit on miner revenue for infinite block size)**.** *For any possibly randomized TFM (in the plain model) that satisfies ϵ-UIC, ϵ-MIC, and ϵ-SCP for miner-user coalitions with 1 user, the expected total miner revenue over a random bid vector sampled from $\mathcal{D}^n$ must be upper bounded by*

$$\mathbf{E}_{\mathbf{b}\sim\mathcal{D}^n}[\mu(\mathbf{b})] \leq 6n \cdot (\epsilon + \sqrt{\epsilon} \cdot \mathbf{E}_{x\sim\mathcal{D}}[\sqrt{x}]),$$

where $\mu(\mathbf{b})$ denotes the total miner revenue under the bid vector $\mathbf{b}$, n is the number of users, $\mathcal{D}_i$ denotes the true value distribution of user $i \in [n]$.

Finite block size. Another natural question is: *can we circumvent the finite-block impossibility under approximate incentive compatibility?* Unfortunately, although it is indeed possible to overcome the finite-block impossibility with approximate incentive compatibility, we prove a new impossibility result that rules out the existence of "useful" mechanisms whose social welfare (i.e., the sum of everyone's utilities) scales up proportionally w.r.t. the bid distribution:

Theorem 1.3 (Scalability barrier for approximate incentive compatibility in the plain model)**.** *Fix any $\epsilon > 0$, and suppose that the block size is k. Any (possibly random) TFM in the plain model that simultaneously satisfies ϵ-UIC, ϵ-MIC, and ϵ-SCP (even when the miner colludes with at most one user) has at most $\widetilde{O}(k^3\epsilon)$ social welfare where k is the block size and $\widetilde{O}(\cdot)$ hides logarithmic factors.*

1.1.2 Can We Circumvent the Finite-Block Impossibility with Cryptography?

Due to the negative result of Theorem 1.3, we want to seek other avenues that allow us to circumvent the finite-block impossibility. Since cryptography is widely deployed in today's blockchains, it is natural to ask whether we can bring cryptography to the design of transaction fee mechanisms, to help us achieve what is otherwise impossible.

New model: MPC-assisted TFM. Consider a scenario henceforth called the MPC-assisted model, where a set of miners jointly run a multi-party computation (MPC) protocol to implement the TFM. One may think of the MPC protocol as providing the following ideal functionality $\mathcal{F}_{\text{TFM}}$:

- Each player (either user or miner) may act as any number of identities (including 0), and on behalf of each identity, submit a bid to $\mathcal{F}_{\text{TFM}}$.
- The ideal functionality $\mathcal{F}_{\text{TFM}}$ executes the prescribed *allocation rule* of the TFM to decide which transactions to include and confirm in the block; it executes the *payment rule* and *miner revenue rule* of the TFM to decide how much each confirmed bid pays and the total miner revenue. $\mathcal{F}_{\text{TFM}}$ then sends to all players the set of bids that are confirmed, what price each confirmed bid pays, and the total miner revenue.

We require that the total miner revenue does not exceed the total payment, and that the total miner revenue is split among the miners.

We assume that there is a separate process to decide the set of miners whose job is to jointly run the MPC protocol. For example, this decision can be made through either proof-of-work or proof-of-stake. In the former case, the total miner revenue is effectively split among the miners proportional to their mining power. In the latter case, the total miner revenue is effectively split among the miners proportional to their stake.

We assume that the majority of the miners are honest and that the MPC provides guaranteed output (i.e., the strategic miners cannot cause the MPC protocol to abort without producing outcome). Note that if we can indeed design an incentive compatible protocol in the MPC-assisted model, then, no miner would be incentivized to deviate from the honest protocol, and this reinforces the honest majority assumption. We discuss how to extend our results to the setting of *majority-miner coalitions* in Remark 1.5.

Intuitively, an MPC-assisted TFM restricts the strategy space for players in comparison with the plain model:

R1 A strategic individual or coalition must decide its strategy without having seen honest users' bids (*c.f.* in the plain model, a strategic individual or coalition can decide their strategy after seeing other players' bids).

R2 Once the set of bids are committed to, the allocation rule must be implemented honestly (*c.f.* in the plain model, the winning miner or block proposer can strategically choose which transactions to include in the block).

Exactly because of the MPC-assisted model imposes the above restrictions on the strategy space, we are hopeful that it may allow us to circumvent impossibilities. Before we explain our results, we first discuss how to define incentive compatibility in the MPC-assisted model.

Remark 1.4 (On the practicality of MPC)**.** We start by assuming generic MPC, since this is a good starting point as an initial feasibility exploration. All the impossibility results in our paper hold even with generic MPC. However, for all the MPC-assisted mechanisms we propose, although we initially describe the feasibility results using generic MPC for conceptual simplicity, it turns out that we actually do not need generic MPC to actually instantiate these mechanisms. We discuss how to efficiently instantiate our MPC-assisted mechanisms in Appendix E.

Remark 1.5 (Extending our results to majority-miner coalitions)**.** All the results in the paper actually hold even when a coalition may control the majority of miners. When the majority of the miners may be malicious, the MPC protocol cannot provide guaranteed output, it can only provide "security with abort". In other words, the ideal functionality that is realized by the MPC now provides the following backdoor: an adversary controlling the majority of miners can send $\perp$ to the ideal functionality, which causes the protocol to abort and not produce any output.

Threfore, if we assume that the coalition can control the majority of miners, essentially the strategy space includes one more move: the strategic coalition can cause the protocol to abort in which case no block is mined, and no on obtains any utility. Obviously, a rational coalition should never make such a move.

Ex post vs. Bayesian notions of incentive compatibility. In the plain model, because a strategic individual or coalition can decide their bids after seeing others' bids, prior works [Rou21, CS21] considered an *ex post* notion of incentive compatibility. In the new MPC-assisted model, since players must submit their bids to $\mathcal{F}_{\text{TFM}}$ without seeing others' bids, it also makes sense to consider a *Bayesian* notion of equilibrium.

Informally, we say that an MPC-assisted TFM satisfies *Bayesian Nash Equilibrium (BNE)* for a strategic coalition (or individual) $\mathcal{C}$, following the honest strategy allows $\mathcal{C}$ to maximize its expected gain, assuming that the bids of users not in $\mathcal{C}$ are drawn independently from some known distribution. If the coalition $\mathcal{C}$ consists of an individual user, we say that the scheme satisfies *Bayesian UIC*. When $\mathcal{C}$ consists of at most ρ fraction of the miners, we say that the scheme satisfies *Bayesian MIC* against a ρ-sized miner-coalition, Finally, when the coalition $\mathcal{C}$ consists of at most ρ fraction of miners as well as at least 1 and at most c users, we say that the scheme satisfies *Bayesian SCP* against a (ρ, c)-sized coalition.

Jumping ahead, for the MPC-assisted model, all our mechanism designs achieve incentive compatibility even in the *ex post* setting — in other words, the incentive compatibility guarantees hold even if $\mathcal{F}_{\text{TFM}}$ leaks other players' bids to the strategic players before they decide their own strategy. On the other hand, all of our impossibilities hold even for the Bayesian setting. This makes both our upper- and lower-bounds stronger.

MPC-assisted TFM under strict incentive compatibility. Unfortunately, as shown in Appendix D, the MPC-assisted model does not help us circumvent the zero miner revenue lower bound, even for Bayesian notions of equilibrium. Instead, the main question we care about here is *whether the MPC-assisted model allows us to circumvent the finite-block impossibility.* It turns out that the answer is not a simple binary one.

First, we show that absent user-user collusion, we can indeed circumvent the strong finite-block impossibility of Chung and Shi [CS21]. Specifically, we can indeed construct a TFM that simultaneously achieves UIC, MIC, and $(\rho, c = 1)$-SCP for any ρ. In particular, consider the following *posted price auction with random selection* — recall that to specify an MPC-assisted TFM, we only need to specify the allocation rule, the payment and miner revenue rules.

MPC-assisted, posted price auction with random selection

Let r be a fixed reserve price. Any bid that is at least r is considered as a candidate. Randomly choose up to block size k candidates to confirm. Any confirmed bid pays r. All payment is burnt and the miner revenue is 0.

Appendix E describes how to instantiate the above MPC-assisted mechanism efficiently without using generic MPC.

Theorem 1.6 (MPC-assisted, posted price auction with random selection)**.** *The above MPC-assisted, posted price auction with random selection satisfies UIC, MIC, and $(\rho, 1)$-SCP in the ex post setting for an arbitrary $\rho \in [0, 1]$.*

Since Theorem 1.6 holds even in the ex post setting, another interpretation is that the enforcement of the allocation rule (i.e., restriction R2, and not R1) is what allows us to circumvent the finite-block impossibility when $c = 1$.

The above posted price auction with random selection works for $c = 1$, i.e. no user-user collusion; however, it fails when the coalition may contain $c \geq 2$ users. Imagine that the number of users $n = k + 1$, and the coalition consists of two users and any fraction of miners. Now, suppose one of the colluding users has true value $v \gg r$, and the other has true value $v' = r$. In this case, the user with true value $v' = r$ should simply drop out and not submit a bid. This guarantees that the friend with large true value will be confirmed, and thus the coalition's joint utility increases.

It turns out that this is no accident. We prove that for $c \geq 2$, no MPC-assisted TFM can achieve UIC, MIC, and SCP for (ρ, c)-sized coalitions at the same time for any choice of ρ. Further, the impossibility holds even assuming Bayesian notions of incentive compatibility.

Theorem 1.7 (Finite-block impossibility in the MPC-assisted model for $c \geq 2$)**.** *Let $c \geq 2$ and let $\rho \in [0, 1]$. No (possibly randomized) MPC-assisted TFM with non-trivial utility can simultaneously achieve Bayesian UIC, Bayesian MIC, and Bayesian SCP for (ρ, c)-sized coalitions, assuming finite block size.*

Table 1: Mathematical landscape of TFM. Results in blue background are shown in this paper. ✗ means impossible and ✓ means possible. $\Theta(\cdot)$ means that we show matching upper and lower bounds — here m is a term that depends on the scale of the bid distribution, and we ignore terms related to c for simplicity. Unless otherwise noted, the impossibilities hold even for $c = 1$.

		plain model	MPC-assisted model
Infinite block	strict	0 miner rev [CS21]	0 miner rev
	approximate	$\Theta(n \cdot (\epsilon + \sqrt{m\epsilon}))$ miner rev	$\Theta(n \cdot (\epsilon + \sqrt{m\epsilon}))$ miner rev
Finite block	strict	✗ [CS21]	✓: $c = 1$, ✗: $c \geq 2$
	approximate	scalability ✗ (ignoring log terms)	scalability ✓

MPC-assisted TFM under approximate incentive compatibility. Recall that in the plain model, even with approximate incentive compatibility, we cannot have scalable TFMs whose social welfare scales w.r.t. the bid distribution (Theorem 1.3). We show that if we consider approximate incentive compatibility in the MPC-assisted model, we can overcome this scalability barrier. Specifically, we construct an MPC-assisted TFM called the "diluted posted price auction" that can achieve up to $\Theta(M \cdot k)$ social welfare when many people's bids are large enough, where M is an upper bound on users' bid.

MPC-assisted, diluted posted price auction

- Let r be a fixed reserve price, let M be the maximum possible value of the bid, and let k be the block size.
- Remove all bids that are less than r, and suppose that there are ℓ bids left — these bids form the candidate pool.
- Let $N = \max\{c \cdot \sqrt{\frac{kM}{2\epsilon}}, k\}$. If $\ell < N$, pad the candidate pool with fake 0 bids such that its size is N.
- Choose k bids at random from the candidate pool. All real bids chosen are confirmed and pay the reserve price r.
- The miner gets $\frac{2\epsilon}{c}$ for each confirmed bid.

Appendix E describes how to instantiate the above MPC-assisted mechanism efficiently without using generic MPC.

In the above mechanism, suppose we set the reserve price $r \leq M/2$, and further, imagine that everyone's true value is M, and they all bid their true value. Further, assume that there are many more users than the block size k. In this case, the block will be filled with k confirmed bids, and for each confirmed bid obtains utility $M/2$. Thus, we can achieve $\Theta(M \cdot k)$ social welfare.

Theorem 1.8 (MPC-assisted, diluted posted price auction)**.** *The above MPC-assisted, diluted posted price auction satisfies strict UIC, strict MIC, and ϵ-SCP for (ρ, c)-sized coalitions in the ex post setting, for any choice of ρ and c. Further, the mechanism is scalable, i.e., it can achieve $\Theta(M \cdot k)$ expected social welfare under some bid configurations.*

Summary of landscape. Summarizing our understanding so far, we present the mathematical landscape of TFM in Table 1. Our results show that cryptography can help us circumvent fundamental impossibilities of the plain model under finite block size. First, for strict incentive compatibility, cryptography allows us to overcome the finite-block impossibility for $c = 1$ (Theorem 1.6). Second, with approximate incentive compatibility, cryptography allows us to overcome the scalability barrier for finite block size in the plain model.

On the other hand, cryptography is also not a panacea. For example, it does not fundamentally help us improve miner revenue in the infinite block size setting.

1.2 Additional Related Work

We review some additional related works besides the most closely related works on transaction mechanism design [LSZ19, Yao, BEOS19, BCD+, Rou20, Rou21, FMPS21, CS21] mentioned earlier.

Earlier, an elegant line of work [HT04, KN08, ADGH06, OPRV09, AL11, ACH11, GKM+13, GKTZ15, GTZ15, Kat08, DR07, GLR10, CGL+18, WAS22, CCWS21, PS17, KMSW22, FW20, EFW22] revealed ways in which cryptography and game theory can help each other. Among them, some works [DR07] showed how to rely on cryptography to remove the trusted mediator assumption in certain game theoretic notions such as correlated equilibrium. Some [HT04, ADGH06, IML05, OPRV09, CGL+18, WAS22] showed that adopting game theoretic notions of fairness rather than the more stringent cryptographic notions of fairness can allow us to circumvent well-known lower bounds. Recently, Ferreira et al. [FW20] and Essaidi et al. [EFW22] showed that using cryptographic commitments can help us circumvent lower bounds pertaining to credible auctions. As Chung and Shi [CS21] explained, credible auction is of a different nature from transaction fee mechanism design. Transaction fee mechanism is a new type of decentralized mechanism design problem, and the new connections between cryptography and mechanism design revealed in our paper differ in nature from the settings in prior works.

2 Model and Definitions

Notation. We use bold letters to denote vectors. For a vector $\mathbf{b} = (b_1, \ldots, b_N)$, we use b_i to represent the i-th entry of vector $\mathbf{b}$. The notation $\mathbf{b}_{-i} = (b_1, b_2, \ldots, b_{i-1}, b_{i+1}, \ldots, b_N)$ represents all except the i-th entry. We often use $(\mathbf{b}_{-i}, b_i)$ and $\mathbf{b}$ interchangeably. Throughout the paper, we use n to denote the number of users, and N to denote the number of bids. N is equal to n if everyone behaves truthfully. However, strategic users may post zero or multiple bids — in this case N may not be equal to n. Given a distribution $\mathcal{D}$, we use the notation $\mathrm{Supp}(D)$ to denote its support. We use $\mathbb{R}^{\geq 0}$ to denote non-negative real numbers.

2.1 Transaction Fee Mechanism in the Plain Model

We first define transaction fee mechanism (TFM) in the plain model. Henceforth, we use $\mathcal{C}$ to denote a coalition of strategic players (or a strategic individual). In particular, $\mathcal{C}$ can be a user, the miner of the present block, or a coalition of the miner and one or more users.

Plain model. In the plain model, a transaction fee mechanism (TFM) describes the following game:

1. Users not in $\mathcal{C}$ submit their bids where each bid is represented by a single real value — let $\mathbf{b}_{-\mathcal{C}}$ denote the resulting bid vector.

2. The coalition $\mathcal{C}$ sees $\mathbf{b}_{-\mathcal{C}}$, and then users in $\mathcal{C}$ submit their bids.
3. The miner of the present block, possibly a member of $\mathcal{C}$, chooses up to k bids to include in the block, where k denotes the maximum block size.
4. Among the at most k bids included in the block, the trusted blockchain decides 1) which of them are confirmed, 2) how much each confirmed bid pays, and 3) how much revenue is paid to the miner.

Therefore, to specify a transaction fee mechanism (TFM) in the plain model, it suffices to specify the following rules which are *possibly randomized* functions:

- *Inclusion rule*: given a bid vector $\mathbf{b}$, the inclusion rule chooses up to k bids to include in the block;
- *Confirmation and payment rules*: Given the at most k bids included in the block, the confirmation rule decides which ones to confirm, and the payment rule decides how much each confirmed user pays.
- *Miner revenue rule*: Given the at most k bids included in the block, the miner revenue rule decides how much the miner earns.

In particular, the inclusion rule is implemented by the miner, and if the miner is strategic, it may not follow the prescribed inclusion rule but instead choose an arbitrary set of bids to include. By contrast, the confirmation, payment, and miner revenue rules are implemented by the blockchain, and honest implementation is guaranteed.

We assume that the (honest) TFM is *symmetric* in the following sense: if we apply any permutation π to an input bid vector $\mathbf{b} = (b_1, \ldots, b_N)$, it does not change the distribution of the *random variable* represented by the *set* $\{(b_i, x_i, p_i)\}_{i \in [N]}$ where x_i and p_i are random variables denoting the probability that bid i is confirmed, and its payment, respectively. An equivalent, more operational view of the above condition is the following. We may assume that the honest mechanism can always be equivalently described in the following manner: given a bid vector $\mathbf{b}$ where each bid may carry some extra information such as identity or timestamp, the honest mechanism always sorts the vector $\mathbf{b}$ by the bid amount first. During this step, if multiple bids have the same amount, then arbitrary tie-breaking rules may be applied, and the tie-breaking can depend on the extra information such as timestamp or identity. At this point, the inclusion rule and the confirmation rules should depend *only* on the amount of the bids and their relative position in the sorted bid vector. Note that our symmetry requirement is natural and quite general — it captures all the mechanisms we know so far [LSZ19, Yao, BEOS19, BCD+, Rou20, Rou21, FMPS21]. In particular, due to possible tie-breaking in the sorting step, our symmetry condition does *not* require two bids of the same amount to receive the same treatment, i.e., the distribution of their outcomes can be different.

Strategy space. A user's truthful behavior is submit a single bid representing its true value. However, strategic users may choose to submit zero to multiple bids, and the bids need not reflect their true value.

An honest miner does not submit any bids and honestly implements the prescribed inclusion rule. A strategic miner, on the other hand, may not honestly implement the prescribed inclusion rule — it can pick an arbitrary set of up to k bids of its choice to include. A strategic miner can also post fake bids. A coalition $\mathcal{C}$'s strategy space is defined in the most natural manner, i.e., it includes any strategic behavior of its members.

$\mathcal{F}_{\text{MPC}}$: **Ideal Functionality**

// The functionality is parametrized with the allocation, payment, and miner revenue rules.

1. Receive a single bid b_i from each identity. Let $\mathbf{b}$ be the resulting bid vector.
2. Run the allocation rule, the payment rule, and the miner revenue rule with the bid vector $\mathbf{b}$. The outputs include a bit vector indicating whether each bid in $\mathbf{b}$ is confirmed or not, a payment vector where all unconfirmed bids must pay 0, and the total miner revenue. Send the outputs to everyone.

Figure 1: Ideal functionality realized by the MPC protocol.

Notably, any strategic player in $\mathcal{C}$ can decide its actions *after* having observed the bids of the remaining users not in $\mathcal{C}$.

2.2 Transaction Fee Mechanism in the MPC-Assisted Model

Imagine that all miners jointly run an multi-party computation (MPC) protocol that implements the TFM. Figure 1 depicts the natural ideal functionality (denoted $\mathcal{F}_{\text{MPC}}$) realized by the MPC protocol. Further, the MPC protocol can achieve full security with guaranteed output as long as the majority of the miners are honest. Therefore, following the modular composition [Can00] paradigm in the standard cryptography literature, we can simply assume that a trusted party $\mathcal{F}_{\text{MPC}}$ exists — this is often referred to as the $\mathcal{F}_{\text{MPC}}$-hybrid model. We defer how to securely realize $\mathcal{F}_{\text{MPC}}$ to Appendix D.

MPC-assisted model. A transaction fee mechanism (TFM) in the MPC-assisted model describes the following game:

1. Every player (i.e., user or user) can take on *zero to multiple* identities, and every identity submits a bid represented by a single real value to $\mathcal{F}_{\text{MPC}}$ defined in Figure 1.
2. $\mathcal{F}_{\text{MPC}}$ decides which bids to confirm, how much each confirmed bid pays, and the total miner revenue. The total miner revenue is split among the miners.

Therefore, to specify a TFM in the MPC-assisted model, we need to specify the allocation rule, the payment rule, and the miner revenue rule — we assume that these rules are *possibly randomized*, polynomial-time algorithms, and the syntax of the rules are evident from $\mathcal{F}_{\text{MPC}}$ in Figure 1. In comparison with the plain model, here the *inclusion* rule and the *confirmation* rule are combined into a single *allocation* rule, since both inclusion and confirmation decisions are made by $\mathcal{F}_{\text{MPC}}$. Just like in the plain model, we assume that the (honest) TFM is symmetric.

Strategy space. A user's honest behavior is to take on a *single* identity, submit a single bid which reflects its true value. However, as mentioned above, any strategic user can take on zero or multiple identities, submit zero or multiple bids that need not be its true value.

An honest miner does not take on any identities or submit any bids. However, a strategic miner can take on one or more identities and submit fake bids. Unlike the plain model, here, a strategic miner can no longer choose which bids to include in the block — the allocation rule (i.e., the counterpart of the inclusion + confirmation rules of the plain model) is enforced by $\mathcal{F}_{\text{MPC}}$.

One technicality is whether the distribution of users' identities matter, and whether choosing identities strategically should be part of the strategy space. Jumping ahead, all of our mechanisms are proven to be incentive compatible even when the strategic individual or coalition can arbitrarily choose their identities as long as they cannot impersonate honest users' identities. On the other hand, all of our impossibility results hold even when the strategic individual or coalition is forced to choose their identities from some a-priori known distribution. This makes both our feasibility and infeasbility results stronger.

2.3 Defining Incentive Compatibility

Utility. Every user $i \in [n]$ has a true value $v_i \in \mathbb{R}^{\geq 0}$ if its transaction is confirmed. If user i's transaction is confirmed and the user pays p_i, then its utility is defined as $v_i - p_i$. A miner's utility is simply its revenue.

The utility of any strategic coalition $\mathcal{C}$ is the sum of the utilities of all members of $\mathcal{C}$. Considering the joint utility of the coalition is appropriate since we assume that the coalition has a *binding* mechanism (e.g., decentralized smart contracts) to split off their gains off the table.

Ex post incentive compatibility. We first define ex post incentive compatibility for both the plain model and the MPC-assisted model. Roughly speaking, ex post incentive compatibility requires that a strategic player or coalition's best response is always to behave honestly, even after observing the remaining users' bids. Similary, ex post ϵ-incentive compatibility requires that no strategy can increase a strategic player or coalition's expected utility by more than ϵ in comparison with the honest strategy, and this should hold even if the coalition can decide its strategy *after* having observed the remaining users' bids.

Below in our formal definitions, we define the *approximate* case that allows ϵ slack. When $\epsilon = 0$, we get *strict* incentive compatibiity — in this case, we can omit writing the ϵ.

Definition 2.1 (Ex post incentive compatibility)**.** We say that a mechanism satisfies *ex post* ϵ-incentive compatibility for a set of players $\mathcal{C}$ (possibly an individual), iff for any bid vector $\mathbf{b}_{-\mathcal{C}}$ posted by users not in $\mathcal{C}$, for any vector of true values $\mathbf{v}_{\mathcal{C}}$ of users in $\mathcal{C}$, no strategy can increase $\mathcal{C}$'s expected utility by more than ϵ in comparison with honest behavior. Specifically,

- *UIC.* We say that a TFM (in either the plain or MPC-assisted model) satisfies *ex post ϵ-user incentive compatibility (UIC)*, iff for any n, for any $i \in [n]$, for any bid vector $\mathbf{b}_{-i}$ of all users other than i, for any true value v_i of user i, no strategy can increase i's expected utility by more than ϵ in comparison with truthful bidding.

- *MIC.* In the plain model, we focus on the miner of the present block when defining miner incentive compatibility. We say a TFM in the plain model satisfies *ex post ϵ-miner incentive compatibility MIC*, iff for any bid vector $\mathbf{b}$, no strategy can increase the miner's expected utility by more than ϵ in comparison with honest behavior. Recall that that here, the miner's honest behavior is to honestly implement the inclusion rule and not inject any fake bids.

  In the MPC-assisted model, we want MIC to hold for any coalition controlling at most ρ fraction of the miners. Therefore, we say that an MPC-assisted TFM satisfies *ex post ϵ-MIC* against ρ-sized coalitions, iff for any coalition controlling at most ρ fraction of the miners, for any bid vector $\mathbf{b}$, no strategy can increase the miner's expected utility by more than ϵ in comparison with honest behavior. In the $\mathcal{F}_{\text{MPC}}$-hybrid world, the miner's honest behavior is simply not to take on any identities and inject any fake bids.

- *SCP.* In the plain model, we want side-contract-proofness to hold for any miner-user coalition that involves the miner of the present block, and up to c users. We say that a TFM in the plain model satisfies ex post ϵ-side-contract-proofness (SCP) for c-sized coalitions, iff for any miner-user coalition consisting of the miner and up to c users, for any bid vector $\mathbf{b}_{-\mathcal{C}}$ posted by users not in $\mathcal{C}$, no strategy can increase $\mathcal{C}$'s expected utility by more than ϵ in comparison with honest behavior.

  In the MPC-assisted model, we want SCP to hold for any miner-user coalition that involves up to ρ fraction of the miners and up to c users. We say that an MPC-assisted TFM satisfies *ex post ϵ-SCP* for (ρ, c)-sized coalitions, iff for any miner-user coalition[2] consisting of at most ρ fraction of the miners and up to c users, for any bid vector $\mathbf{b}_{-\mathcal{C}}$ posted by users not in $\mathcal{C}$, no strategy can increase the coalition's utility by more than ϵ in comparison with honest behavior.

Bayesian incentive compatibility. For the MPC-assisted model, it also makes sense to consider a Bayesian notion of incentive compatibility. In particular, the MPC-assisted model requires that the strategic player or coalition decides its strategy without having seen the remaining users' bids. We may assume that the strategic player or coalition has some a-prior belief of each honest user's true value distribution. We assume that all honest users' true values are independently and identically distributed (i.i.d.) and sampled from some distribution $\mathcal{D}$. In Bayesian incentive compatibility, we imagine that a strategic individual or coalition cares about maximizing its expected utility where the expectation is taken over not just the random coins of the mechanism, but also the remaining honest users' bids.

Henceforth, we denote the bid vector as $\mathbf{b}$. Since the strategic players can choose to inject fake bids or drop out, the length of $\mathbf{b}$ is not necessarily equal to the number of users. Given a set $\mathcal{C}$ of users, we use $\mathbf{b}_{-\mathcal{C}}$ to denote the bids from users outside the coalition and $\mathcal{D}_{-\mathcal{C}}$ to denote the joint distribution $\mathbf{b}_{-\mathcal{C}}$. That is, $\mathcal{D}_{-\mathcal{C}} = \mathcal{D}^h$, where h is the number of honest users outside the coalition. Similarly, for any fixed individual i, we use $\mathbf{b}_{-i}$ to denote the bids from the remaining users and $\mathcal{D}_{-i}$ to denote the joint distribution of $\mathbf{b}_{-i}$. Again, we define ϵ-incentive compatibility for the Bayesian setting below, where the corresponding strict incentive compatibility notions can be obtained by setting $\epsilon = 0$.

Definition 2.2 (Bayesian incentive compatibility)**.** We say that an MPC-assisted TFM satisfies Bayesian ϵ-incentive compatibility for a coalition or individual $\mathcal{C}$, iff for any $\mathbf{v}_{\mathcal{C}}$ denoting the true values of users in $\mathcal{C}$, sample $\mathbf{b}_{-\mathcal{C}} \sim \mathcal{D}_{-\mathcal{C}}$, then, no strategy can increase $\mathcal{C}$'s expected utility by more than ϵ in comparison with honest bevavior, where the expectation is taken over randomness of the honest users bids $\mathbf{b}_{-\mathcal{C}}$, as well as random coins consumed by the TFM. Specifically,

- *UIC.* We say that an MPC-assisted TFM satisfies Bayesian ϵ-UIC, iff for any n, for any user $i \in [n]$, for any true value $v_i \in \mathbb{R}^{\geq 0}$ of user i, for any strategic bid vector $\mathbf{b}_i$ from user i which could be empty or consist of multiple bids,

$$\mathop{\mathbf{E}}_{\mathbf{b}_{-i} \sim \mathcal{D}_{-i}} \left[\mathsf{util}^i(\mathbf{b}_{-i}, v_i) \right] \geq \mathop{\mathbf{E}}_{\mathbf{b}_{-i} \sim \mathcal{D}_{-i}} \left[\mathsf{util}^i(\mathbf{b}_{-i}, \mathbf{b}_i) \right] - \epsilon$$

  where $\mathsf{util}^i(\mathbf{b})$ denotes the expected utility (taken over the random coins of the TFM) of user i when the bid vector is $\mathbf{b}$.

[2] We require the miner-user coalition to consist of a non-zero fraction the miners and at least one user — otherwise the definition would degenerate to UIC or MIC.

- *MIC.* We say that an MPC-assisted TFM satisfies Bayesian ϵ-MIC for ρ-sized coalitions, iff for any miner coalition $\mathcal{C}$ controlling at most ρ fraction of the miners, for any strategic bid vector $\mathbf{b}'$ injected by the miner,

$$\mathop{\mathbf{E}}_{\mathbf{b}_{-\mathcal{C}} \sim \mathcal{D}_{-\mathcal{C}}} \left[\mathsf{util}^{\mathcal{C}}(\mathbf{b}_{-\mathcal{C}}) \right] \geq \mathop{\mathbf{E}}_{\mathbf{b}_{-\mathcal{C}} \sim \mathcal{D}_{-\mathcal{C}}} \left[\mathsf{util}^{\mathcal{C}}(\mathbf{b}_{-\mathcal{C}}, \mathbf{b}') \right] - \epsilon$$

  where $\mathsf{util}^{\mathcal{C}}(\mathbf{b})$ denotes the expected utility (taken over the random coins of the TFM) of the coalition $\mathcal{C}$ when the input bid vector is $\mathbf{b}$.

- *SCP.* We say that an MPC-assisted TFM satisfies Bayesian ϵ-SCP for (ρ, c)-sized coalitions, iff for any miner-user coalition consisting of at most ρ fraction of the miners and at most c users, for any true value vector $\mathbf{v}_{\mathcal{C}}$ of users in $\mathcal{C}$, for any strategic bid vector $\mathbf{b}_{\mathcal{C}}$ of the coalition (whose length may not be equal to the number of users in $\mathcal{C}$),

$$\mathop{\mathbf{E}}_{\mathbf{b}_{-\mathcal{C}} \sim \mathcal{D}_{-\mathcal{C}}} \left[\mathsf{util}^{\mathcal{C}}(\mathbf{b}_{-\mathcal{C}}, \mathbf{v}_{\mathcal{C}}) \right] \geq \mathop{\mathbf{E}}_{\mathbf{b}_{-\mathcal{C}} \sim \mathcal{D}^{-\mathcal{C}}} \left[\mathsf{util}^{\mathcal{C}}(\mathbf{b}_{-\mathcal{C}}, \mathbf{b}_{\mathcal{C}}) \right] - \epsilon$$

Note that the Bayesian notions of incentive compatibility do not make sense in the plain model, since in the plain model, the strategic individual or coalition can decide its move *after* having observed the remaining honest users' bids. This is why we adopt only the ex post notion in the plain model. Formally, it is easy to show that any mechanism that satisfies Bayesian incentive compatibility in the plain model also satisfies ex post incentive compatibility.

In the MPC-assisted model, both notions make sense, and the ex post notions are strictly stronger than the Bayesian counterparts. Jumping ahead, all of our impossibility results for the MPC-assisted model work even for the Bayesian notions, and all of our mechanism designs in the MPC-assisted model work even for the ex post notions. This makes both our lower- and upper-bounds stronger.

3 Approximate Incentive Compatibility for Infinite Block Size

In the plain model, no UIC and SCP mechanism (even for $c = 1$ and infinite block size) can achieve positive miner revenue [CS21]. In Appendix C.1, we show that the same zero miner revenue lower bound holds even in the MPC-assisted model. Therefore, we consider how to get meaningful miner revenue using the relaxed notion of approximate incentive compatibility. In this section, we give a tight characterization of approximate incentive compatibility for infinite block size. This tight characterization applies to both the MPC-assisted model and the plain model.

3.1 Bounds on Miner Revenue

We first prove a limit on miner revenue in the MPC-assisted model, which holds even for in the Bayesian setting. The same limit applies to the plain model for the ex post setting — to see this, observe that the strategy space is strictly larger in the plain model, and moreover, for the plain model, we only care about $\rho = 1$.

We now show an MPC-assisted mechanism simultaneously satisfies ϵ-UIC, ϵ-MIC and ϵ-SCP even for the Bayesian setting and even for $c = 1$ and an arbitray choice $\rho \in (0, 1]$, then the miner can gain at most $O(n \cdot (\epsilon + \cdot\sqrt{m^* \cdot \epsilon}))$-miner revenue, where n is the number of users, and m^* is a term that depends on the "scale" of the bid distribution.

To prove the limit on the miner revenue, we care only about the probability of each bid being confirmed, the expected payment of each bid, and the miner revenue. Therefore, we introduce the

following notations to denote the outputs of the allocation, payment, and miner revenue rules — we assume that each user's true value is drawn i.i.d. from some distribution $\mathcal{D}$ since we are considering the Bayesian setting:

- **Allocation rule**: given a bid vector $\mathbf{b} = (b_1, \ldots, b_N)$, the allocation rule outputs a vector $\mathbf{x}(\mathbf{b}) := (x_1, \ldots, x_N) \in [0,1]^N$, where each x_i denotes the probability of b_i being confirmed.
- **Payment rule**: given a bid vector $\mathbf{b} = (b_1, \ldots, b_N)$, the payment rule outputs a vector $\mathbf{p}(\mathbf{b}) := (p_1, \ldots, p_N) \in \mathbb{R}^N$, where each p_i denotes the expected payment of b_i.
- **Miner revenue rule**: given a bid vector $\mathbf{b} = (b_1, \ldots, b_N)$, the miner revenue rule outputs $\mu(\mathbf{b}) \in \mathbb{R}$, denoting the amount paid to the miner.

We also define $\mathcal{D}_{-i} := \mathcal{D}^{N-1}$, and for the i-th user, we define

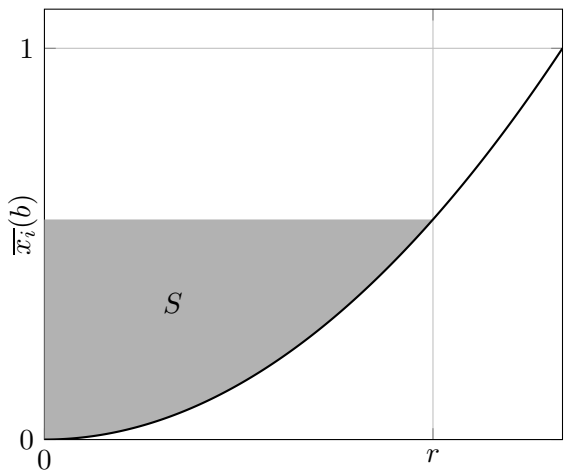
$$\overline{x_i}(\cdot) = \mathop{\mathbf{E}}_{\mathbf{b}_{-i} \sim \mathcal{D}_{-i}}[\mathbf{x}_i(\mathbf{b}_{-i}, \cdot)], \quad \overline{p_i}(\cdot) = \mathop{\mathbf{E}}_{\mathbf{b}_{-i} \sim \mathcal{D}_{-i}}[\mathbf{p}_i(\mathbf{b}_{-i}, \cdot)], \quad \overline{\mu_i}(\cdot) = \mathop{\mathbf{E}}_{\mathbf{b}_{-i}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

Henceforth, we often use $(\mathbf{x}, \mathbf{p}, \mu)$ to denote a TFM in the MPC-assisted model. The crux of our proof is to characterize how miner revenue changes when we lower one user's bid to 0 (Lemma 3.3). We then apply this argument n times, and lower each user's bid one by one to 0 to get the desired bound. To make the second step work, we need to use approximate MIC to remove a user's bid from consideration once we have lowered it to zero — this ensures that in any step of our inductive argument, the non-strategic users' bids are always i.i.d. sampled from $\mathcal{D}$.

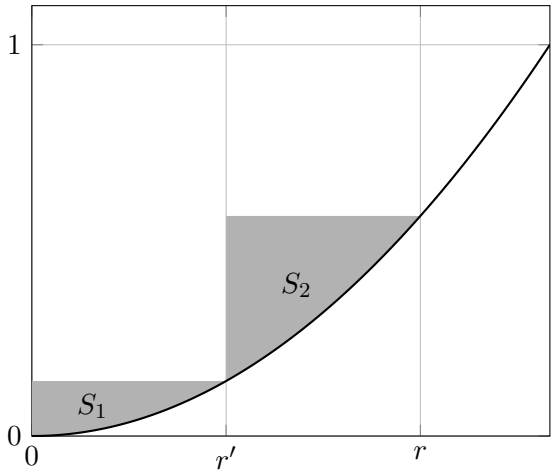
Warmup. To understand how much the miner revenue changes when one user lowers its bid to 0, we start from a simplified case where a TFM $(\mathbf{x}, \mathbf{p}, \mu)$ is Bayesian *strict*-UIC and Bayesian ϵ-SCP for $c = 1$ and some $\rho \in (0, 1]$. By Myerson's Lemma [Mye81], strict-UIC implies that, for any user i, the allocation rule $x_i(\cdot)$ must be non-decreasing. Moreover, the expected payment when bidding b is specified as

$$\overline{p_i}(b) = b \cdot \overline{x_i}(b) - \int_0^b \overline{x_i}(t)dt.$$

We care about how much the miner revenue can increase when user i bids r instead of 0. One trivial upper bound can be obtained as follows. Imagine that user i's true value is 0, but it bids r instead. In this case, the user's loss in utility (in comparison with truthful bidding) is represented by the area of the gray triangle S in Figure 2a. Due to ϵ-SCP, the miner revenue increase when user i bids r instead of 0 must be upper bounded by $S + \epsilon$. This bound, however, is not tight. To make it tighter, we consider bounding it in two steps by introducing a mid-point $r' \in (0, r)$. If user i's true value is 0, but it bids r' instead, its utility loss is the area S_1 of Figure 2b. By ϵ-SCP, we conclude that $\overline{\mu_i}(r') - \overline{\mu_i}(0) \geq S_1 + \epsilon$. Now, imagine user i's true value is r' but it bids r instead. Using a similar argument, we conclude that $\overline{\mu_i}(r) - \overline{\mu_i}(r') \geq S_2 + \epsilon$ (see Figure 2b). Summarizing the above, we have that $\overline{\mu_i}(r) - \overline{\mu_i}(0) \geq S_1 + S_2 + 2\epsilon$.



(a) When user i changes its bid from 0 to r, it loses utility S. Therefore, miner revenue changes by no more than $S + \epsilon$.



(b) When user i changes its bid from 0 to r', it loses utility S_1. Then when it changes its bid from r' to r, it loses utility S_2.

Figure 2: User's utility change

To get a tight bound, the key is how to choose the optimal number of steps L we use in the above argument. Taking more steps makes the total area of the gray triangles smaller; however, every step incurs an extra ϵ. Given the number of steps L, the sum of the L triangles is upper bounded by r/L, and since each step incurs an additive ϵ term, our goal is to minimize the expression $r/L + \epsilon L$. Picking $L = \sqrt{\frac{r}{\epsilon}}$ minimizes the expression and thus we have that $\overline{\mu_i}(r) - \overline{\mu_i}(0) \leq 2\sqrt{r\epsilon}$.

Full proof. The above warmup argument works for strict-UIC and ϵ-SCP. We want to prove a limitation on miner revenue for Bayesian ϵ-UIC and ϵ-SCP. The challenge is that for ϵ-UIC, Myerson's lemma no longer holds — in particular, the allocation rule may not even be monotone any more. The key idea our proof is to give a generalization of Myerson's lemma to account for the ϵ slack in incentive compatibility. We first prove a generalization of Myerson's *payment difference sandwich* for ϵ-UIC.

Lemma 3.1. *Given any (possibly randomized) MPC-assisted TFM that is Bayesian ϵ-UIC, it must be that for any user i, for any $y \leq z$,*

$$z \cdot [\overline{x_i}(z) - \overline{x_i}(y)] + \epsilon \geq \overline{p_i}(z) - \overline{p_i}(y) \geq y \cdot [\overline{x_i}(z) - \overline{x_i}(y)] - \epsilon. \tag{1}$$

Proof. The proof is similar to the proof of Myerson's Lemma. Note that user i's expected utility is $v \cdot \overline{x_i}(b) - \overline{p_i}(b)$ if its true value is v and its bid is b. By the definition of Bayesian ϵ-UIC, it must be that

$$z \cdot \overline{x_i}(z) - \overline{p_i}(z) + \epsilon \geq z \cdot \overline{x_i}(y) - \overline{p_i}(y).$$

Otherwise if user i's true value is z, bidding y can bring it strictly more than ϵ utility compared to bidding truthfully, which contradicts Bayesian ϵ-UIC. By the same reasoning, we have

$$y \cdot \overline{x_i}(y) - \overline{p_i}(y) + \epsilon \geq y \cdot \overline{x_i}(z) - \overline{p_i}(z).$$

The lemma thus follows by combining these two inequalities.

□

Based on this payment difference sandwich, we have the following result about the expected miner's revenue for approximate incentive compatibility.

Lemma 3.2. *Fix any $\rho \in (0, 1]$. For any (possibly randomized) MPC-assisted TFM that is Bayesian ϵ_u-UIC and Bayesian ϵ_s-SCP against a $(\rho, 1)$-sized coalition, it must be that for any user i, for any $y \le z$,*

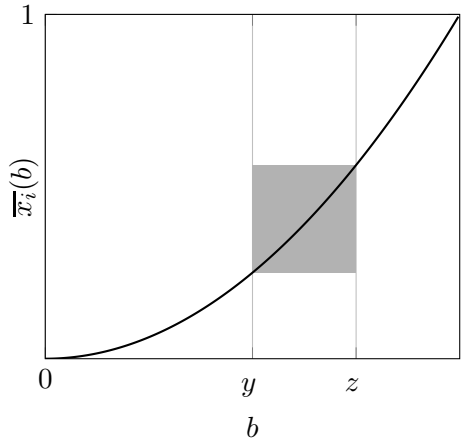
$$\overline{\mu_i}(z) - \overline{\mu_i}(y) \le \frac{1}{\rho}(\epsilon_u + \epsilon_s + S(y, z)), \tag{2}$$

where $S(y, z) = (z - y)[\overline{x_i}(z) - \overline{x_i}(y)]$.

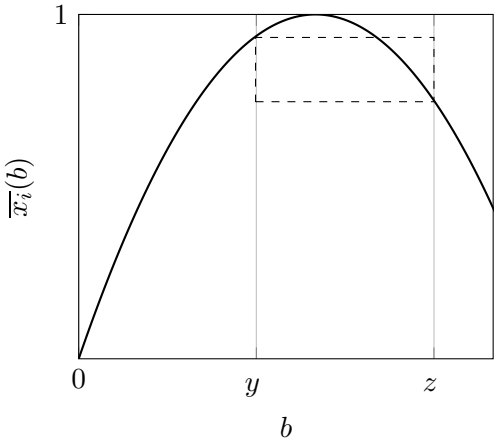
Proof. The utility of user i is $v \cdot \overline{x_i}(b) - \overline{p_i}(b)$ if its true value is v and it bids b. Imagine that the user i's true value is y. If user i overbids $z > y$ instead of its true value y, then its expected utility decreases by

$$\begin{aligned}
\Delta &= y \cdot \overline{x_i}(y) - \overline{p_i}(y) - [y \cdot \overline{x_i}(z) - \overline{p_i}(z)] \\
&= -y \cdot [\overline{x_i}(z) - \overline{x_i}(y)] + (\overline{p_i}(z) - \overline{p_i}(y)) \\
&\le -y \cdot [\overline{x_i}(z) - \overline{x_i}(y)] + z \cdot [\overline{x_i}(z) - \overline{x_i}(y)] + \epsilon_u \qquad \text{By Bayesian } \epsilon_u\text{-UIC and (1)} \\
&= (z - y) \cdot [\overline{x_i}(z) - \overline{x_i}(y)] + \epsilon_u = S(y, z) + \epsilon_u.
\end{aligned}$$

A graphical description of $S(y, z)$ is shown in Figure 3 — note that $S(y, z)$ can be *negative* since the allocation rule $\overline{x_i}(\cdot)$ may not be monotone under approximate UIC.



(a) An illustrative example of $S(y, z)$ in increasing function. The size of the gray area in the figure is exactly $S(y, z)$.



(b) When the function decreases, $S(y, z)$ can be negative. $S(y, z)$ is the negative of the dashed rectangle area.

Figure 3: User's utility change

By Bayesian ϵ_s-SCP, it must be that $\rho\overline{\mu_i}(z) - \rho\overline{\mu_i}(y) \le \Delta + \epsilon_s$; otherwise, a strategic player controlling ρ fraction of the miners can collude with user i, and ask user i to bid z instead of its true value y. This increases the coalition's utility by strictly more than ϵ_s compared to the honest strategy, which contradicts Bayesian ϵ_s-SCP. □

Lemma 3.3. *Fix any $\rho \in (0, 1]$. For any (possibly randomized) MPC-assisted TFM that is Bayesian ϵ_u-UIC and Bayesian ϵ_s-SCP against a $(\rho, 1)$-sized coalition, for any user i, for any value r, it must be that*

$$\overline{\mu_i}(r) - \overline{\mu_i}(0) \le \begin{cases} \frac{2}{\rho}(\epsilon_s + \epsilon_u), & \text{if } r \le \epsilon_s + \epsilon_u \\ \frac{2}{\rho}(\sqrt{r(\epsilon_s + \epsilon_u)}), & \text{if } r > \epsilon_s + \epsilon_u. \end{cases} \tag{3}$$

Proof. Let $\epsilon' = \epsilon_s + \epsilon_u$. To prove this Lemma, we consider the following two cases.

Case 1: If $r \le \epsilon'$. In this case, by Lemma 3.2, we have that

$$\overline{\mu_i}(r) - \overline{\mu_i}(0) \le \frac{1}{\rho}\left(\epsilon_u + \epsilon_s + S(0,r)\right) \le \frac{1}{\rho}\left(\epsilon_u + \epsilon_s + r\right) \le \frac{2\epsilon'}{\rho}.$$

Case 2: If $r > \epsilon'$. We choose a sequence of points that partitions the interval $[0, r]$ as follows. Let $L = \lfloor \sqrt{\frac{r}{\epsilon'}} \rfloor$. Set $r_0 = 0$ and $r_{L+1} = r$. For $l = 1, \ldots, L$, we set $r_l = l \cdot \sqrt{r\epsilon'}$. Each segment except the last one is of length $\sqrt{r\epsilon'}$, while the last one has length no more than $\sqrt{r\epsilon'}$.

Now we proceed to bound $\overline{\mu_i}(r) - \overline{\mu_i}(0)$. Note that

$$\begin{aligned}
\overline{\mu_i}(r) - \overline{\mu_i}(0) &= \sum_{l=0}^{L} [\overline{\mu_i}(r_{l+1}) - \overline{\mu_i}(r_l)] \\
&\le \sum_{l=0}^{L} \frac{1}{\rho}[\epsilon' + S(r_l, r_{l+1})] && \text{By Lemma 3.2} \\
&= \frac{L\epsilon'}{\rho} + \frac{1}{\rho}\sum_{l=0}^{L} (r_{l+1} - r_l) \cdot [\overline{x_i}(r_{l+1}) - \overline{x_i}(r_l)] \\
&\le \frac{L\epsilon'}{\rho} + \frac{1}{\rho}\sqrt{r\epsilon'}\sum_{l=0}^{L} [\overline{x_i}(r_{l+1}) - \overline{x_i}(r_l)] && \text{By the choice of } r_l \\
&\le \frac{L\epsilon'}{\rho} + \frac{1}{\rho}\sqrt{r\epsilon'} && \text{By } \overline{x_i}(r) \le 1
\end{aligned}$$

Since $L = \lfloor \sqrt{\frac{r}{\epsilon'}} \rfloor \le \sqrt{\frac{r}{\epsilon'}}$, we have that

$$\overline{\mu_i}(r) - \overline{\mu_i}(0) \le \frac{2\sqrt{r\epsilon'}}{\rho}.$$

□

Now, we want to bound the miner revenue by lowering each user's bid to 0 one by one, and apply Lemma 3.3 in each step. To make this argument work, one key insight is to rely on approximate MIC to remove a user's bid from consideration after lowering it to zero — see Equation (5) in the proof of Theorem 3.4 below. This ensures that in any step of the induction, any honest user's bid is sampled from $\mathcal{D}$.

Theorem 3.4 (Limit on miner revenue for approximate incentive compatibility)**.** *Suppose that there are n users, whose true values are drawn i.i.d. from some distribution $\mathcal{D}$. Given any (possibly randomized) MPC-assisted TFM that is Bayesian ϵ_u-UIC, Bayesian ϵ_m-MIC against a ρ-sized miner coalition and Bayesian ϵ_s-SCP against a $(\rho, 1)$-sized coalition, it must be that*

$$\mathop{\mathbf{E}}_{\mathbf{b}\sim\mathcal{D}^n}[\mu(\mathbf{b})] \le \frac{2n}{\rho}\left(\epsilon + C_{\mathcal{D}}\sqrt{\epsilon}\right), \tag{4}$$

where $\epsilon = \epsilon_s + \epsilon_u + \epsilon_m$, and $C_{\mathcal{D}} = \mathbf{E}_{X\sim\mathcal{D}}[\sqrt{X}]$ is a term that depends on the "scale" of the distribution $\mathcal{D}$.

Proof. Since the TFM is Bayesian ϵ_m-MIC, it must be that for any ℓ,

$$\mathop{\mathbf{E}}_{\mathbf{b}\sim\mathcal{D}^\ell}[\rho\mu(\mathbf{b},0)] \le \mathop{\mathbf{E}}_{\mathbf{b}\sim\mathcal{D}^\ell}[\rho\mu(\mathbf{b})] + \epsilon_m. \tag{5}$$

Otherwise, the strategic miner can inject a bid 0 and increase its miner revenue by strictly more than ϵ_m, while it does not need pay anything for injecting this 0-bid. This violates Bayesian ϵ_m-MIC.

Let $f(\cdot)$ be the p.d.f. of distribution $\mathcal{D}$. By the law of total expectation,

$$\mathop{\mathbf{E}}_{\mathbf{b}\sim\mathcal{D}^n}[\mu(\mathbf{b})] = \int_0^\infty \mathop{\mathbf{E}}_{\mathbf{b}'\sim\mathcal{D}^{n-1}}[\mu(\mathbf{b}',r)]f(r)dr.$$

Let $\epsilon' = \epsilon_s + \epsilon_u$. Since the mechanism is Bayesian ϵ_u-UIC and Bayesian ϵ_s-SCP against $(\rho,1)$-sized coalition, by Lemma 3.3, it must be that

$$\int_0^{\epsilon'} \mathop{\mathbf{E}}_{\mathbf{b}'\sim\mathcal{D}^{n-1}}[\mu(\mathbf{b}',r)]f(r)dr \le \int_0^{\epsilon'} \left[\mathop{\mathbf{E}}_{\mathbf{b}'\sim\mathcal{D}^{n-1}}[\mu(\mathbf{b}',0)] + \frac{2\epsilon'}{\rho}\right] f(r)dr;$$
$$\int_{\epsilon'}^\infty \mathop{\mathbf{E}}_{\mathbf{b}'\sim\mathcal{D}^{n-1}}[\mu(\mathbf{b}',r)]f(r)dr \le \int_{\epsilon'}^\infty \left[\mathop{\mathbf{E}}_{\mathbf{b}'\sim\mathcal{D}^{n-1}}[\mu(\mathbf{b}',0)] + \frac{2\sqrt{r\epsilon'}}{\rho}\right] f(r)dr.$$

Summing up the two inequalities above, we can bound the expected miner revenue with

$$\begin{aligned}
&\mathop{\mathbf{E}}_{\mathbf{b}\sim\mathcal{D}^n}[\mu(\mathbf{b})] \\
&= \int_0^{\epsilon'} \mathop{\mathbf{E}}_{\mathbf{b}'\sim\mathcal{D}^{n-1}}[\mu(\mathbf{b}',r)]f(r)dr + \int_{\epsilon'}^\infty \mathop{\mathbf{E}}_{\mathbf{b}'\sim\mathcal{D}^{n-1}}[\mu(\mathbf{b}',r)]f(r)dr \\
&\le \int_0^{\epsilon'} \left[\mathop{\mathbf{E}}_{\mathbf{b}'\sim\mathcal{D}^{n-1}}[\mu(\mathbf{b}',0)] + \frac{2\epsilon'}{\rho}\right] f(r)dr + \int_{\epsilon'}^\infty \left[\mathop{\mathbf{E}}_{\mathbf{b}'\sim\mathcal{D}^{n-1}}[\mu(\mathbf{b}',0)] + \frac{2\sqrt{r\epsilon'}}{\rho}\right] f(r)dr \\
&\le \mathop{\mathbf{E}}_{\mathbf{b}'\sim\mathcal{D}^{n-1}}[\mu(\mathbf{b}',0)] + \frac{2\epsilon'}{\rho}\int_0^{\epsilon'} f(r)dr + \frac{2\sqrt{\epsilon'}}{\rho}\int_{\epsilon'}^\infty \sqrt{r}f(r)dr
\end{aligned}$$

By (5), we have that $\mathop{\mathbf{E}}_{\mathbf{b}'\sim\mathcal{D}^{n-1}}[\mu(\mathbf{b}',0)] \le \mathop{\mathbf{E}}_{\mathbf{b}'\sim\mathcal{D}^{n-1}}[\mu(\mathbf{b}')] + \frac{\epsilon_m}{\rho}$. Therefore,

$$\begin{aligned}
&\mathop{\mathbf{E}}_{\mathbf{b}\sim\mathcal{D}^n}[\mu(\mathbf{b})] \\
&\le \mathop{\mathbf{E}}_{\mathbf{b}'\sim\mathcal{D}^{n-1}}[\mu(\mathbf{b}',0)] + \frac{2\epsilon'}{\rho}\int_0^{\epsilon'} f(r)dr + \frac{2\sqrt{\epsilon'}}{\rho}\int_{\epsilon'}^\infty \sqrt{r}f(r)dr \\
&\le \mathop{\mathbf{E}}_{\mathbf{b}'\sim\mathcal{D}^{n-1}}[\mu(\mathbf{b}')] + \frac{\epsilon_m}{\rho} + \frac{2\epsilon'}{\rho} + \frac{2\sqrt{\epsilon'}}{\rho}\mathbf{E}_{X\sim\mathcal{D}}[\sqrt{X}] \\
&\le \mathop{\mathbf{E}}_{\mathbf{b}'\sim\mathcal{D}^{n-1}}[\mu(\mathbf{b}')] + \frac{2\epsilon}{\rho} + \frac{2C_{\mathcal{D}}\sqrt{\epsilon}}{\rho},
\end{aligned}$$

where the last step comes from the fact that $\epsilon = \epsilon_s + \epsilon_u + \epsilon_m$. The theorem follows by induction on n, where in each induction step we repeat the argument above. □

It is easy to see that the same miner revenue limit of Theorem 3.4 also holds in the plain model, as stated in the following corollary.

Corollary 3.5. *Suppose that there are n users, whose true values are drawn i.i.d. from some distribution $\mathcal{D}$. Given any (possibly randomized) TFM in the plain model that is ϵ_u-UIC, ϵ_m-MIC, and ϵ_s-SCP even for $c = 1$, it must be that*

$$\mathop{\mathbf{E}}_{\mathbf{b}\sim\mathcal{D}^n}[\mu(\mathbf{b})] \le 2n\left(\epsilon + C_{\mathcal{D}}\sqrt{\epsilon}\right), \tag{6}$$

where $\epsilon = \epsilon_s+\epsilon_u+\epsilon_m$, and $C_{\mathcal{D}} = \mathbf{E}_{X\sim\mathcal{D}}[\sqrt{X}]$ is a term that depends on the "scale" of the distribution $\mathcal{D}$.

Proof. Follows directly from Theorem 3.4 which holds in particular for $\rho = 1$, and the fact that the strategy space in the plain model is strictly larger than in the MPC-assisted model. □

3.2 Achieving Optimal Revenue: Proportional Auction

We now show that the limit on miner revenue in Theorem 3.4 is asymptotically tight, i.e., we can indeed design a TFM, even in the plain model, whose miner revenue asymptotically matches Equation (4) for some natural bid distribution.

Proportional Auction (plain model)

Parameters: the slack ϵ, the reserved price r where $r \ge 2\epsilon$.

Input: a bid vector $\mathbf{b} = (b_1, \ldots, b_N)$.

Mechanism:

- *Inclusion rule.* Include all bids in $\mathbf{b}$.
- *Confirmation rule.* For each bid b, if $b < r$, it is confirmed with the probability b/r; otherwise, if $b \ge r$, it is confirmed with probability 1.
- *Payment rule.* For each confirmed bid b, if $b < r$, it pays $b/2$; otherwise, it pays $r/2$.
- *Miner revenue rule.* For each confirmed bid b, if $b \ge \sqrt{2r\epsilon}$[a], then miner is paid $\frac{\sqrt{2r\epsilon}}{2}$.

[a] This guarantees that the miner revenue does not exceed the total payment.

The above mechanism is called the proportional mechanism since the user's confirmation probability is proportional to the bid in the region $[0, r]$, and any bid that is at least r is confirmed with probability 1.

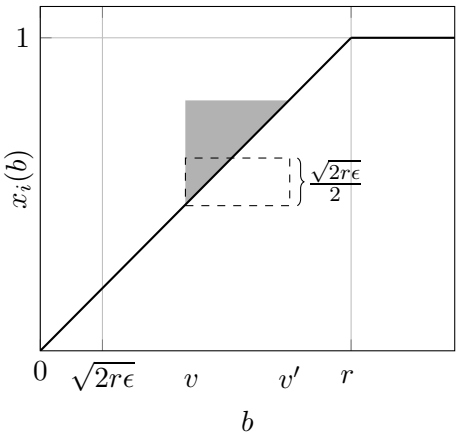
Theorem 3.6. *The above proportional auction in the plain model is UIC, MIC and $\frac{5}{4}c\epsilon$-SCP against c-sized coalitions for arbitrary $c \ge 1$.*

Proof intuition. We provide the proof intuition and defer the full proof to Appendix A. First, UIC and MIC are easy to prove. Observe that the allocation rule (i.e., the union of the inclusion and confirmation rules) is monotone, and by design, the payment rule is the unique one that satisfies Myerson's Lemma. Therefore, the mechanism satisfies UIC. It is easy to see that injecting a bid does not help the miner, since each bid's contribution to the miner revenue is independent and limited by the payment amount.

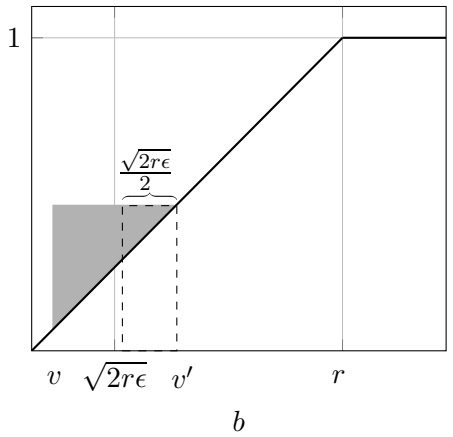
Proving that the mechanism satisfies $\frac{5}{4}c\epsilon$-SCP is more technical. Here we give an illustrative explanation to show that the joint utility of each user and the miner can increase by at most $\frac{5}{4}\epsilon$. Since underbidding does not increase the user's utility or the miner's revenue, we focus on overbidding. Note that overbidding does not increase the joint utility for a user whose true value is $v \ge r$. Therefore, we focus in the case where the colluding user has true value $v < r$ and overbids.

If $v \geq \sqrt{2r\epsilon}$, the user's utility loss when overbidding to v' is represented by the gray triangle in Figure 4a. Meanwhile, the miner's expected revenue increases by $\frac{\sqrt{2r\epsilon}}{2}(\frac{v'}{r} - \frac{v}{r})$, which is the area of the dashed rectangle in Figure 4a. Therefore, when the user overbids by $v' - v = \frac{\sqrt{2r\epsilon}}{2}$, the coalition's utility increase is maximized and equals to $\frac{\epsilon}{4}$.

If $v < \sqrt{2r\epsilon}$ and the colluding user overbids to $v' \geq \sqrt{2r\epsilon}$, then the user's utility loss when overbidding to v' is represented by the area of the gray triangle in Figure 4b. The miner's revenue now increases by $\frac{v'}{r} \cdot \frac{\sqrt{2r\epsilon}}{2}$, because the user's utility would be 0 if the user behaves honestly. The increase in the miner's revenue is represented by the dashed rectangle in Figure 4b. The increase in the joint utility of the coalition is maximized when v is arbitrarily close to $\sqrt{2r\epsilon}$ and the user overbids by $v' - v = \frac{\sqrt{2r\epsilon}}{2}$. In this case, the joint utility of the coalition increases by $\frac{5}{4}\epsilon$.



(a) An illustrative example of the coalition's joint utility change when the user's true value $v \geq \sqrt{2r\epsilon}$.



(b) An illustrative example of the coalition's joint utility change when the user's true value $v < \sqrt{2r\epsilon}$.

Figure 4: Coalition's joint utility change when the miner colluding with one user

4 Characterization of Finite Block Size in the Plain Model

In real-world blockchains, we do not have an infinite block size. Chung and Shi [CS21] showed that no non-trivial plain-model TFM can achieve strict UIC and strict SCP (even when $c = 1$) for finite block size. In this section, we show that although approximate incentive compatibility can help us overcome this impossibility, nonetheless we cannot get useful mechanisms whose social welfare scales with the bid distribution (ignoring logarithmic terms).

Theorem 4.1. *Suppose the block size is upper bounded by k. Fix any $\epsilon > 0$. Given any TFM in the plain model that satisfies ϵ-UIC, ϵ-MIC and ϵ-SCP when the miner can collude with at most $c = 1$ user, and given any bid vector $\mathbf{b}$, let $M = \max(\mathbf{b})$ be the maximum bid of any user, it must be that*

- *the miner's expected revenue is upper bounded by $12k^2\epsilon \log\left(\frac{M}{\epsilon} + 1\right) + 2k\epsilon$;*
- *every user's expected utility is upper bounded by $12k^2\epsilon \log\left(\frac{M}{\epsilon} + 1\right) + (2k + 1)\epsilon$ conditioned on the bid being included in the block, and assuming the bid reflects its true value;*
- *the expected social welfare is upper bounded by $O\left(k^3\epsilon \log\left(\frac{M}{\epsilon} + 1\right) + k^2\epsilon\right)$.*

A direct corollary of Theorem 4.1 is that there is no non-trivial mechanism that satisfies approximate incentive compatibility if the user's true value is unbounded. This implies that there is no universal mechanism that works for all bid distributions. Formally,

Corollary 4.2. *Suppose the block size is upper bounded by k. Fix any $\epsilon > 0$. If users' true values are unbounded, then no (possibly randomized) non-trivial TFM in the plain model can simultaneously satisfy ϵ-UIC and ϵ-SCP, even if the miner colludes with only one user.*

Proof. For the sake of contradiction, assume that there exists an $\epsilon > 0$, such that there exists a non-trivial TFM satisfying ϵ-UIC and ϵ-SCP. Recall that $x_i(\mathbf{b})$ denotes the probability of user i's bid being confirmed given that the world consists of the bid vector $\mathbf{b}$ (assuming the mechanism is honestly implemented). We define $\widetilde{x}_i(\mathbf{b}')$ to be the probability of user i's bid being confirmed conditioned on its bid being included in the block configuration $\mathbf{b}'$. According to the assumption that the mechanism is non-trivial, there must exist an $i \in [k]$ and a block configuration $\mathbf{b}' = (b^*, \mathbf{b}_{-i})$ such that b^* has a positive probability $\widetilde{x}_i(\mathbf{b}')$ of being confirmed.

Now imagine the world consists of the bid vector $\mathbf{b}$ where

$$\mathbf{b} = (b_1, b_2, \ldots, b_{k-1}, \underbrace{M, M, \ldots, M}_{T}),$$

where $T \geq \frac{2k}{\widetilde{x}_i(\mathbf{b}')}$ and M is some large number (larger than $\max\{b_1, \ldots, b_k\}$) that we will specify later.

Since the block size is bounded by k, there must exist a user j whose true value is M yet its probability of being confirmed is no more than $\frac{k}{T} \leq \frac{1}{2}\widetilde{x}_i(\mathbf{b}')$ by our choice of T. Therefore, user j's utility (assuming the mechanism is honestly implemented) is at most $M \cdot \frac{1}{2}\widetilde{x}_i(\mathbf{b}')$. Now consider the coalition of the miner and user j. By Theorem 4.1, their joint utility when behaving honestly is at most

$$M \cdot \frac{1}{2}\widetilde{x}_i(\mathbf{b}') + 12k^2\epsilon \log\left(\frac{M}{\epsilon} + 1\right) + 2k\epsilon.$$

However, the miner can ask user j to bid b^* instead of its true value M and include $(b_1, \ldots, b_{k-1}, b^*)$ into the block, where the bid b^* comes from user j. Since the payment cannot exceed the bid, now the utility of user j is at least

$$M \cdot \widetilde{x}_i(\mathbf{b}') - b^*.$$

As long as M is large enough such that

$$M \cdot \widetilde{x}_i(\mathbf{b}') - b^* \geq M \cdot \frac{1}{2}\widetilde{x}_i(\mathbf{b}') + 12k^2\epsilon \log\left(\frac{M}{\epsilon} + 1\right) + 2k\epsilon + \epsilon,$$

the coalition gains ϵ more joint utility comparing to honest strategy. This contradicts ϵ-SCP. Note that since user's true value can be unbounded, such M must exist. Therefore, there does not exist a non-trivial mechanism that satisfies ϵ-UIC and ϵ-SCP simultaneously. □

The rest of Section 4 is dedicated to proving Theorem 4.1.

4.1 Proof Roadmap

We first explain the blueprint. To prove that the total social welfare is small, we first show that the miner revenue must be $\widetilde{O}(k^2\epsilon)$ for any bid configuration. If we can show this, then given that the block size is finite, we can show that every user i's utility conditioned on being included is small, which then allows us to bound the total social welfare. Suppose this is not the case,

i.e., suppose that under some bid configuration $\mathbf{b} := (b_1, \ldots, b_N)$, there is a user i with expected utility (conditioned on being included) significantly larger than the maximum possible expected miner revenue (which is upper bounded by $\widetilde{O}(k^2\epsilon)$). Then, imagine a world consisting of $\mathbf{b}$ and additionally (infinitely) many users whose true value is the same as b_i. In this case, there must be one such user j whose expected utility is almost 0. Thus, if j is the miner's colluding friend, the miner would be willing to sacrifice all of its revenue, pretend that the world consists of $\mathbf{b}$ where the i-th coordinate is replaced with j's bid, and run the honest mechanism subject to j being included. In this case, the coalition can increase its expected joint utility since user j would be doing much better than the honest case.

The crux of our proof, therefore, is to show that the expected miner revenue must be bounded for any bid vector. To show this, we take two main steps. First, we show that if the world consists of only bids of value M, the expected miner revenue must be small (see Lemma 4.5). Using the above as base case, we then go through an inductive argument to show that in fact, for any bid vector where users do not necessarily bid M, the miner revenue must be small too (see Lemma 4.6). Note that showing the first step itself relies on another inductive argument that inducts on the length of the bid vector.

4.2 Detailed Proof

4.2.1 Individual User's Influence on Miner Revenue is Bounded

Before proving Theorem 4.1, we introduce some useful lemmas. The following lemma states that if, given some bid configuration, a user's expected utility is not too large, then, the miner's expected revenue should not drop too much when we lower that user's bid to 0.

Lemma 4.3. *Given any (possibly randomized) TFM in the plain model that satisfies ϵ-UIC, ϵ-MIC and ϵ-SCP against 1-sized coalition, for any $\mathbf{b}_{-i}$ and v, we have the following where $\mathsf{util}^i(\mathbf{b})$ denotes user i's expected utility and $\mu(\mathbf{b})$ is the expected miner revenue when the bid vector is $\mathbf{b}$:*

$$\mu(\mathbf{b}_{-i}, v) - \mu(\mathbf{b}_{-i}, 0) \leq \begin{cases} 4\epsilon, & v \leq 2\epsilon \\ \mathsf{util}^i(\mathbf{b}_{-i}, v) + 3\epsilon \log \frac{v}{\epsilon} + 4\epsilon, & v > 2\epsilon. \end{cases}$$

Proof. Henceforth, we use $\mathbf{x}(\mathbf{b})$ to denote the vector of probabilities that each bid in $\mathbf{b}$ is included and confirmed, and let $\mathbf{p}(\mathbf{b})$ denote the vector of expected payments for every user when the bid vector is $\mathbf{b}$.

First, observe that Lemma 3.1 and Equation (7) still hold in the plain model where the terms $\overline{x}_i(\cdot)$, $\overline{p}_i(\cdot)$, and $\overline{\mu}(\cdot)$ are now replaced with $x_i(\mathbf{b}_{-i}, \cdot)$ $p_i(\mathbf{b}_{-i}, \cdot)$, and $\mu(\mathbf{b}_i, \cdot)$ respectively, i.e., we now fix an arbitrary fixed $\mathbf{b}_{-i}$ rather than taking expectation over the random choice $\mathbf{b}_{-i}$.

Specifically, Lemma 3.1 implies that for any $\mathbf{b}_{-i}$, for any $b \leq b'$,

$$b' \cdot [x_i(\mathbf{b}_{-i}, b') - x_i(\mathbf{b}_{-i}, b)] + \epsilon \geq p_i(\mathbf{b}_{-i}, b') - p_i(\mathbf{b}_{-i}, b) \geq b \cdot [x_i(\mathbf{b}_{-i}, b') - x_i(\mathbf{b}_{-i}, b)] - \epsilon. \quad (7)$$

Lemma 3.2 implies that for any $\mathbf{b}_{-i}$, for any $b \leq b'$,

$$\mu(\mathbf{b}_{-i}, b') - \mu(\mathbf{b}_{-i}, b) \leq 2\epsilon + (b' - b) \cdot [x_i(\mathbf{b}_{-i}, b') - x_i(\mathbf{b}_{-i}, b)]. \quad (8)$$

Henceforth in this proof, we always fix an arbitrary $\mathbf{b}_{-i}$. For simplicity, in this proof, we omit $\mathbf{b}_{-i}$ and use the short-hand notations $x_i(v) := x_i(\mathbf{b}_{-i}, v)$, $p_i(v) := p_i(\mathbf{b}_{-i}, v)$, and $\mu(v) := \mu(\mathbf{b}_{-i}, v)$.

For $v \le 2\epsilon$, the lemma directly follows from (8). In the rest of the proof, we focus on the case where $v > 2\epsilon$. Define a function $u_i(b)$ such that $\int_0^b u_i(t)dt = b \cdot x_i(b) - p_i(b)$. For any $b \le b'$, the payment when bidding b is

$$p_i(b) = b \cdot x_i(b) - \int_0^b u_i(t)dt.$$

Since we do not have the guarantee that the utility increases with the bids, it can be that $u_i(b) \le 0$ for some b. However, we have that guarantee that at any point, $\int_0^b u_i(t)dt$ is non-negative.

By Equation (7), we know that for any $b \le b'$, we have $p_i(b') - p_i(b) \le b'[x_i(b') - x_i(b)] + \epsilon$, i.e.,

$$\left[b' \cdot x_i(b') - \int_0^{b'} u_i(t)dt\right] - \left[b \cdot x_i(b) - \int_0^b u_i(t)dt\right] \le b' \cdot [x_i(b') - x_i(b)] + \epsilon,$$

which is equivalent to

$$\xi(b, b') := (b' - b) \cdot x_i(b) - \int_b^{b'} u_i(t)dt \le \epsilon. \tag{9}$$

Intuitively, the meaning of $\xi(b, b')$ is how much we are over-estimating if we use a rectangle of width $b' - b$ and height $x_i(b)$ to approximate the area-under-curve[3] for u_i, between b and b'. For example, the blue area in Figure 5a represents $\xi(b, b')$, whereas the red area *minus* the gray area is $\xi(b'', v)$.

Now consider the following sequence: $b_l = v - \frac{v}{2^l}$ for $l = 0, \ldots, L$ where $L = \lceil \log \frac{v}{2\epsilon} \rceil$. By (8), the miner revenue

$$\mu(b_l) - \mu(b_{l-1}) \le 2\epsilon + S(b_{l-1}, b_l),$$

where $S(b_{l-1}, b_l) := (b_l - b_{l-1}) \cdot [x_i(b_l) - x_i(b_{l-1})]$. Summing up the miner revenue difference together, we have

$$\begin{aligned}
&\mu(v) - \mu(0) = \mu(v) - \mu(b_L) + \sum_{l=1}^{L} \mu(b_l) - \mu(b_{l-1}) \\
\le &2\epsilon + (v - b_L) \cdot [x_i(v) - x_i(b_L)] + \sum_{l=1}^{L} (S(b_{l-1}, b_l) + 2\epsilon) && \text{By (8)} \\
\le &4\epsilon + 2L\epsilon + \sum_{l=1}^{L} S(b_{l-1}, b_l). && \text{By } v - b_L \le 2\epsilon
\end{aligned}$$

Now we proceed to bound the sum $\sum_{l=1}^{L} S(b_{l-1}, b_l)$. For each $l = 1, \ldots, L$, by the choice of the sequence, we have

$$b_l - b_{l-1} = \frac{v}{2^l} = v - b_l, \quad \text{and} \quad S(b_{l-1}, b_l) = (v - b_l) \cdot [x_i(b_l) - x_i(b_{l-1})]$$

For simplicity, let $b_{L+1} := v$. We have the following:

$$\begin{aligned}
&\sum_{l=1}^{L} S(b_{l-1}, b_l) = \sum_{l=1}^{L} (v - b_l) \cdot [x_i(b_l) - x_i(b_{l-1})] \\
=&(v - b_L) \cdot x_i(b_L) + \sum_{l=1}^{L-1} (b_{l+1} - b_l) \cdot x_i(b_l) \\
=&\sum_{l=1}^{L} (b_{l+1} - b_l) \cdot x_i(b_l). && \text{By } v = b_{L+1}
\end{aligned}$$

[3] We may assume that any area under 0 contributes negatively to the area-under-curve.

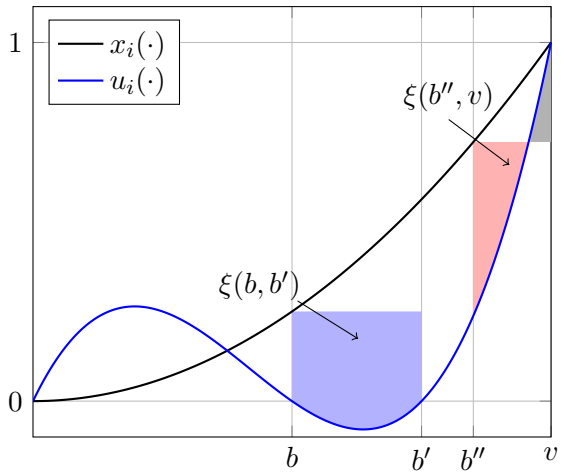
In other words, the sum $\sum_{l=1}^{L} S(b_{l-1}, b_l)$ is equal to the total area of the dashed rectangles in Figure 5b. We want to show that the sum $\sum_{l=1}^{L} S(b_{l-1}, b_l)$ is not significantly greater than $\mathsf{util}^i(v)$, i.e., the area under the u_i-curve. The follow calculation says that this difference is upper bounded by $\sum_{l=1}^{L} \xi(b_l, b_{l+1})$. Formally,

$$\begin{aligned}
&\sum_{l=1}^{L} S(b_{l-1}, b_l) - \int_0^v u_i(t)dt = \sum_{l=1}^{L}(b_{l+1} - b_l)x_i(b_l) - \int_0^v u_i(t)dt \\
\leq &\sum_{l=1}^{L}\left\{(b_{l+1} - b_l)\cdot x_i(b_l) - \int_{b_l}^{b_{l+1}} u_i(t)dt\right\} && \text{By } \int_0^{b_1} u_i(t) \geq 0 \\
= &\sum_{l=1}^{L} \xi(b_l, b_l + 1) \leq \sum_{l=1}^{L} \epsilon = L\epsilon. && \text{By (9)}
\end{aligned}$$

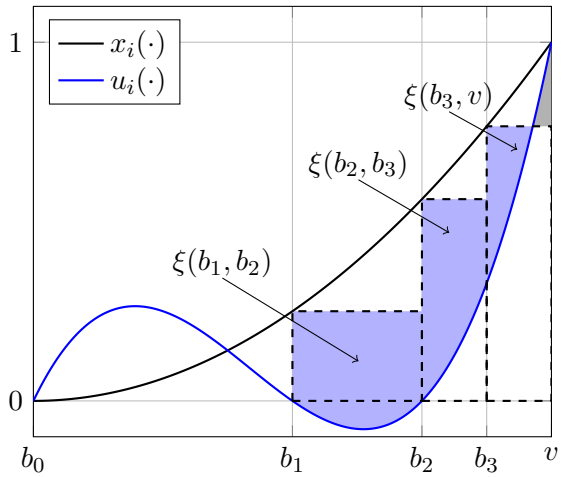
Putting it together, the change in miner revenue $\mu(v) - \mu(0)$ is upper bounded by

$$\begin{aligned}
&\mu(v) - \mu(0) \leq 4\epsilon + 2L\epsilon + \sum_{l=1}^{L} S(b_{l-1}, b_l) \\
\leq &4\epsilon + 2L\epsilon + L\epsilon + \int_0^v u_i(t)dt \leq \mathsf{util}^i(\mathbf{b}_{-i}, v) + 3\epsilon \log \frac{v}{\epsilon} + 4\epsilon,
\end{aligned}$$

where the last step comes from the fact that $L \leq \log \frac{v}{\epsilon}$ by our choice of L. □



(a) The blue area denotes $\xi(b, b')$, and the red area minus the gray area denotes $\xi(b'', v)$.



(b) The sum of the dashed rectangles is equal to $\sum_{l=1}^{L} S(b_l, b_{l+1})$. The difference between $\sum_{l=1}^{L} S(b_l, b_{l+1})$ and the area under the $u_i(\cdot)$ curve is upper bounded by $\sum_{l=1}^{L} \xi(b_l, b_{l+1})$, represented by the sum of the blue areas minus the gray area.

Figure 5: Graphical explanation of the proof to Lemma 4.3

Because the miner can inject a bid 0 for free, Lemma 4.3 implies the following corollary, which says that if we remove a bid, the miner revenue should not be affected by too much.

Corollary 4.4. *Let $(\mathbf{x}, \mathbf{p}, \mu)$ denote any (possibly randomized) TFM in the plain model that satisfies ϵ-UIC, ϵ-MIC and ϵ-SCP against 1-sized coalition. For any $\mathbf{b}_{-i}$ and v,*

$$\mu(\mathbf{b}_{-i}, v) - \mu(\mathbf{b}_{-i}) = \begin{cases} 5\epsilon, & v \leq 2\epsilon \\ \mathsf{util}^i(\mathbf{b}_{-i}, v) + 3\epsilon \log \frac{v}{\epsilon} + 5\epsilon, & v > 2\epsilon. \end{cases}$$

Proof. Because the miner can inject a bid 0 for free, by ϵ-MIC, it must be

$$\mu(\mathbf{b}_{-i}, 0) - \mu(\mathbf{b}_{-i}) \leq \epsilon. \tag{10}$$

The corollary is now directly implied by Equation (10) and Lemma 4.3. □

4.2.2 Bounds on Miner Revenue

We now prove bounds for the miner's revenue. To do this, we first prove a bound on miner revenue when everyone bids the same value M (see Lemma 4.5). Then, we generalize to the case when everyone's bids need not be the same (see Lemma 4.6).

Notation. Henceforth, for $t \in \mathbb{N} \cup \{0\}$, we define $\mathbf{m}_t := (M, \ldots, M)$ where $|\mathbf{m}_t| = t$; that is, $\mathbf{m}_t$ consists of t copies of M. Recall that $\mu(\mathbf{b})$ denote the expected miner revenue given that the world consists of the bid vector $\mathbf{b}$ (assuming the mechanism is honestly implemented). We define $\widetilde{\mu}(\mathbf{b}')$ to be the expected miner revenue given that *the block configuration* is $\mathbf{b}'$.

Lemma 4.5. *Suppose that the block size is upper bounded by k. Fix an arbitrary any $\epsilon > 0$ and $M > 2\epsilon$ and let $\mathbf{m}_t := (M, M, \ldots, M)$ be a vector containing t repetitions of M. Then, for any (possibly randomized) TFM in the plain model that satisfies ϵ-UIC, ϵ-MIC and ϵ-SCP even when the miner colludes with at most $c = 1$ user, it holds that $\widetilde{\mu}(\mathbf{m}_t) \leq 12k^2\epsilon \log \frac{M}{\epsilon}$ for all $t \leq k$.*

Proof. Imagine the world consists of the bid vector $\mathbf{m}_K$ where $K > \frac{Mk}{\epsilon}$ is sufficiently large. Let $\mathbf{m}_{t^*}$ be the block configuration that gives the miner optimal revenue; that is $t^* = \arg\max_{t \leq k} \widetilde{\mu}(\mathbf{m}_t)$. Clearly, it must be $\widetilde{\mu}(\mathbf{m}_{t^*}) \geq \mu(\mathbf{m}_K)$. Because of ϵ-MIC, we have $\mu(\mathbf{m}_{t^*}) \geq \widetilde{\mu}(\mathbf{m}_{t^*}) - \epsilon$. Otherwise, if $\mu(\mathbf{m}_{t^*}) < \widetilde{\mu}(\mathbf{m}_{t^*}) - \epsilon$, when the world is $\mathbf{m}_{t^*}$, the miner could simply choose $\mathbf{m}_{t^*}$ as the block configuration so that the revenue becomes $\widetilde{\mu}(\mathbf{m}_{t^*})$, which is more than ϵ higher than its honest utility $\mu(\mathbf{m}_{t^*})$. Combining the two inequalities, we have $\mu(\mathbf{m}_{t^*}) \geq \mu(\mathbf{m}_K) - \epsilon$.

Recall that $\mathsf{util}^i(\mathbf{b})$ denotes user i's expected utility when the bid vector is $\mathbf{b}$. Next, we will show that for any $t \leq K$ and any user $i \in [t]$, it must be

$$\mu(\mathbf{m}_t) + \mathsf{util}^i(\mathbf{m}_t) \leq \mu(\mathbf{m}_K) + 2\epsilon. \tag{11}$$

For the sake of reaching a contradiction, suppose there is an integer t and user i such that $\mu(\mathbf{m}_t) + \mathsf{util}^i(\mathbf{m}_t) > \mu(\mathbf{m}_K) + 2\epsilon$. Imagine that the world is $\mathbf{m}_K$, where $K > \frac{Mk}{\epsilon}$. There must exist a user j whose confirmation probability is at most $x_j(\mathbf{m}_K) \leq \frac{k}{K} < \frac{\epsilon}{M}$, as at most k bids can be included in a block. Therefore, user j's utility is at most $\mathsf{util}^j(\mathbf{m}_K) \leq x_j(\mathbf{m}_K) \cdot M < \epsilon$. Imagine that the miner now colludes with user j. The miner implements the inclusion rule as if the world consists of the bid vector $\mathbf{m}_t$ where the i-th position is occupied by user j's bid. Since the TFM is symmetric, and both users bid M, user j's expected utility is now $\mathsf{util}^i(\mathbf{m}_t)$. The joint utility of the coalition now is $\mu(\mathbf{m}_t) + \mathsf{util}^i(\mathbf{m}_t) > \mu(\mathbf{m}_K) + 2\epsilon > \mu(\mathbf{m}_K) + \mathsf{util}^j(\mathbf{m}_K) + \epsilon$, which contradicts ϵ-SCP. Consequently, Equation (11) must hold for any $t \leq K$ and any user $i \in [t]$.

According to Equation (11), we have $\mu(\mathbf{m}_{t^*}) + \mathsf{util}^i(\mathbf{m}_{t^*}) \le \mu(\mathbf{m}_K) + 2\epsilon$ for any user i. As we have shown, it must be $\mu(\mathbf{m}_{t^*}) \ge \mu(\mathbf{m}_K) - \epsilon$. Combining these two inequalities, we have

$$\mathsf{util}^i(\mathbf{m}_{t^*}) \le \mu(\mathbf{m}_K) + 2\epsilon - \mu(\mathbf{m}_{t^*}) \le \mu(\mathbf{m}_K) + 2\epsilon - \mu(\mathbf{m}_K) + \epsilon = 3\epsilon.$$

Since the utility of user i is bounded, by applying Corollary 4.4, it must be

$$\mu(\mathbf{m}_{t^*}) - \mu(\mathbf{m}_{t^*-1}) \le \mathsf{util}^i(\mathbf{m}_{t^*}) + 3\epsilon \log \frac{M}{\epsilon} + 5\epsilon \le 8\epsilon + 3\epsilon \log \frac{M}{\epsilon}. \tag{12}$$

Consequently, we have

$$\begin{aligned}
&\mathsf{util}^i(\mathbf{m}_{t^*-1}) \le \mu(\mathbf{m}_K) + 2\epsilon - \mu(\mathbf{m}_{t^*-1}) && \text{By (11)}\\
\le&\mu(\mathbf{m}_K) + 2\epsilon - \mu(\mathbf{m}_{t^*}) + 8\epsilon + 3\epsilon \log \frac{M}{\epsilon} && \text{By (12)}\\
\le&\mu(\mathbf{m}_K) + 2\epsilon - \mu(\mathbf{m}_K) + \epsilon + 8\epsilon + 3\epsilon \log \frac{M}{\epsilon} && \text{By } \mu(\mathbf{m}_{t^*}) \ge \mu(\mathbf{m}_K) - \epsilon\\
=&11\epsilon + 3\epsilon \log \frac{M}{\epsilon}.
\end{aligned}$$

Then, we can apply Corollary 4.4 again, and we have

$$\mu(\mathbf{m}_{t^*-1}) - \mu(\mathbf{m}_{t^*-2}) \le \mathsf{util}_i(\mathbf{m}_{t^*-1}) + 3\epsilon \log \frac{M}{\epsilon} + 5\epsilon \le 16\epsilon + 6\epsilon \log \frac{M}{\epsilon}.$$

By the same reason, for any $r \le t^*$, we have

$$\mu(\mathbf{m}_{t^*-r}) - \mu(\mathbf{m}_{t^*-r-1}) \le (8r+8)\epsilon + (3r+3) \cdot \epsilon \log \frac{M}{\epsilon}. \tag{13}$$

Since $M \ge 2\epsilon$, we have $\epsilon \log \frac{M}{\epsilon} \ge \epsilon$. By Eq.(13), we have

$$\begin{aligned}
&\mu(\mathbf{m}_{t^*}) - \mu(\mathbf{m}_0) = \sum_{r=0}^{t^*-1} \mu(\mathbf{m}_{t^*-r}) - \mu(\mathbf{m}_{t^*-r-1})\\
\le&\left(8t^* + 4(t^*-1)t^*\right)\epsilon + \left(3t^* + \frac{3(t^*-1)t^*}{2}\right) \cdot \epsilon \log \frac{M}{\epsilon}\\
\le&11(t^*)^2 \epsilon \log \frac{M}{\epsilon}. && \text{By } \epsilon \log \frac{M}{\epsilon} \ge \epsilon \text{ and } t^* \ge 1
\end{aligned}$$

Notice that $\mu(\mathbf{m}_0) = 0$, so we have

$$\mu(\mathbf{m}_{t^*}) \le 11(t^*)^2 \epsilon \log \frac{M}{\epsilon}.$$

Recall that we define $t^* = \arg\max_{t \le k} \widetilde{\mu}(\mathbf{m}_t)$. By definition, $\widetilde{\mu}(\mathbf{m}_t) \le \widetilde{\mu}(\mathbf{m}_{t^*})$ for all $t \le k$. As we have shown at the beginning, it must be $\mu(\mathbf{m}_{t^*}) \ge \widetilde{\mu}(\mathbf{m}_{t^*}) - \epsilon$. Thus, we have $\widetilde{\mu}(\mathbf{m}_t) \le \widetilde{\mu}(\mathbf{m}_{t^*}) \le \mu(\mathbf{m}_{t^*}) + \epsilon$ for all $t \le k$. Combine the arguments above, we have $\widetilde{\mu}(\mathbf{m}_t) \le 11k^2\epsilon \log \frac{M}{\epsilon} + \epsilon \le 12k^2\epsilon \log \frac{M}{\epsilon}$ for all $t \le k$.

□

Lemma 4.6. *Suppose the block size is upper bounded by k. Fix any $\epsilon > 0$. For any (possibly randomized) TFM in the plain model that satisfies ϵ-UIC, ϵ-MIC and ϵ-SCP (even when the miner only colludes with one user), for any block configuration $\mathbf{b}$, the following must hold where M is the maximum bid amount in the bid vector $\mathbf{b}$:*

$$\widetilde{\mu}(\mathbf{b}) \le \begin{cases} 2k\epsilon, & \text{if } M < 2\epsilon,\\ 12k^2\epsilon \log \frac{M}{\epsilon} + 2k\epsilon, & \text{if } M \ge 2\epsilon. \end{cases}$$

Proof. Given any block configuration $\mathbf{b}$, the miner revenue must be upper bounded by the sum of the bids in $\mathbf{b}$. Thus, if $M < 2\epsilon$, the miner revenue is upper bounded by $2k\epsilon$.

Henceforth, we focus on the case $M \geq 2\epsilon$. Throughout the proof, we say that a bid b is a *low bid* if $b < M$. Then, any block configuration, up to reordering, can be represented by $(\mathbf{m}_t, \mathbf{L})$ for some $t \geq 1$, where $\mathbf{m}_t$ consists of t repetitions of M, $\mathbf{L}$ which is possibly of length 0, contains only low bids. We prove the following claim by induction on the length of $\mathbf{L}$:

> *For any $\mathbf{L}$ consisting of only low bids, for any t such that $t + |\mathbf{L}| \leq k$, the miner revenue $\widetilde{\mu}(\mathbf{m}_t, \mathbf{L}) \leq \tau + 2|\mathbf{L}|\epsilon$, where we set $\tau := 12k^2\epsilon \log \frac{M}{\epsilon}$.*

For the base case where $|\mathbf{L}| = 0$, i.e. the block does not contain any low bid, it is proven by Lemma 4.5.

Now, suppose we have proven that for any $\mathbf{L}'$ of length R, for any t, the miner revenue $\widetilde{\mu}(\mathbf{m}_t, \mathbf{L}') \leq \tau + 2R\epsilon$. We are going to show that for any $\mathbf{L}$ of length $R+1$, for any t, the miner revenue $\widetilde{\mu}(\mathbf{m}_t, \mathbf{L}) \leq \tau + 2(R+1)\epsilon$.

For the sake of contradiction, suppose there exists a bid $\mathbf{L}$ of length $R+1$ and there exists a t, such that for the block configuration $(\mathbf{m}_t, \mathbf{L}) = (\mathbf{m}_t, d_1, \ldots, d_R, d_{R+1})$, the miner's revenue is $\tau + 2(R+1)\epsilon + \delta$ for some $\delta > 0$. Now, imagine that the world consists of $(\mathbf{m}_K, d_1, \ldots, d_R)$, where $K > \frac{kM}{\epsilon}$. In this case, the block configuration output by the honest inclusion rule must be of the form $(\mathbf{m}_{t^*}, \mathbf{d})$ for some $t^* \leq k - |\mathbf{d}|$ and $\mathbf{d} \subseteq \{d_1, \ldots, d_R\}$ consists of only low bids. Since $(\mathbf{m}_{t^*}, \mathbf{d})$ only contains at most R low bids, the miner revenue $\widetilde{\mu}(\mathbf{m}_{t^*}, \mathbf{d}) \leq \tau + 2R\epsilon$ by induction hypothesis.

By our choice of K, there must exist a user i with true value M, whose confirmation probability $x_i(\mathbf{m}_K, d_1, \ldots, d_R) \leq \frac{k}{K} < \frac{\epsilon}{M}$ when the miner is honest. Thus, user i's utility is at most $M \cdot x_i(\mathbf{m}_K, d_1, \ldots, d_R) < \epsilon$. Now the miner can collude with user i, ask user i to bid d_{R+1} instead of its true value M and include $(\mathbf{m}_t, d_1, \ldots, d_R, d_{R+1})$ in the block. Since $d_{R+1} < M$ and the payment never exceeds the bid, user i's utility is at least zero. This implies that the decrease of the utility of user i is strictly less than ϵ. Now the miner revenue is $\tau + 2(R+1)\epsilon + \delta$ by our assumption, whereas the miner revenue in the honest case is at most $\tau + 2R\epsilon$. Thus, the miner revenue increases by more than 2ϵ compared to the honest case. Thus, the joint utility of the coalition increases by more than ϵ, which contradicts ϵ-SCP. Therefore, by induction, we have that $\mu(\mathbf{m}_t, \mathbf{L}) \leq \tau + 2|\mathbf{L}|\epsilon$ for any $\mathbf{L}$ and any t where $|\mathbf{L}| + t \leq k$. Finally, since $|\mathbf{L}| \leq k$, we conclude that $\widetilde{\mu}(\mathbf{b}) \leq 12k^2\epsilon \log \frac{M}{\epsilon} + 2k\epsilon$. □

4.2.3 Completing the Proof of Theorem 4.1

We now complete the proof of Theorem 4.1. To do so, we prove that each user's utility conditioned on being included must be bounded given that the miner revenue is bounded (see Lemma 4.7), which then leads to our conclusion that the total social welfare must be small.

Lemma 4.7. *Suppose that the block size is upper bounded by k. Fix any $\epsilon > 0$. For any (possibly randomized) TFM in the plain model satisfies ϵ-UIC and ϵ-SCP (even when the miner colludes with only one user), for any bid vector $\mathbf{b}$ where $M := \max(\mathbf{b})$, for and any user i, conditioned on user i being included in the block, user i's utility must be upper bounded by $U + \epsilon$ where $U = \max_{|\mathbf{b}'| \leq k, \max(\mathbf{b}') \leq M} \widetilde{\mu}(\mathbf{b}')$, i.e., U is the maximum possible revenue the miner can get among all possible block configurations where all bids are at most M.*

Proof. For the sake of contradiction, suppose that under some bid vector $\mathbf{b}'$ where all bids are at most M, some user j's expected utility conditioned on being included in the block is strictly more than $U + \epsilon$. This implies that there must exist a block configuration $\mathbf{b} = (b_1, \ldots, b_{|\mathbf{b}|})$ where all

bids are at most M, and some $i \le |\mathbf{b}|$, such that under conditioned on the block configuration being $\mathbf{b}$, the i-th bid b_i in the block has expected utility at least $U + \epsilon + \delta$ for some positive δ. Let $T = \lceil \frac{b_i k}{\delta} \rceil + 1$. Imagine that the world consists of the bid vector $\mathbf{b}'$ of length $T + |\mathbf{b}|$ where

$$\mathbf{b}' = (\mathbf{b}, \underbrace{b_i, b_i, \ldots, b_i}_{T}).$$

Because the block size is upper bounded by k, there must exist a user j whose bid is b_i while its confirmation probability is at most $\frac{k}{T}$. Therefore, if user j bids truthfully, its utility is at most $b_i \cdot \frac{k}{T} < \delta$. By our assumption, the miner revenue is at most U under any block configuration where bids are upper bounded by M. Thus, when behaving honestly, the miner and user j have joint utility strictly less than $U + \delta$. However, the miner can collude with user j and prepare the block where the block configuration is $\mathbf{b}$ and the i-th position is replaced with user j's bid instead. In this case, user j's utility is $U + \epsilon + \delta$. Because the coalition does not inject any fake bid, the miner's utility is at least zero. Thus, by deviating from the mechanism, the joint utility of the coalition becomes at least $U + \epsilon + \delta$, which exceeds the honest case by more than ϵ. This contradicts ϵ-SCP. □

Proof of Theorem 4.1. Suppose the world consists of an arbitrary bid vector $\mathbf{b}$. Let $M = \max(\mathbf{b})$. If $M < 2\epsilon$, the miner can have at most $2k\epsilon$-miner revenue by Lemma 4.6. For any user i who is bidding truthfully, its true value must be upper bounded by M since $M = \max(\mathbf{b})$. Moreover, each confirmed user's utility is at most its true value, which is upper bounded by $M < 2\epsilon$. Since there are at most k number of confirmed user, the expected social welfare is $\sum_i \mathsf{util}^i(\mathbf{b})$ plus the miner's expected utility, which is upper bounded by $4k\epsilon$.

In the rest of the proof, we assume $M \ge 2\epsilon$ and we define $\widehat{\mathsf{util}}^i(\mathbf{b})$ to be the utility of user i conditioned on being confirmed when the world consists of the bid vector $\mathbf{b}$. By Lemma 4.6, the miner can have at most $\left(12k^2\epsilon \log \frac{M}{\epsilon} + 2k\epsilon\right)$-miner revenue. By Lemma 4.7, for any i, $\widehat{\mathsf{util}}^i(\mathbf{b}) \le 12k^2\epsilon \log \frac{M}{\epsilon} + (2k+1)\epsilon$. Let γ_i be the probability that user i is included in the block given the bid vector $\mathbf{b}$. Observe that $\sum_i \gamma_i \le k$ for any $\mathbf{b}$. Therefore, the expected total utility of all users is upper bounded by

$$\sum_i \mathsf{util}^i(\mathbf{b}) = \sum_i \widehat{\mathsf{util}}^i(\mathbf{b}) \cdot \gamma_i \le \left(12k^2\epsilon \log \frac{M}{\epsilon} + (2k+1)\epsilon\right) \cdot \sum_i \gamma_i = O\left(k^3\epsilon \log \frac{M}{\epsilon}\right).$$

The expected social welfare is $\sum_i \mathsf{util}^i(\mathbf{b})$ plus the miner's expected utility. Clearly, it is also upper bounded by $O\left(k^3\epsilon \log \frac{M}{\epsilon}\right)$.

Combine the argument above, because $\log\left(\frac{M}{\epsilon} + 1\right)$ is always non-negative, the theorem follows.

5 Characterization for Finite Block Size in the MPC-Assisted Model

5.1 Characterization for Strict Incentive Compatibility

In this section, we give a characterization of strict incentive compatibility in the MPC-assisted model for finite block size. We show that cryptography helps us overcome the finite-block impossibility [CS21] for $c = 1$, but for $c \ge 2$, the impossibility still holds.

5.1.1 Feasibility for $c = 1$

In the MPC-assisted model, we indeed can have a mechanism that achieves UIC, MIC, and $(\rho, 1)$-SCP against a coalition controlling $\rho \in (0, 1]$ fraction of the miners and $c = 1$ user.

MPC-assisted, posted price auction with random selection

Parameters: the reserved price r, and a block size k.
Input: a bid vector $\mathbf{b} = (b_1, \ldots, b_N)$.
Mechanism:

- *Allocation rule.* Any bid that is at least r is considered as a candidate. Randomly select k bids from the candidates to confirm.
- *Payment rule.* Each confirmed bid pays r.
- *Miner revenue rule.* Miner gets 0 revenue.

In the above mechanism, the miner gains zero revenue. This is inevitable as shown in Theorem C.5 of Appendix C.1. Even in the MPC-assisted model, the miner must have zero revenue if we insist on strict incentive compatibility (even under Bayesian notions of equilibrium).

Theorem 5.1. *Assuming a finite block size k. The above MPC-assisted, posted price auction with random selection in the MPC-assisted model satisfies UIC, MIC, and $(\rho, 1)$-SCP (in the ex post setting) for arbitrary $\rho \in (0, 1]$.*

Proof. We will prove the three incentive compatibility properties separately.

UIC. Let v_i denote the true value of user i. First, refusing to bid cannot increase its utility. Moreover, injecting bids does not help either. To see this, assume that user i bids its true value v_i and injects a bid b'. If $b' < r$, then it does not influence user i's utility. If $b' \geq r$, it either decreases the probability of user i being confirmed if $v_i \geq r$, or it brings user i negative expected utility if $v_i < r$.

Thus, we only need to argue that overbidding or underbidding does not increase the user's utility. If user i's true value $v_i < r$, then its utility when overbidding $b \geq r$ is $q \cdot (v_i - r) < 0$, where q is the probability of b being confirmed. If user i's true value $v_i \geq r$, then underbidding $b < r$ brings it 0-utility, whereas the honest utility $q(v_i - r)$ is positive. Therefore, no matter how user i deviates from the protocol, its utility does not increase.

MIC. Since the total miner revenue is always 0, injecting fake bids does not increase the colluding miner's utility. The miner cannot increase its utility by deviating from the protocol.

SCP. No matter how the coalition deviates, the colluding miner's revenue is always 0. Therefore, the joint utility of the coalition is at most the utility of the colluding user. By strict UIC, the joint utility does not increase. □

Note that the above mechanism does not work for $c = 2$. Imagine that the miner colludes with two users i and j, where user i has true value exactly r and user j has a sufficiently large true value. User i may choose not to bid to increase the probability of user j being confirmed. This brings the coalition strictly more utility than behaving honestly.

5.1.2 Impossibility for $c \geq 2$

Unfortunately, even in the MPC-assisted model, no mechanism with non-trivial utility can achieve UIC, MIC, and $(\rho, 2)$-SCP, even for Bayesian notions of incentive compatibility. To see this, observe that under the strict incentive compatible notion, (ρ, c)-SCP implies that any coalition of $\leq c$ users cannot benefit from any deviation[4], since the miner revenue has to be 0 by Theorem C.5 of Appendix C.1. Similar to the proof in Goldberg and Hartline [GH05], we show that any mechanism that is Bayesian UIC and Bayesian SCP against a $(\rho, 2)$-sized coalition (for an arbitrary $\rho \in (0, 1]$ must satisfy the following condition: no matter how a user j changes its bid, user i's utility should not change. Formally,

Lemma 5.2. *Given any (possibly random) mechanism in the MPC-assisted model that is Bayesian UIC and Bayesian SCP against $(\rho, 2)$-sized coalition for some $\rho \in (0, 1]$, and suppose each user's true value is drawn i.i.d. from a distribution $\mathcal{D}$. Then, for any user i and user j, for any bid b_j and b'_j, it must be that for any $\ell \geq 1$,*

$$\mathop{\mathbf{E}}_{(v, \mathbf{b}_{-i,j}) \sim \mathcal{D}^\ell}[\mathsf{util}^i(v, b_j, \mathbf{b}_{-i,j})] = \mathop{\mathbf{E}}_{(v, \mathbf{b}_{-i,j}) \sim \mathcal{D}^\ell}[\mathsf{util}^i(v, b'_j, \mathbf{b}_{-i,j})],$$

where $\mathbf{b}_{-i,j}$ represents all except user i and user j's bids.

The proof of this lemma is deferred to Appendix C.2. This lemma implies that no matter how user j changes its bid, the expected utility of user i should not change if user i's true value is sampled randomly from $\mathcal{D}$. Consequently, we have the following result stating that user i's utility should remain the same when bidding its true value, regardless of how many users are there.

Lemma 5.3. *Given any (possibly randomized) mechanism in the MPC-assisted model that achieves Bayesian UIC and Bayesian SCP against $(\rho, 2)$-sized coalition for some $\rho \in (0, 1]$, it holds that for any user i and j, for any bid b_j, for any $\ell \geq 1$,*

$$\mathop{\mathbf{E}}_{(v_i, \mathbf{b}_{-i,j}) \sim \mathcal{D}^\ell}[\mathsf{util}^i(v_i, b_j, \mathbf{b}_{-i,j})] \leq \mathop{\mathbf{E}}_{(v_i, \mathbf{b}_{-i,j}) \sim \mathcal{D}^\ell}[\mathsf{util}^i(v_i, \mathbf{b}_{-i,j})],$$

where v_{id} (b_{id}) denotes a bid v (b) coming from identity id.

Proof roadmap for Lemma 5.3. By Lemma 5.2, $\mathop{\mathbf{E}}_{v_i, \mathbf{b}_{-i,j} \sim \mathcal{D}^\ell}[\mathsf{util}^i(v_i, b_j, \mathbf{b}_{-i,j})] = \mathop{\mathbf{E}}_{v_i, \mathbf{b}_{-i,j} \sim \mathcal{D}^\ell}[\mathsf{util}^i(v_i, 0_j, \mathbf{b}_{-i,j})]$. Therefore, to prove Lemma 5.3, it suffices to prove that

$$\mathop{\mathbf{E}}_{\mathbf{b}_{-i,j} \sim \mathcal{D}^{\ell-1}}[\mathsf{util}^i(v_i, 0_j, \mathbf{b}_{-i,j})] \leq \mathop{\mathbf{E}}_{\mathbf{b}_{-i,j} \sim \mathcal{D}^{\ell-1}}[\mathsf{util}^i(v_i, \mathbf{b}_{-i,j})].$$

This claim is relatively easy to prove if we are willing to assume a *strong* symmetry assumption explained below. With a technically more involved proof, we can eventually get rid of this *strong* symmetry assumption and prove it under our current (much weaker) symmetry assumption defined in Section 2.1.

Strong symmetry assumption. On top of our current symmetric assumption defined in Section 2.1, we additionally assume that for any bid vector $\mathbf{b} := (b_1, \ldots, b_N)$, if for $i \neq j$, $b_i = b_j$, then the random variables (x_i, p_i) and (x_j, p_j) are identically distributed, where (x_i, p_i) are random variables denoting i's confirmation probability and i's payment, respectively, and (x_j, p_j) are similarly defined.

[4] We credit Bahrani, Garimidi, Roughgarden, Shi, and Weinberg for making this observation.

In other words, the strong symmetry assumption additionally assumes that two bids of the same amount receive the same treatment, on top of our existing symmetry assumption — note that this is a very strong assumption, and this is why we want to get rid of it eventually. If the above strong symmetry assumption holds, then we have that for any identity i' that injects a 0 bid,

$$\mathop{\mathbf{E}}_{\mathbf{b}_{-i,j}\sim\mathcal{D}^{\ell-1}}[\mathsf{util}^i(v_i, 0_j, \mathbf{b}_{-i,j})] = \mathop{\mathbf{E}}_{\mathbf{b}_{-i,j}\sim\mathcal{D}^{\ell-1}}[\mathsf{util}^i(v_i, 0_{i'}, \mathbf{b}_{-i,j})],$$

This is because under the strong symmetry assumption, anyone who bids the same amount as i has the same expected utility, and moreover, this utility is not affected by whether the 0 bid is posted by j or i'. Finally, we have for any v_i,

$$\mathop{\mathbf{E}}_{\mathbf{b}_{-i,j}\sim\mathcal{D}^{\ell-1}}[\mathsf{util}^i(v_i, 0_{i'}, \mathbf{b}_{-i,j})] \le \mathop{\mathbf{E}}_{\mathbf{b}_{-i,j}\sim\mathcal{D}^{\ell-1}}[\mathsf{util}^i(v_i, \mathbf{b}_{-i,j})].$$

Otherwise, user i can inject a 0-bid using an arbitrary identity i', which strictly increases its utility. This contradicts Bayesian UIC. We refer the reader to Appendix C.3 for a full proof of Lemma 5.3 without relying on the strong symmetry assumption.

Lemma 5.4. *Given any (possibly randomized) mechanism in the MPC-assisted model that achieves Bayesian UIC, MIC and Bayesian SCP against $(\rho, 2)$-sized coalitions for some $\rho \in (0, 1]$, it holds that for any user i, any value v_i, for any $\ell \ge 1$,*

$$\mathop{\mathbf{E}}_{v_i,\mathbf{b}_{-i}\sim\mathcal{D}^{\ell}}[\mathsf{util}^i(v_i, \mathbf{b}_{-i})] = \mathop{\mathbf{E}}_{v_i\sim\mathcal{D}}\mathsf{util}^i(v_i).$$

Proof. We first show that for any j, user i's expected utility should not change if user j refuses to bid. Formally, for any b_j,

$$\mathop{\mathbf{E}}_{v_i,\mathbf{b}_{-i,j}\sim\mathcal{D}^{\ell}}[\mathsf{util}^i(v_i, b_j, \mathbf{b}_{-i,j})] = \mathop{\mathbf{E}}_{v_i,\mathbf{b}_{-i,j}\sim\mathcal{D}^{\ell}}[\mathsf{util}^i(v_i, \mathbf{b}_{-i,j})]. \tag{14}$$

To see this, by Lemma 5.3, we have

$$\mathop{\mathbf{E}}_{v_i,\mathbf{b}_{-i,j}\sim\mathcal{D}^{\ell}}[\mathsf{util}^i(v_i, b_j, \mathbf{b}_{-i,j})] \le \mathop{\mathbf{E}}_{v_i,\mathbf{b}_{-i,j}\sim\mathcal{D}^{\ell}}[\mathsf{util}^i(v_i, \mathbf{b}_{-i,j})].$$

Next, we are going to show that

$$\mathop{\mathbf{E}}_{v_i,\mathbf{b}_{-i,j}\sim\mathcal{D}^{\ell}}[\mathsf{util}^i(v_i, b_j, \mathbf{b}_{-i,j})] = \mathop{\mathbf{E}}_{v_i,\mathbf{b}_{-i,j}\sim\mathcal{D}^{\ell}}[\mathsf{util}^i(v_i, 0_j, \mathbf{b}_{-i,j})] \ge \mathop{\mathbf{E}}_{v_i,\mathbf{b}_{-i,j}\sim\mathcal{D}^{\ell}}[\mathsf{util}^i(v_i, \mathbf{b}_{-i,j})].$$

To see why this holds, note that the first equality follows from Lemma 5.2. The inequality comes from 2-SCP: Since by MIC, it must be that $\mathop{\mathbf{E}}_{\mathbf{b}_{-i,j}\sim\mathcal{D}^{\ell}}[\mu(v_i, 0_j, \mathbf{b}_{-i,j})] \le \mathop{\mathbf{E}}_{\mathbf{b}_{-i,j}\sim\mathcal{D}^{\ell}}[\mu(v_i, \mathbf{b}_{-i,j})]$, therefore, it must be that $\mathop{\mathbf{E}}_{v_i,\mathbf{b}_{-i,j}\sim\mathcal{D}^{\ell}}[\mathsf{util}^i(v_i, 0_j, \mathbf{b}_{-i,j})] \ge \mathop{\mathbf{E}}_{v_i,\mathbf{b}_{-i,j}\sim\mathcal{D}^{\ell}}[\mathsf{util}^i(v_i, \mathbf{b}_{-i,j})]$. Otherwise, if there exists a v_i such that this does not hold, the miner can collude with user i and user j with true value 0, and ask user j not to bid. This strategy strictly increases the coalition's joint utility and thus contradicts Bayesian SCP against $(\rho, 2)$-sized coalition. Equation (14) thus follows.

Let $f(\cdot)$ denote the p.d.f. of $\mathcal{D}$. By definition of expectation,

$$\begin{aligned}
&\mathop{\mathbf{E}}_{v_i,\mathbf{b}_{-i}\sim\mathcal{D}^{\ell}}[\mathsf{util}^i(v_i, \mathbf{b}_{-i})] = \int_0^\infty \mathop{\mathbf{E}}_{v_i,\mathbf{b}_{-i,1}\sim\mathcal{D}^{\ell-1}}[\mathsf{util}^i(v_i, z_1, \mathbf{b}_{-i,1})] f(z_1) dz_1 \\
&= \int_0^\infty \mathop{\mathbf{E}}_{v_i,\mathbf{b}_{-i,1}\sim\mathcal{D}^{\ell-1}}[\mathsf{util}^i(v_i, \mathbf{b}_{-i,1})] f(z_1) dz_1 && \text{By Equation (14)} \\
&= \mathop{\mathbf{E}}_{v_i,\mathbf{b}_{-i,1}\sim\mathcal{D}^{\ell-1}}[\mathsf{util}^i(v_i, \mathbf{b}_{-i,1})].
\end{aligned}$$

The lemma follows by repeating the above argument. □

Now we are ready to prove the theorem stating that there is no mechanism that gives non-zero utility to either users or miners and yet satisfies Bayesian UIC and SCP against $(\rho, 2)$-sized coalitions.

Theorem 5.5. *Suppose the block size is k. No MPC-assisted mechanism with non-trivial utility simultaneously achieves Bayesian UIC, MIC and Bayesian SCP against $(\rho, 2)$-sized coalitions.*

Proof. By Theorem C.5, the miner-revenue has to be 0. Therefore, it suffices to prove that every user must have 0-utility.

Consider a crowded world with K number of users and all of their bids are sampled independently at random from $\mathcal{D}$. There must exist a user j^* whose probability of being confirmed is at most k/K, and thus its expected utility is at most $\max(\mathcal{D}) \cdot k/K$ where k is the block size. Thus, $\mathop{\mathbf{E}}_{\mathbf{b}\sim\mathcal{D}^K}[\mathsf{util}^{j^*}(\mathbf{b})] = 0$ by taking K to be arbitrarily large.

By Lemma 5.4, it must be that $\mathop{\mathbf{E}}_{v_{j^*}\sim\mathcal{D}} \mathsf{util}^i(v_{j^*}) = 0$. Since $\mathop{\mathbf{E}}_{v_i\sim\mathcal{D}} \mathsf{util}^i(v_i) = \mathop{\mathbf{E}}_{v_j\sim\mathcal{D}} \mathsf{util}^j(v_j)$ for all user i and j, every user's expected utility must be the 0 where the expectation is taken over the randomness of bids as well as the randomness of the mechanism, i.e., for any user i, any $\ell \geq 1$, we have $\mathop{\mathbf{E}}_{\mathbf{b}\sim\mathcal{D}^\ell}[\mathsf{util}^i(\mathbf{b})] = 0$. Since user's utility is non-negative, this implies that $\mathop{\mathbf{E}}_{\mathbf{b}_{-i}\sim\mathcal{D}^{\ell-1}}[\mathsf{util}^i(v, \mathbf{b}_{-i})] = 0$.

□

5.2 Feasibility of Approximate Incentive Compatibility

Although strict (even Bayesian) incentive compatibility is impossible to achieve for $c \geq 2$ in the MPC-assisted model, we have meaningful feasibility results if we allow ϵ additive slack. Still, we use k to denote the finite block size and M to denote the upper bound of the true values. Specifically, we can achieve $\Theta(kM)$ social welfare as long as many people place high enough bids, which is asymptotically the best possible social welfare one can hope for.

MPC-assisted, Diluted Posted Price Auction

Parameters: the block size k, an upper bound c of the number of users colluding with the miner, an upper bound M of users' true values, a slack $\epsilon \geq 0$, and a posted-price r such that $r \geq \frac{\epsilon}{2c}$.

Input: a bid vector $\mathbf{b} = (b_1, \ldots, b_N)$.

Mechanism:

1. *Allocation rule.*
   - Given a bid vector $\mathbf{b} = (b_1, \ldots, b_N)$, remove all bids which are smaller than r. Let $\widetilde{\mathbf{b}} = (\widetilde{b}_1, \ldots, \widetilde{b}_\ell)$ denote the resulting vector.
   - Let $T = \max\left(2c\sqrt{\frac{kM}{\epsilon}}, k\right)$. If $\ell \geq T$, let $\mathbf{d} = \widetilde{\mathbf{b}}$. Else, let $\mathbf{d} = (\widetilde{b}_1, \ldots, \widetilde{b}_\ell, 0, \ldots, 0)$ such that $|\mathbf{d}| = T$. In other words, $\mathbf{d}$ is $\widetilde{\mathbf{b}}$ appended with $T - \ell$ zeros.
   - Randomly choose a set S of size k from $\mathbf{d}$, and every non-zero bid in S is confirmed.
2. *Payment rule.* For each confirmed bid b, it pays r.
3. *Miner revenue rule.* For each confirmed bid b, the miner is paid $\frac{\epsilon}{2c}$.

Theorem 5.6. *Suppose there exists an upper bound M on users' true values. The above MPC-assisted, diluted posted price auction satisfies UIC, MIC, and ϵ-SCP (in the ex post setting) against (ρ, c)-sized coalitions for arbitrary $\rho \in (0, 1]$ and $c \geq 1$.*

Proof. We will prove the three incentive compatibility properties separately. Note that in this mechanism, refusing to bid is equivalent to underbidding some value less than r. So we mainly focus on the strategy space of bidding untruthfully and injecting bids. When we say the expected utility of a user, the randomness is taken over the randomness in the mechanism.

UIC. Fix any user i, and let v denote the true value of user i. In the mechanism, any confirmed bid pays r and any bid less than r must be unconfirmed. Thus, if $v \leq r$, bidding untruthfully cannot give a postive utility, so bidding truthfully and getting 0-utility is optimal.

Below we focus on the case when $v > r$. In this case, the bid has a non-negative probability of being confirmed and it pays r. So following the honest strategy leads to positive utility. Bidding less than r will cause the bid to be unconfirmed and will not help the user. Therefore, we may assume that the user bids at least r and may inject some fake bids. Observe that any bid that is at least r is treated the same by the mechanism. Moreover, injecting fake bids either make no difference (when $\ell \leq T$ after injecting), or it reduces the probability of bid v being elected into the set S (when $\ell > T$ after injecting). Therefore, bidding untruthfully and/or injecting fake bids does not help the user.

MIC. By injecting fake bids, the strategic miner cannot increase the expected number of real bids in the vector $\mathbf{d}$. Thus, injecting fake bids cannot increase other bids' contribution towards the miner's revenue. Therefore, the expected gain in miner revenue must be upper bounded by the fake bids' contribution towards miner revenue minus the expected payments of the fake bids. For each confirmed bid, the miner revenue is fixed to $\frac{\epsilon}{2c}$, which is no more than the payment of the bid. Thus, the expected miner revenue cannot increase through injecting fake bids, i.e., the mechanism is MIC.

ϵ-SCP. First, we argue that injecting bids does not help the coalition. Specifically, using a similar proof as UIC, injecting bids does not help improve the utility of any user in the coalition. Using a similar argument as MIC, injecting bids does not improve the miner's revenue minus the payment of the injected bids. Therefore, injecting bids will not increase the coalition's joint utility.

Now it suffices to argue that underbidding or overbidding does not increase the coalition's joint utility by more than ϵ. Suppose when bidding honestly, the number of bids in $\widetilde{\mathbf{b}}$ is ℓ. Each bid in $\widetilde{\mathbf{b}}$ is confirmed with probability $\frac{k}{\max\{T,\ell\}}$. Assume that by bidding untruthfully, the coalition changes the length of $\widetilde{\mathbf{b}}$ to ℓ'. Now each bid in $\widetilde{\mathbf{b}}$ is confirmed with probability $\frac{k}{\max\{T,\ell'\}}$.

We partition the players in the coalition into the following groups:

- Those whose true values are less than r and bid less than r. Their expected utility does not change.
- Those whose true values are less than r and bid higher than or equal to r. Their expected utility does not increase.
- Those whose true values are at least r and bid less than r. Their expected utility does not increase.

- Those whose true values are at least r and bid at least r. For each of these users, its expected utility increases by at most

$$(v-r)\frac{k}{\max\{T,\ell'\}} - (v-r)\frac{k}{\max\{T,\ell\}}. \tag{15}$$

Note that for $\ell' \geq \ell$, then (15) ≤ 0. Therefore, we only need to consider the case where $\ell' < \ell$. If $\ell \leq T$, then (15) is 0. If $\ell > T$, then (15) is upper bounded by

$$\begin{aligned}(15) &\leq (v-r)\left[\frac{k}{\ell'} - \frac{k}{\ell}\right] \\ &\leq (v-r)\left[\frac{k}{\ell-c} - \frac{k}{\ell}\right] \leq (v-r)\frac{ck}{\ell(\ell-c)} \\ &\leq M \cdot \frac{ck}{T(T-c)}.\end{aligned}$$

By the choice of T, we have that $T(T-c) \geq \frac{1}{2}T^2$. Thus,

$$(15) \leq M \cdot \frac{ck}{T(T-c)} \leq \frac{2Mck}{T^2} \leq \frac{\epsilon}{2c}.$$

This implies that each user's utility can increase by at most $\frac{\epsilon}{2c}$. Meanwhile, for each user in the coalition, it can increase the miner's revenue by no more than $\frac{\epsilon}{2c}$ via bidding untruthfully. Since there are at most c users in the coalition, the coalition can gain at most ϵ more utility in total, no matter how they deviate.

□

Acknowledgments

This work is in part supported by NSF awards 2212746, 2044679, 1704788, a Packard Fellowship, a generous gift from the late Nikolai Mushegian, a gift from Google, and an ACE center grant from Algorand Foundation. The authors would like to thank the anonymous reviewers for their helpful comments. We also thank Matt Weinberg for helpful technical discussions regarding how to efficiently instantiate our MPC-assisted mechanisms.

References

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

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

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

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

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

[Can00] Ran Canetti. Security and composition of multiparty cryptographic protocols. *Journal of Cryptology*, 2000.

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

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

[CS21] Hao Chung and Elaine Shi. Foundations of transaction fee mechanism design. *arXiv preprint arXiv:2111.03151*, 2021.

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

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

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

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

[GH05] Andrew V. Goldberg and Jason D. Hartline. Collusion-resistant mechanisms for single-parameter agents. In *SODA 2005*, pages 620–629, 2005.

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

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

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

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

[GO14] Jens Groth and Rafail Ostrovsky. Cryptography in the multi-string model. *Journal of cryptology*, 27(3):506–543, 2014.

[GPS19] Yue Guo, Rafael Pass, and Elaine Shi. Synchronous, with a chance of partition tolerance. In *Annual International Cryptology Conference*, pages 499–529. Springer, 2019.

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

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

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

[IML05] Sergei Izmalkov, Silvio Micali, and Matt Lepinski. Rational secure computation and ideal mechanism design. In *FOCS*, 2005.

[Kat08] Jonathan Katz. Bridging game theory and cryptography: Recent results and future directions. In *TCC*, 2008.

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

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

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

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

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

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

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

[RBO89] T. Rabin and M. Ben-Or. Verifiable secret sharing and multiparty protocols with honest majority. In *STOC*, 1989.

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

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

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

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

A Full Proof of Theorem 3.6

We now prove Theorem 3.6 of Section 3.2, i.e., the propotional auction in the plain model satisfies UIC, MIC, and $\frac{5}{4}c\epsilon$-SCP against any miner-user coalition with an arbitrary $c \geq$ number of users.

Proof of Theorem 3.6. We prove the three properties individually.

UIC. Because the confirmation and the payment of each bid are independent of other bids, injecting fake bids does not help to increase any user's utility. Next, suppose user i's true value is v_i. If user i bids b_i, its expected utility is

$$\begin{cases} \left(v_i - \frac{b_i}{2}\right)\frac{b_i}{r}, & \text{if } b_i < r, \\ v_i - \frac{r}{2}, & \text{if } b_i \geq r. \end{cases}$$

By direct calculation, the expected utility is maximized when $b_i = v_i$. Thus, proportional auction is strict UIC.

MIC. Since the block size is infinite, the miner's best strategy is to include all bids to maximize its revenue. Notice that the confirmation of each bid and the miner revenue of each bid are independent of other bids. Thus, injecting fake bids does not change the miner revenue from "other bids." Moreover, for each confirmed bid, the miner revenue is upper bounded by the payment of that bid. Thus, the increment of the miner revenue never exceeds the cost of the injected fake bids. Thus, the miner revenue cannot increase by injecting fake bids, so the mechanism is strict MIC.

$\frac{5}{4}c\epsilon$-SCP. As we have shown in the argument for strict UIC and strict MIC, injecting fake bids does not change the colluding miner's revenue. Because the confirmation and the payment of each bid are independent of other bids, injecting fake bids does not help to increase any user's utility. Thus, in the rest of the proof, we assume the only deviation of the coalition is to change the bids from colluding users' true values to other values. Let user i be a colluding user. We will show that the joint utility increases at most by $\frac{5}{4}\epsilon$ if user i changes its bid from its true value to other values, no matter what other bids are. Because there are at most c colluding users, the mechanism is $\frac{5}{4}c\epsilon$-SCP for all c.

Let user i be a colluding user with true value v_i, and let b_i be user i's bid. We now proceed to analyze the utility of coalition based on how users in the coalition bid untruthfully.

1. Underbidding. Suppose $b_i < v_i$. Notice that the miner can get the payment from b_i only when b_i is confirmed, and the miner is paid $\frac{\sqrt{2r\epsilon}}{2}$ if $b_i \geq \sqrt{2r\epsilon}$. When user i underbids, the miner's revenue can not increase. Because the mechanism is strict UIC, underbidding does not increase user i's utility either. Thus, the joint utility does not increase if $b_i < v_i$.

2. Overbidding. Suppose $b_i > v_i$. We first consider the following cases based on whether the true value v_i is less than r.

- **If $v_i \geq \sqrt{2r\epsilon}$.** If $v_i \geq r$, bidding truthfully already guarantees user i's bid to be confirmed, and the miner is paid $\sqrt{\frac{r\epsilon}{2}}$. Thus, when $v_i \geq r$, overbidding does not increase the joint utility.

In the following, we assume $v_i < r$. Let $\Delta = \min(b_i - v_i, r - v_i) > 0$. If user i bids truthfully, its bid is confirmed with the probability $\frac{v_i}{r}$, so its expected utility is

$$\left(v_i - \frac{v_i}{2}\right)\frac{v_i}{r}.$$

Next, suppose user i bids $b_i > v_i$. Then, b_i is confirmed with the probability $\frac{v_i+\Delta}{r}$, and the payment is $\frac{v_i+\Delta}{2}$ if b_i is confirmed. Thus, user i's expected utility is

$$\left(v_i - \frac{v_i+\Delta}{2}\right)\frac{v_i+\Delta}{r}.$$

Hence, compared to bidding truthfully, user i's expected utility decreases by

$$\left(v_i - \frac{v_i}{2}\right)\frac{v_i}{r} - \left(v_i - \frac{v_i+\Delta}{2}\right)\frac{v_i+\Delta}{r} = \frac{\Delta^2}{2r} > 0.$$

On the other hand, if user i bids truthfully, the miner's expected revenue is $\frac{v_i}{r}\sqrt{\frac{r\epsilon}{2}}$. If user i bids $b_i > v_i$, the miner's expected revenue is $\frac{v_i+\Delta}{r}\sqrt{\frac{r\epsilon}{2}}$. Thus, compared to bidding truthfully, the miner's expected utility increases by

$$\frac{v_i+\Delta}{r}\sqrt{\frac{r\epsilon}{2}} - \frac{v_i}{r}\sqrt{\frac{r\epsilon}{2}} = \frac{\Delta}{r}\sqrt{\frac{r\epsilon}{2}}.$$

Combine the argument above, the joint utility increases by

$$\frac{\Delta}{r}\sqrt{\frac{r\epsilon}{2}} - \frac{\Delta^2}{2r}. \tag{16}$$

The maximum of Eq.(16) is $\frac{\epsilon}{4}$, so overbidding b_i can only increase the joint utility by $\frac{\epsilon}{4}$.

- **If $v_i < \sqrt{2r\epsilon}$.** Because the mechanism is strict-UIC, overbidding does not increase user i's utility. If $b_i < \sqrt{2r\epsilon}$, the miner revenue is still zero. Thus, we assume $b_i \geq \sqrt{2r\epsilon}$. From the argument in the previous case, we know that compared to bidding truthfully, user i's expected utility decreases by $\frac{\Delta^2}{2r}$. However, if user i bids truthfully, the miner's revenue is zero. If user i bids $b_i > v_i$, the miner's expected revenue is $\frac{v_i+\Delta}{r}\sqrt{\frac{r\epsilon}{2}}$. Thus, compared to bidding truthfully, the miner's expected revenue increases by $\frac{v_i+\Delta}{r}\sqrt{\frac{r\epsilon}{2}}$. Consequently, the joint utility increases by

$$\frac{v_i}{r}\sqrt{\frac{r\epsilon}{2}} + \frac{\Delta}{r}\sqrt{\frac{r\epsilon}{2}} - \frac{\Delta^2}{2r}. \tag{17}$$

Because the maximum of Eq.(16) is $\frac{\epsilon}{4}$, the maximum of Eq.(17) when $v_i < \sqrt{2r\epsilon}$ is at most $\frac{5\epsilon}{4}$. Thus, overbidding b_i can only increase the joint utility by $\frac{5\epsilon}{4}$.

To sum up, among all cases, overbidding b_i can only increase the joint utility by at most $\frac{5}{4}\epsilon$. The theorem thus follows.

Proportional auction for the MPC-assisted model. In the MPC-assisted model, we want to ensure incentive compatibility for any miner-user coalition controlling at most ρ fraction of the miners and at most c users — recall that the total miner revenue is split among the miners. By contrast, in the plain model, effectively ρ is always equal to 1 since we always focus on the miner of the present block. Therefore, to make the proportional auction work in the MPC-assisted model, we make a small modification to the scheme and proof. For the scheme, the only modification is that we now allow the miner revenue to scale up w.r.t. $\frac{1}{\rho}$ (up to the total user payment), such that the miner revenue can be larger if we only want to be resilient against coalitions controlling small fraction of the miners — see the formal description below.

Proportional Auction for the MPC-assisted Model

Parameters: the approximate factor ϵ, upper bound ρ on the fraction of the colluding miners and the reserved price r such that $r \geq 2\epsilon$.

Input: a bid vector $\mathbf{b} = (b_1, \ldots, b_N)$.

Mechanism:

- *Allocation rule.* For each bid b, if $b < r$, it is confirmed with the probability b/r; otherwise, if $b \geq r$, it is confirmed with probability 1.
- *Payment rule.* For each confirmed bid b, if $b < r$, it pays $b/2$; otherwise, if $b \geq r$, it pays $r/2$.
- *Miner revenue rule.* For each confirmed bid b, let p be the payment of b, and the miner is paid $\min\left(p, \frac{\sqrt{2r\epsilon}}{2\rho}\right)$.[a]

[a] The minimum guarantees that the miner revenue never exceed the payment.

It is not hard to see that our proof of Theorem 3.6 can be easily modified to work for the MPC-assisted model. The only difference is that when the colluding user's true value v_i is smaller than the threshold $\frac{\sqrt{2r\epsilon}}{\rho}$, overbidding to $v_i + \Delta < \frac{\sqrt{2r\epsilon}}{\rho}$ also increases the joint utility. There are two cases:

- When a user with true value v_i overbids to $v_i + \Delta < \frac{\sqrt{2r\epsilon}}{\rho}$, the coalition of the miner and this colluding user can gain at most ϵ more utility if $\Delta = \sqrt{2r\epsilon}$.
- When a user with true value v_i overbids to $v_i + \Delta \geq \frac{\sqrt{2r\epsilon}}{\rho}$, the coalition of the miner and this colluding user can gain at most $\frac{9}{4}\epsilon$ more utility when $\Delta = \sqrt{2r\epsilon}$ and v_i is arbitrarily close to $\frac{\sqrt{2r\epsilon}}{\rho}$.

B Feasibility: Approximate Incentive Compatibility for Finite Blocks

In this section, we give a mechanism, called staircase mechanism, that is ϵ-UIC, MIC, and ϵ-SCP for $c = 1$ in the plain model. The staircase mechanism can in the best case achieve $\Theta(k^2\epsilon)$ social welfare. Recall that in Theorem 4.1, we showed that any plain-model mechanism that works for finite block size suffers from poor scaling of the social welfare w.r.t. the bid distribution. In particular, we showed that the social welfare is upper bounded by $O(k^3\epsilon \log(M/\epsilon))$ where M is an upper bound on the social welfare. Our staircase mechanism can achieve $\Theta(k^2\epsilon)$ social welfare in the best case. Thus, we still have a gap between the upper and lower bounds. Bridging this gap is an interesting open problem.

Staircase Mechanism

Parameters: the block size k, the upper bound c of the colluding users, the upper bound M of the true value, the approximate factor ϵ.

Notations: We define

$$F_0 = \begin{cases} M - k\epsilon, & \text{if } \lfloor \frac{M}{\epsilon} \rfloor \geq k, \\ M - \lfloor \frac{M}{\epsilon} \rfloor \epsilon, & \text{if otherwise.} \end{cases}$$

For all $i = 1, \ldots, k$, we define

$$F_i = F_0 + i \cdot \epsilon.$$

Input: a bid vector $\mathbf{b} = (b_1, \ldots, b_N)$.

Mechanism:

1. *Inclusion rule.* Given the bid vector $\mathbf{b} = (b_1, \ldots, b_N)$, choose the top k bids.
2. *Confirmation rule.*
   - Let $\mathbf{c} = (c_1, \ldots, c_{N'})$ denote the bid vector in the block, where $c_1 \geq c_2 \geq \cdots \geq c_{N'}$ and $N' \leq k$.
   - If $c_1 < F_1$, set $t = 0$. Otherwise, set $t = \max_i \{i : c_i \geq F_i\}$.
   - If $t = 0$, no one is confirmed. Otherwise, $c_1, \ldots, c_t$ are confirmed.
3. *Payment rule.* For each confirmed bid, it pays F_t.
4. *Miner revenue rule.* Miner is paid $t \cdot \epsilon$.

In the staircase mechanism, the more bids confirmed, the higher the price. For example, let $M = 10$ be the maximum possible bid, let $\epsilon = 1$, and let the block size be $k = 5$. Thus, if only one user is confirmed, then the price would be set to 6; if two users are confirmed, the price would be 7; and so on. Now, if the bid vector is $10, 9, 5, 3, 1$, the mechanism would confirm the top two bids and they each pay 7. One can see that the mechanism achieves at least $\Theta(k^2\epsilon)$ social welfare in the best case: suppose $\lfloor \frac{M}{\epsilon} \rfloor \geq k$ and $k/2$ users have true value M while the remaining users have a value of 0. Then, all the $k/2$ bids at M will be confirmed and each bid pays $F_{k/2} = M - (k\epsilon/2)$. In this case, the mechanism achieves $\Theta(k^2\epsilon)$ social welfare.

Notice that the miner's revenue grows linearly in t, the number of the confirmed bids in the block. On the other hand, any confirmed user's payment also grows linearly in t, so each confirmed user's utility actually decreases linearly in t. The miner's revenue and any user's utility as the functions of t can be visualized by Figure 6, which explains why the mechanism is called "staircase".

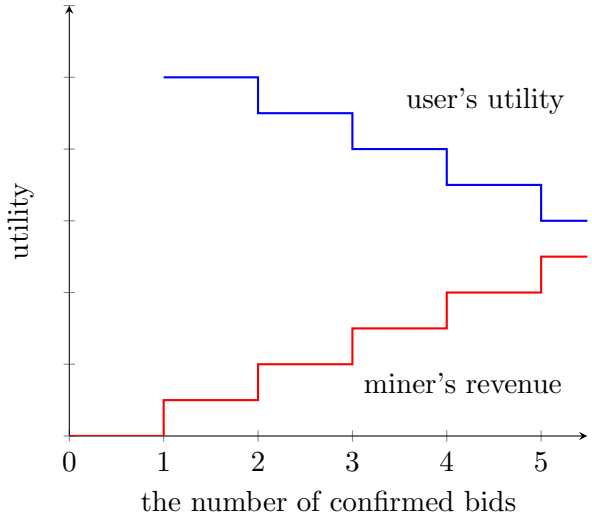


Figure 6: The miner's revenue and any user's utility as the functions of the number of the confirmed bids in the block.

Intuitively, for any coalition consisting of the miner and a user, they do not have incentive to manipulate the number of confirmed bids, as the increase in miner revenue cancels out the decrease in the colluding user's utility. The following example shows that a user or a miner-user coalition may have ϵ extra utility by deviation. Suppose $M = k = 10$, and $\epsilon = 1$. In this case, $F_0 = 0$. Imagine that there are five users with the true values $8, 7, 6, 4.95, 4.9$, respectively. If everyone bids truthfully, then $8, 7, 6, 4.95$ will be confirmed, since $F_4 = 4$ and $F_5 = 5$. Notice that the fifth user (with the true value 4.9) is unconfirmed, so its utility is zero. However, if the fifth user bids 4.96 instead, its bid will be confirmed, and 4.95 will be unconfirmed. The fifth user pays $F_4 = 4$, and gets the utility $4.9 - 4 = 0.9$. Notice that the number of the confirmed bids does not change, so the miner is always paid 4ϵ. Thus, if the miner colludes with the fifth user, their utility increases by 0.9. One can easily modify the true values so that the strategic gain is arbitrarily close to ϵ.

The following theorem shows that a strategic user or miner-user coalition cannot gain more than ϵ.

Theorem B.1. *The staircase mechanism above satisfies ϵ-UIC, strict-MIC, and ϵ-SCP when the miner colludes with at most* 1 *user.*

Proof. We prove the three incentive compatibility properties separately.

ϵ-UIC. Let v_i be user i's true value. Without loss of generality, we assume a strategic user i always first injects some fake bids, and then changes its true bid (not the fake bids) from the true value to some other value. We will show that user i's utility does not increase in either step. Consequently, user i's utility can never increase even if it plays strategically.

First, we show that regardless of the current bid vector. If user i injects one more fake bid, its utility does not increase. Suppose $\mathbf{b} = (b_1, \ldots, b_N)$ is the current bid vector, where some bids might be fake bids injected by user i, and $b_i = v_i$ is user i's true bid. If b_i is already confirmed, i.e. $x_i(\mathbf{b}_{-i}, v_i) = 1$, injecting another fake bid can never lower t. Thus, user i's payment can never be lower after injecting another fake bid. On the other hand, if b_i is unconfirmed, i.e. $x_i(\mathbf{b}_{-i}, v_i) = 0$, b_i must still be unconfirmed after injecting another fake bid. Thus, injecting fake bids does not increase user i's utility.

Second, we show that no matter what the current bid vector is, if user i changes its true bid from the true value to some other value, its utility does not increase. Suppose $\mathbf{b} = (b_1, \ldots, b_N)$ is the current bid vector, where some of the bids might be the fake bids injected by user i, and $b_i = v_i$ is user i's true bid. Let $t^* = \max_i\{i : b_i \geq F_i\}$. There are two cases.

- **Case 1: b_i is confirmed under the bid vector b.** Notice that the payment never exceeds the bid, so user i's utility is always non-negative when user i bids truthfully. Thus, if user i's bid becomes unconfirmed after changing the bid, user i's utility does not increase. On the other hand, if user i's bid is still confirmed after changing the bid, the number of confirmed bids in the block is still t^* because changing the bid only permutes the order of the top t^* bids. Thus, user i's payment is still F_{t^*}, so user i's utility does not change.

- **Case 2: b_i is unconfirmed when the bid vector is b.** If user i underbids, its bid must still be unconfirmed. If user i overbids, the number of the confirmed bids must be at least t^*, so the payment for each confirmed bid is at least F_{t^*}. Because $x_i(\mathbf{b}_{-i}, v_i) = 0$, it must be $v_i \leq F_{t^*+1} = F_{t^*} + \epsilon$. Thus, if user i's bid becomes confirmed because of overbidding, user i's utility is at most $v_i - F_{t^*} \leq \epsilon$.

Strict-MIC. Without loss of generality, we assume the strategic miner prepares the block in the following order: the miner chooses a subset of the bids $\mathbf{c}' = (c'_1, \ldots, c'_{\ell'})$ from the bid vector (not necessarily the top k) where $c'_1 \geq \cdots \geq c'_{\ell'}$; then, the miner injects some fake bids into $\mathbf{c}'$. We will show that the miner's utility does not increase in either step.

First, let $\mathbf{c} = (c_1, \ldots, c_k)$ denote the top k bids in the current bid vector (if the number of bids is less than k, append zeros), where $c_1 \geq c_2 \geq \cdots \geq c_k$. Because $c_i \geq c'_i$ for all i, the number of the confirmed bids in $\mathbf{c}'$ cannot be more than the number of the confirmed bids in $\mathbf{c}$. Thus, not choosing the top k bids into the block never increases miner's revenue.

Second, let $\mathbf{d} = (d_1, \ldots, d_r)$ denote the bid vector that the miner prepares, where $d_1 \geq d_2 \geq \cdots \geq d_r$ for some r. Here, $\mathbf{d}$ may or may not contain fake bids injected by the miner. We will show that if the miner injects one more fake bid f, its utility does not increase. Let $t^* = \max_i\{i : d_i \geq F_i\}$. In this case, it must be $d_{t^*+1} < F_{t^*+1}$. To increase the miner's utility, the number of the confirmed bids after injecting f must increase, so we assume it is the case. Because $d_{t^*+1} < F_{t^*+1}$, if the number of the confirmed bids increases, it must be that f is confirmed and $f \geq F_{t^*+1}$. Moreover, because the miner only injects one more fake bid to $\mathbf{d}$, the number of the confirmed bids after injecting the fake bid is at most $t^* + 1$. Thus, the revenue that the miner gets increases by at most ϵ. The extra cost for injecting f is $F_{t^*+1} \geq \epsilon$ for any $t \geq 0$. Therefore, the overall utility does not increase.

ϵ-SCP. Let user j be the colluding user. Let $\mathbf{d} = (d_1, \ldots, d_k)$ denote the top k bids that the miner includes if both the miner and user j are honest, where $d_1 \geq \cdots \geq d_k$. Let $t = \max_i\{i : d_i \geq F_i\}$; that is, if the miner is honest, t bids will be confirmed. Next, suppose the coalition strategically includes the bids $\mathbf{d}' = (d'_1, \ldots, d'_r)$ for some r, where $d'_1 \geq d'_2 \geq \ldots \geq d'_r$. Let $t' = \max_i\{i : d'_i \geq F_i\}$.

First, to increase the joint utility of the coalition, user j's bid must be confirmed when the block is $\mathbf{d}'$ — if user j's bid is not included under $\mathbf{d}'$, then by strict-MIC, the miner's utility cannot increase when it chooses $\mathbf{d}'$ to be the block, and obviously user j's utility cannot increase either if it is not confirmed under $\mathbf{d}'$. Henceforth, we assume user j's bid is confirmed when the block is $\mathbf{d}'$. There are two possible cases.

- **Case 1: User j's bid is confirmed if the block is d.** In this case, user j's bid is confirmed under both $\mathbf{d}$ and $\mathbf{d}'$, so the change of user j's utility only depends on its payment.

The payment changes from F_t to $F_{t'}$, so user j's utility increases by $F_t - F_{t'} = (t - t')\epsilon$ — if $t - t'$ is negative, user j's utility actually decreases. On the other hand, the miner's revenue decreases by $(t - t')\epsilon$. Therefore, the increase in user j's utility cancels out the decrease in the miner's revenue, and their joint utility does not change.

- **Case 2: User j's bid is unconfirmed if the block is $\mathbf{d}$.** In this case, user j's true value v_j must be smaller than F_{t+1}. Since user j's bid is unconfirmed when the block is $\mathbf{d}$, its utility is zero. Since user j's bid is confirmed when the block is $\mathbf{d}'$, its utility now becomes $v_j - F_{t'} < F_{t+1} - F_{t'}$. Thus, user j's utility increases by $v_j - F_{t'} < F_{t+1} - F_{t'} = (t + 1 - t')\epsilon$. On the other hand, the miner's revenue decreases by $(t - t')\epsilon$. Therefore, the joint utility increases by at most ϵ.

□

C Deferred Proofs of Section 5

C.1 Strict Incentive Compatibility in MPC-Assisted Model: Necessity of Zero Miner Revenue

Chung and Shi [CS21] showed that the posted-price auction with burning gives strict incentive compatibility in the plain model, assuming infinite block size. One may hope that with the Bayesian notion of incentive compatibility, we can achieve larger miner revenue. Unfortunately, in this section we show that zero-miner revenue is the best we can hope for strict incentive compatibility, even in the Bayesian setting.

Prelimary: Myerson's lemma for the Bayesian setting. We first review the Bayesian version of Myerson's lemma. Recall that $\mathbf{b}_{-i}$ denotes all but user i's bid, and $(\mathbf{b}_{-i}, b_i) = \mathbf{b}$. We also let $\mathcal{D}_{-i}$ to denote $\mathcal{D}_1 \times \cdots \times \mathcal{D}_{i-1} \times \mathcal{D}_{i+1} \times \cdots \times \mathcal{D}_n$, which denotes the distribution of other users' true values.

Lemma C.1 (Myerson's Lemma [Mye81])**.** *Let $\mathcal{D} = \mathcal{D}_1 \times \cdots \times \mathcal{D}_n$ be the joint distribution of users' true values. Let $(\mathbf{x}, \mathbf{p}, \mu)$ be a single-parameter TFM that is Bayesian UIC. Then, it must be that*

1. *The allocation rule $\mathbf{x}$ is monotonically non-decreasing. Formally, for any user i, and any $b_i' > b_i$, it must be that $\mathbf{E}_{\mathbf{b}_{-i} \sim \mathcal{D}_{-i}}[x_i(\mathbf{b}_{-i}, b_i')] \geq \mathbf{E}_{\mathbf{b}_{-i} \sim \mathcal{D}_{-i}}[x_i(\mathbf{b}_{-i}, b_i)]$.*
2. *The payment rule $\mathbf{p}$ is defined as follows. For any user i, and bid b_i from user i, it must be*

$$\mathbf{E}_{\mathbf{b}_{-i} \sim \mathcal{D}_{-i}}[p_i(\mathbf{b}_{-i}, b_i)] = \mathbf{E}_{\mathbf{b}_{-i} \sim \mathcal{D}_{-i}}\left[b_i \cdot x_i(\mathbf{b}_{-i}, b_i) - \int_0^{b_i} x_i(\mathbf{b}_{-i}, t)dt\right]. \tag{18}$$

Lemma C.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*

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

Necessity of zero miner revenue. Henceforth we use the following simplified notation.

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

The following technical lemma was given in [CS21].

Lemma C.3 (Lemma 4.8 in [CS21])**.** *Let $(\mathbf{x},\mathbf{p},\mu)$ be any (possibly randomized) TFM in the Bayesian setting. If $(\mathbf{x},\mathbf{p},\mu)$ is Bayesian SCP against a $(\rho,1)$-sized coalition, then for any bid vector $\mathbf{b}$, user i, and r,r' such that $r<r'$, it must be*

$$r\cdot\left(\overline{x_i}(r')-\overline{x_i}(r)\right)\le\pi(r')-\pi(r)\le r'\cdot\left(\overline{x_i}(r')-\overline{x_i}(r)\right),$$

where $\pi(r):=\overline{p_i}(r)-\rho\overline{\mu_i}(r)$.

The following result shows that if we allow the strategic players to inject fake bids, then the miner's revenue can only be 0 if the mechanism is UIC, MIC, and 1-SCP. Actually, in the proof of the lower bound, we only need the deviation where the miners in the coalition injecting fake bids, and colluding users only bid untruthfully.

We first show that if a TFM is Bayesian UIC and Bayesian SCP against $(\rho,1)$-sized coalition, then the miner revenue must be independent from each user's bid. Without loss of generality, we assume that 0 is the minimum value in the support of $\mathcal{D}_i$ for $i\in[n]$.

Lemma C.4. *Let $\mathcal{D}=\mathcal{D}_1\times\cdots\times\mathcal{D}_n$ be the joint distribution of users' true values. Let $(\mathbf{x},\mathbf{p},\mu)$ be any (possibly randomized) TFM in the MPC model. If $(\mathbf{x},\mathbf{p},\mu)$ is Bayesian UIC and Bayesian SCP against a $(\rho,1)$-sized miner-user coalition, then for any user i, any bid b, it must be*

$$\overline{\mu_i}(b)=\overline{\mu_i}(0). \tag{19}$$

In other words, the miner's revenue is a constant that is independent of user i's bid b when other bids $\mathbf{b}_{-i}$ are drawn from the distribution $\mathcal{D}_{-i}$.

Proof. Define $\widetilde{\pi}(r)$ as

$$\widetilde{\pi}(r)=\overline{p_i}(r)-\rho\overline{\mu_i}(r)-(\overline{p_i}(0)-\rho\overline{\mu_i}(0)).$$

By Lemma C.3, and the fact that definition of $\widetilde{\pi}(r)$ and $\pi(r)$ differs by only a fixed constant, it must be that

$$r\cdot\left(\overline{x_i}(r')-\overline{x_i}(r)\right)\le\widetilde{\pi}(r')-\widetilde{\pi}(r)\le r'\cdot\left(\overline{x_i}(r')-\overline{x_i}(r)\right). \tag{20}$$

Therefore, we have the following two inequalities:

$$r\cdot[\overline{x_i}(r')-\overline{x_i}(r)]\le\widetilde{\pi}(r')-\widetilde{\pi}(r)$$
$$r\cdot[\overline{x_i}(r')-\overline{x_i}(r)]\ge\widetilde{\pi}(r')-\widetilde{\pi}(r)$$

Now, observe that the above expression strictly agrees with the "payment sandwich" in the proof of Myerson's Lemma [Mye81, Har]. Furthermore, we have that $\widetilde{\pi}(0)=0$ by definition; and $\mathbf{x}$ must be monotone because the TFM is UIC and satisfies Myerson's Lemma. Due to Lemma C.2, it must be that $\widetilde{\pi}(\cdot)$ obeys the unique payment rule specified by Myerson's Lemma; that is,

$$\widetilde{\pi}(r)=\left[b_i\cdot\overline{x_i}(b_i)-\int_0^{b_i}\overline{x_i}(t)dt\right].$$

On the other hand, since the TFM is UIC, its payment rule itself must also satisfy the same expression (Eq.(18)), that is,

$$\overline{p_i}(b_i) = b_i \cdot \overline{x_i}(b_i) - \int_0^{b_i} \overline{x_i}(t)dt.$$

We therefore have that

$$\widetilde{\pi}(r) = \overline{p_i}(b_i).$$

In other words, $\rho\overline{\mu_i}(r) = \rho\overline{\mu_i}(0) - \overline{p_i}(0)$. Because $\overline{p_i}(0) = 0$, we conclude $\overline{\mu_i}(r) = \overline{\mu_i}(0)$. □

Note that the result in Lemma C.4 holds even if users do not inject any fake bids. This provides a stronger impossibility result.

Now we show that, if in addition the mechanism $(\mathbf{x}, \mathbf{p}, \mu)$ is Bayesian MIC, then the total miner revenue can only be 0.

Theorem C.5. *Let $\{\mathcal{D}^{(n)}\}_n$ be a sequence of distributions where $\mathcal{D}^{(n)} = \mathcal{D}_1 \times \cdots \times \mathcal{D}_n$ is the joint distribution of n users' true values, where user i's true value is drawn from $\mathcal{D}_i$ independently. Let $(\mathbf{x}, \mathbf{p}, \mu)$ be any (possibly randomized) TFM in the MPC model. If $(\mathbf{x}, \mathbf{p}, \mu)$ is Bayesian UIC, Bayesian MIC against ρ-sized miner coalition and Bayesian SCP against $(\rho, 1)$-sized miner-user coalition, then*

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

Proof. For any $n \geq 2$, we have the following claim:

Lemma C.6. *If $(\mathbf{x}, \mathbf{p}, \mu)$ is Bayesian MIC against $(\rho, 1)$-sized miner-user coalition, then* $\mathop{\mathbf{E}}_{\mathbf{b}\sim\mathcal{D}^{(n)}}[\mu(\mathbf{b})] \leq \mathop{\mathbf{E}}_{\mathbf{b}'\sim\mathcal{D}^{(n-1)}}[\mu(\mathbf{b}')]$.

For now assume Lemma C.6 holds and we explain why Theorem C.5 follows from it. The proof of Lemma C.6 appears right afterwards. By induction on n, we have that

$$\mathop{\mathbf{E}}_{\mathbf{b}\sim\mathcal{D}^{(n)}}[\mu(\mathbf{b})] \leq \mathop{\mathbf{E}}_{b\sim\mathcal{D}_1}[\mu(b)].$$

By Lemma C.4, for any $b \in \mathrm{Supp}(\mathcal{D}_1)$, it should be that $\mu(b) = \mu(0)$. Therefore,

$$\mathop{\mathbf{E}}_{b\sim\mathcal{D}_1}[\mu(b)] \leq \mu(0) = 0,$$

where the last equality comes from the requirement that the miner's revenue cannot exceeds the payment of the single identity, who will pay at most what it bids. Theorem C.5 thus follows. □

Proof of Lemma C.6 Since $(\mathbf{x}, \mathbf{p}, \mu)$ is Bayesian-SCP against $(\rho, 1)$-sized coalition, it must be that for any user i,

$$\mathop{\mathbf{E}}_{\mathbf{b}\sim\mathcal{D}^{(n-1)}}[\rho\mu(\mathbf{b}, 0)] \leq \mathop{\mathbf{E}}_{\mathbf{b}\sim\mathcal{D}^{(n-1)}}[\rho\mu(\mathbf{b})]. \tag{21}$$

Otherwise, the miners can collude with user i, ask user i to bid, and inject 0 and increase the coalition's miner revenue while it does not need to pay anything for injecting the 0-bid. This violates the MIC condition.

By the law of total expectation, we have that

$$\begin{aligned}
\mathop{\mathbf{E}}_{\mathbf{b}\sim\mathcal{D}^{(n)}}[\mu(\mathbf{b})] &= \int_0^{+\infty} \mathop{\mathbf{E}}_{\mathbf{b}'\sim\mathcal{D}^{(n-1)}}[\mu(\mathbf{b}', r)]f(r)dr \\
&= \int_0^{+\infty} \mathop{\mathbf{E}}_{\mathbf{b}'\sim\mathcal{D}^{(n-1)}}[\mu(\mathbf{b}', 0)]f(r)dr && \text{By Lemma C.4} \\
&= \mathop{\mathbf{E}}_{\mathbf{b}'\sim\mathcal{D}^{(n-1)}}[\mu(\mathbf{b}', 0)] \le \mathop{\mathbf{E}}_{\mathbf{b}'\sim\mathcal{D}^{(n-1)}}[\mu(\mathbf{b}')] && \text{By (21)}
\end{aligned}$$

Lemma C.6 thus follows.

C.2 Proof of Lemma 5.2

Lemma C.7 (Restatement of Lemma 5.2)**.** *Let $(\mathbf{x}, \mathbf{p}, \mu)$ be any (possibly random) mechanism that is Bayesian UIC and Bayesian SCP against $(\rho^*, 2)$-sized coalition for some $\rho^* \in (0, 1]$. Suppose each user's true value is drawn i.i.d. from a distribution $\mathcal{D}$. Then for any user i and j, for any bid b_j and b'_j, it must be that for any $\ell \ge 1$,*

$$\mathop{\mathbf{E}}_{v,\mathbf{b}\sim\mathcal{D}^\ell}[\mathsf{util}^i(v, b_j, \mathbf{b})] = \mathop{\mathbf{E}}_{v,\mathbf{b}\sim\mathcal{D}^\ell}[\mathsf{util}^i(v, b'_j, \mathbf{b})].$$

Proof. In this proof, we use the following notations for simplicity. For any fixed $\ell \ge 1$, for any user i and j, we define the following notations:

$$\overline{x_i}(\cdot,\cdot) = \mathop{\mathbf{E}}_{\mathbf{b}_{-i,j}\sim\mathcal{D}^{\ell-1}}[x_i(\cdot,\cdot,\mathbf{b})], \quad \overline{p_i}(\cdot,\cdot) = \mathop{\mathbf{E}}_{\mathbf{b}_{-i,j}\sim\mathcal{D}^{\ell-1}}[p_i(\cdot,\cdot,\mathbf{b})], \quad \overline{\mu}(\cdot,\cdot) = \mathop{\mathbf{E}}_{\mathbf{b}_{-i,j}\sim\mathcal{D}^{\ell-1}}[\mu(\cdot,\cdot,\mathbf{b})].$$

Imagine that user i has true value v and user j has true value y. Then for any feasible $\rho \le \rho^*$, it must be that

$$\begin{aligned}
&\text{Honest utility} = [v \cdot \overline{x_i}(v,y) - \overline{p_i}(v,y)] + [y \cdot \overline{x_j}(v,y) - \overline{p_j}(v,y)] + \rho\overline{\mu}(v,y) \\
\ge&\text{Overbid utility} = [v \cdot \overline{x_i}(v,z) - \overline{p_i}(v,z)] + [y \cdot \overline{x_j}(v,z) - \overline{p_j}(v,z)] + \rho\overline{\mu}(v,z).
\end{aligned}$$

Otherwise, the miner can collude with user i with true value v and user j with true value y and ask user j to overbid to some $z \ge y$. This will increase the joint utility of the coalition, which contradicts $(\rho^*, 2)$-SCP. For the same reason, if user j's true value is z, then

$$\begin{aligned}
&\text{Honest utility} = [v \cdot \overline{x_i}(v,z) - \overline{p_i}(v,z)] + [z \cdot \overline{x_j}(v,z) - \overline{p_j}(v,z)] + \rho\overline{\mu}(v,z) \\
\ge&\text{Underbid utility} = [v \cdot \overline{x_i}(v,y) - \overline{p_i}(v,y)] + [z \cdot \overline{x_j}(v,y) - \overline{p_j}(v,y)] + \rho\overline{\mu}(v,y).
\end{aligned}$$

Combining these two inequalities together, we get the following payment difference sandwich. For any $z \ge y$, we have

$$\begin{aligned}
&v[\overline{x_i}(v,z) - \overline{x_i}(v,y)] + z[\overline{x_j}(v,z) - \overline{x_j}(v,y)] + \rho[\overline{\mu}(v,z) - \overline{\mu}(v,y)] \\
\ge&\overline{p_i}(v,z) - \overline{p_i}(v,y) + \overline{p_j}(v,z) - \overline{p_j}(v,y) \\
\ge&v[\overline{x_i}(v,z) - \overline{x_i}(v,y)] + y[\overline{x_j}(v,z) - \overline{x_j}(v,y)] + \rho[\overline{\mu}(v,z) - \overline{\mu}(v,y)]
\end{aligned}$$

Divide the inequality with $z - y$ and take limit $y \to z$, we get

$$v \cdot \frac{\partial}{\partial z}\overline{x_i}(v,z) + z \cdot \frac{\partial}{\partial z}\overline{x_j}(v,z) + \rho\frac{\partial}{\partial z}\overline{\mu}(v,z) = \frac{\partial}{\partial z}\overline{p_i}(v,z) + \frac{\partial}{\partial z}\overline{p_j}(v,z). \tag{22}$$

Note that Equation (22) should hold for at least two different values of $\rho \le \rho^*$. Hence, it must be that $\frac{\partial}{\partial z}\overline{\mu}(v,z) = 0$. Equation (22) thus becomes

$$\begin{aligned} & v[\overline{x_i}(v,z) - \overline{x_i}(v,y)] + z[\overline{x_j}(v,z) - \overline{x_j}(v,y)] \\ \ge & \overline{p_i}(v,z) - \overline{p_i}(v,y) + \overline{p_j}(v,z) - \overline{p_j}(v,y) \\ \ge & v[\overline{x_i}(v,z) - \overline{x_i}(v,y)] + y[\overline{x_j}(v,z) - \overline{x_j}(v,y)]. \end{aligned} \tag{23}$$

This is equivalent to say: when user j changes its bid, the joint utility of user i and user j should not increase. That means, for any v, if a user j with true value y changes its bid from y to z, it must be that

$$\begin{aligned} \text{i-gain}(v, y \to z) &:= \mathop{\mathbf{E}}_{\mathbf{b}_{-i,j}\sim\mathcal{D}^{\ell-1}} \mathsf{util}^i(v, z, \mathbf{b}_{-i,j}) - \mathop{\mathbf{E}}_{\mathbf{b}_{-i,j}\sim\mathcal{D}^{\ell-1}} \mathsf{util}^i(v, y, \mathbf{b}_{-i,j}) \\ &\le \mathop{\mathbf{E}}_{\mathbf{b}_{-i,j}\sim\mathcal{D}^{\ell-1}} \mathsf{util}^j(v, y, \mathbf{b}_{-i,j}) - \mathop{\mathbf{E}}_{\mathbf{b}_{-i,j}\sim\mathcal{D}^{\ell-1}} \mathsf{util}^j(v, z, \mathbf{b}_{-i,j}) := \text{j-loss}(v, y \to z) \end{aligned}$$

Since the mechanism is UIC, by the same proof as of [GH05], we get:

$$\begin{aligned} \mathop{\mathbf{E}}_{v\sim\mathcal{D}}[\text{i-gain}(v, y \to z)] &\le \mathop{\mathbf{E}}_{v\sim\mathcal{D}}[\text{j-loss}(v, y \to z)] \\ &\le \mathop{\mathbf{E}}_{v\sim\mathcal{D}}[(z-y)(\overline{x_j}(v,z) - \overline{x_j}(v,y))]. \end{aligned}$$

Now consider the situation where user j changes its bid from b_j to b'_j. Without loss of generality, we assume that $b'_j \ge b_j$. If we divide the interval between $[b_j, b'_j]$ into L equally sized segments $b_j^{(0)}, \ldots, b_j^{(L)}$, then the total gain for user i can be bounded by

$$\begin{aligned} \mathop{\mathbf{E}}_{v\sim\mathcal{D}}[\text{i-gain}(v, b_j \to b'_j)] &= \sum_{l=0}^{L-1} \mathop{\mathbf{E}}_{v\sim\mathcal{D}}[\text{i-gain}(v, b_j^{(l)} \to b_j^{(l+1)})] \\ &\le \sum_{l=0}^{L-1} (b_j^{(l+1)} - b_j^{(l)}) \mathop{\mathbf{E}}_{v\sim\mathcal{D}} \left[\overline{x_j}(v, b_j^{(l+1)}) - \overline{x_j}(v, b_j^{(l)})\right] \\ &= \frac{b'_j - b_j}{L} \mathop{\mathbf{E}}_{v\sim\mathcal{D}} \left[\overline{x_j}(v, b'_j) - \overline{x_j}(v, b_j)\right]. \end{aligned}$$

This holds for any L. Taking limit for $L \to \infty$, we have that

$$\mathop{\mathbf{E}}_{v\sim\mathcal{D}}[\text{i-gain}(v, b_j \to b'_j)] \le 0.$$

Since $\mathop{\mathbf{E}}_{v\sim\mathcal{D}}[\text{i-gain}(v, b_j \to b'_j)] = -\mathop{\mathbf{E}}_{v\sim\mathcal{D}}[\text{i-gain}(v, b'_j \to b_j)]$, we have that $\mathop{\mathbf{E}}_{v\sim\mathcal{D}}[\text{i-gain}(v, b_j \to b'_j)] = 0$, for arbitrary b_j and b'_j. The lemma thus follows. □

C.3 Full Proof of Lemma 5.3

In this section, we provide a full proof of Lemma 5.3 assuming the symmetry assumption in Section 2.1.

By Lemma 5.2, for any $i, j, b_j, \ell \ge 1$, we have

$$\mathop{\mathbf{E}}_{v_i, \mathbf{b}_{-i,j}\sim\mathcal{D}^\ell}[\mathsf{util}^i(v_i, b_j, \mathbf{b}_{-i,j})] = \mathop{\mathbf{E}}_{v_i, \mathbf{b}_{-i,j}\sim\mathcal{D}^\ell}[\mathsf{util}^i(v_i, 0_j, \mathbf{b}_{-i,j})].$$

Therefore, it suffices to prove that for any i, j, ℓ, $\mathop{\mathbf{E}}_{v_i, \mathbf{b}_{-i,j} \sim \mathcal{D}^\ell}[\mathsf{util}^i(v_i, 0_j, \mathbf{b}_{-i,j})] \leq \mathop{\mathbf{E}}_{v_i, \mathbf{b}_{-i,j} \sim \mathcal{D}^\ell}[\mathsf{util}^i(v_i, \mathbf{b}_{-i,j})]$.

Suppose for the sake of contradiction, the above statement is not true, that is, there exist some i, j, ℓ, such that $\mathop{\mathbf{E}}_{v_i, \mathbf{b}_{-i,j} \sim \mathcal{D}^\ell}[\mathsf{util}^i(v_i, 0_j, \mathbf{b}_{-i,j})] > \mathop{\mathbf{E}}_{v_i, \mathbf{b}_{-i,j} \sim \mathcal{D}^\ell}[\mathsf{util}^i(v_i, \mathbf{b}_{-i,j})]$. Then there must exists a v_i, such that $\mathop{\mathbf{E}}_{\mathbf{b}_{-i,j} \sim \mathcal{D}^{\ell-1}}[\mathsf{util}^i(v_i, 0_j, \mathbf{b}_{-i,j})] > \mathop{\mathbf{E}}_{\mathbf{b}_{-i,j} \sim \mathcal{D}^{\ell-1}}[\mathsf{util}^i(v_i, \mathbf{b}_{-i,j})]$.

Consider an arbitrary fake identity m registered by the miner. There are two possible cases.

Good identity m: $\mathop{\mathbf{E}}_{\mathbf{b}_{-i,j} \sim \mathcal{D}^{\ell-1}}[\mathsf{util}^i(v_i, 0_m, \mathbf{b}_{-i,j})] > \mathop{\mathbf{E}}_{\mathbf{b}_{-i,j} \sim \mathcal{D}^{\ell-1}}[\mathsf{util}^i(v_i, \mathbf{b}_{-i,j})]$.

Bad identity m: $\mathop{\mathbf{E}}_{\mathbf{b}_{-i,j} \sim \mathcal{D}^{\ell-1}}[\mathsf{util}^i(v_i, 0_m, \mathbf{b}_{-i,j})] \leq \mathop{\mathbf{E}}_{\mathbf{b}_{-i,j} \sim \mathcal{D}^{\ell-1}}[\mathsf{util}^i(v_i, \mathbf{b}_{-i,j})] < \mathop{\mathbf{E}}_{\mathbf{b}_{-i,j} \sim \mathcal{D}^{\ell-1}}[\mathsf{util}^i(v_i, 0_j, \mathbf{b}_{-i,j})]$.

Now, suppose the miner samples a fake identity m. Over the choice of m, either Pr(Good identity m) $\geq \frac{1}{2}$ or Pr(Bad identity m) $\geq \frac{1}{2}$. If Pr(Good identity m) $\geq \frac{1}{2}$, then suppose that the world consists of ℓ users not including j, and the miner forms a coalition with user i whose true value is v_i. The miner can sample a random identity m, and if it is a good identity, the miner can inject a fake bid 0_m, and the coalition can strictly gain. This violates SCP when $c = 1$.

Henceforth, we focus on the case when Pr(Bad identity m) $\geq \frac{1}{2}$. In this case, there are two possibilities, either with probability at least $1/4$ over the choice of the identity m, for all v_i',

$$\mathop{\mathbf{E}}_{\mathbf{b}_{-i,j} \sim \mathcal{D}^{\ell-1}}[\mathsf{util}^i(v_i', 0_m, \mathbf{b}_{-i,j})] \leq \mathop{\mathbf{E}}_{\mathbf{b}_{-i,j} \sim \mathcal{D}^{\ell-1}}[\mathsf{util}^i(v_i', 0_j, \mathbf{b}_{-i,j})], \tag{24}$$

or with probability at least $1/4$ over the choice of m, there exists some v_i' such that $\mathop{\mathbf{E}}_{\mathbf{b}_{-i,j} \sim \mathcal{D}^{\ell-1}}[\mathsf{util}^i(v_i', 0_m, \mathbf{b}_{-i,j})] > \mathop{\mathbf{E}}_{\mathbf{b}_{-i,j} \sim \mathcal{D}^{\ell-1}}[\mathsf{util}^i(v_i', 0_j, \mathbf{b}_{-i,j})]$. If it is the latter case, then, consider a scenario where the miner colludes with user i whose true value is v_i', and user j whose true value is 0, and the rest of the world is a random variable $\mathbf{b}_{-i,j}$. Now, the miner can sample a random fake identity m, and see if dropping 0_j and injecting 0_m can help its friend i. If so, it performs this strategic behavior. This strategy can strictly help the coalition which violates SCP for $c = 2$.

It suffices to rule out the former case, that is, with probability at least $1/4$ over the choice of the identity m, for all v_i', Equation (24) is satisfied. Recall also, for v_i specifically, we have strict inequality, that is, $\mathop{\mathbf{E}}_{\mathbf{b}_{-i,j} \sim \mathcal{D}^{\ell-1}}[\mathsf{util}^i(v_i, 0_m, \mathbf{b}_{-i,j})] < \mathop{\mathbf{E}}_{\mathbf{b}_{-i,j} \sim \mathcal{D}^{\ell-1}}[\mathsf{util}^i(v_i, 0_j, \mathbf{b}_{-i,j})]$. Thus, $\mathop{\mathbf{E}}_{\mathbf{b}_{-j} \sim \mathcal{D}^{\ell}}[\mathsf{util}^i(0_m, \mathbf{b}_{-j})] < \mathop{\mathbf{E}}_{\mathbf{b}_{-j} \sim \mathcal{D}^{\ell}}[\mathsf{util}^i(0_j, \mathbf{b}_{-j})]$.

For every bad identity m that additionally satisfies Equation (24), there must exist some $i' \neq i$ and $i \neq j$, and some $b_{i'} > 0$, such that

$$\mathop{\mathbf{E}}_{\mathbf{b}_{-j,i'} \sim \mathcal{D}^{\ell-1}}[\mathsf{util}^{i'}(0_m, b_{i'}, \mathbf{b}_{-j,i'})] > \mathop{\mathbf{E}}_{\mathbf{b}_{-j,i'} \sim \mathcal{D}^{\ell-1}}[\mathsf{util}^{i'}(0_j, b_{i'}, \mathbf{b}_{-j,i'})] \tag{25}$$

We can prove the above claim by contradiction. Suppose for the sake of contradiction that for all $i' \neq i$ and $i \neq j$, and for all $b_{i'}$, $\mathop{\mathbf{E}}_{\mathbf{b}_{j,i'} \sim \mathcal{D}^{\ell-1}}[\mathsf{util}^{i'}(0_m, b_{i'}, \mathbf{b}_{-j,i'})] \leq \mathop{\mathbf{E}}_{\mathbf{b}_{j,i'} \sim \mathcal{D}^{\ell-1}}[\mathsf{util}^{i'}(0_j, b_{i'}, \mathbf{b}_{-j,i'})]$. Therefore, it must be that for any $i' \neq i$ and $i \neq j$, $\mathop{\mathbf{E}}_{\mathbf{b}_{-j} \sim \mathcal{D}^{\ell}}[\mathsf{util}^{i'}(0_m, \mathbf{b}_{-j})] \leq \mathop{\mathbf{E}}_{\mathbf{b}_{-j} \sim \mathcal{D}^{\ell}}[\mathsf{util}^{i'}(0_j, \mathbf{b}_{-j})]$. Therefore, we have that

$$\mathop{\mathbf{E}}_{\mathbf{b}_{-j} \sim \mathcal{D}^{\ell}}[\mathsf{USW}(0_m, \mathbf{b}_{-j})] < \mathop{\mathbf{E}}_{\mathbf{b}_{-j} \sim \mathcal{D}^{\ell}}[\mathsf{USW}(0_j, \mathbf{b}_{-j})] \tag{26}$$

where $\mathsf{USW}(\mathbf{b})$ denotes the social welfare for all users (i.e., sum of all user utilities) when the bid vector is $\mathbf{b}$. However, by our symmetry assumption in Section 2.1, it must be that $\mathop{\mathbf{E}}_{\mathbf{b}_{-j}\sim\mathcal{D}^\ell}[\mathsf{USW}(0_m, \mathbf{b}_{-j})] = \mathop{\mathbf{E}}_{\mathbf{b}_{-j}\sim\mathcal{D}^\ell}[\mathsf{USW}(0_j, \mathbf{b}_{-j})]$, which contradicts Equation (26).

Let i' be a user such that Equation (25) happens with probability at least $1/4(\ell+1)$ over the choice of m — clearly, such a user must exist since we are assuming that with probability at least $1/4$ over the choice of m, where m is a bad identity satisfying Equation (24). Now, imagine that the world consists of $\ell+1$ users including both i and j, and the miner forms a coalition with users i' and j. The miner samples a random fake identity m, and if the identity helps i' in the sense that Equation (25) holds, then the coalition replaces j's bid 0_j with 0_m. This strategy strictly increases the coalition's joint utility, and this violates SCP when $c = 2$.

D Multi-Party Computation Protocol Realizing $\mathcal{F}_{\text{MPC}}$

So far in the paper, we have assumed that the transaction fee mechanism is implemented by a trusted ideal functionality $\mathcal{F}_{\text{MPC}}$. In this section, we show how to instantiate $\mathcal{F}_{\text{MPC}}$ in the real world with cryptography. The protocol described in this section uses generic MPC. However, as mentioned in Remark 1.4, the MPC-assisted mechanisms described in this paper actually need not employ generic MPC to be instantiated in practice — we describe efficient instantiations for our specific protocols in Appendix E.

Terminology and model. Imagine that there are m *miners* and a set of user *identities*. Since each user can assume multiple identities, henceforth, we often use the term *identities* to refer to the set of purported user identities. We assume that the miners can communicate with each other through a pairwise private channel. Further, every user identity can communicate with every miner through a pairwise private channel. Morever, there is a broadcast channel among the miners and the user identities. We assume that all channels are authenticated, i.e., every message is marked with the true sender. Further, we assume a synchronous model of communication, i.e., the protocol proceeds in rounds and messages sent by honest parties will be received by honest recipients at the beginning of the next round.

We assume that at the beginning of the protocol, the miners have reached consensus on the set of user identities that will participate in the auction. For example, the consensus can be achieved in the following manner: every user identity announces itself to all miners. Then, each of the m miners broadcasts to all miners a candidate set consisting of the identities it has heard. Any identity that appears in the majority of the miners' candidate sets will be permitted into the auction. As long as the majority of the miners are honest, then any honest user identity will be included in the final permitted list.

The parties now execute an interactive protocol at the end of which all parties, including the miners and user identities, learn the outcome of the auction, including which identities' bids are confirmed and how much each confirmed bid pays. In our actual protocol, the user identities need not communicate with each other. Each user identity communicates only with the miners — either it sends a direct message to a miner over the pairwise private channel, or it broadcasts a message which can be seen by all miners.

During the protocol, if a subset of parties (miners or user identities) form a coalition, we assume that the coalition has the advantage of performing a so-called "rushing attack". Specifically, in any round r, parties in the coalition can observe honest parties' messages sent to coalition members or

posted on the broadcast channel, before deciding what messages coalition members want to send in the same round r.

D.1 Building Blocks

We first introduce some building blocks used in the protocol.

D.1.1 Commitment Scheme

A commitment scheme, parametrized by a security parameter λ, and a message space $\{0,1\}^{\ell(\lambda)}$ where $\ell(\cdot)$ is a polynomial in λ, has two phases:

- Commitment phase: the committer who has a message $X \in \{0,1\}^{\ell(\lambda)}$ samples some random coins $r \xleftarrow{\$} \{0,1\}^\lambda$, and computes the commitment $\widehat{X} \leftarrow \mathsf{comm}(X, r)$. It sends the commitment $\widehat{X}$ to the receiver.
- Open phase: The committer sends the pair (X, r) to the receiver. The receiver outputs "accept" if $\mathsf{comm}(X, r) = \widehat{X}$; otherwise, it outputs "reject".

In our protocol, we require that the commitment scheme must satisfy the following two properties.

- **Perfect binding**: for any $X \neq X'$, and for any r, r', it must be that
$$\mathsf{comm}(X, r) \neq \mathsf{comm}(X', r').$$
- **Computationally hiding**: for any X and X', it must be that
$$\{\mathsf{comm}(X, r), r \xleftarrow{\$} \{0,1\}^\lambda\} \equiv_c \{\mathsf{comm}(X', r), r \xleftarrow{\$} \{0,1\}^\lambda\},$$
where $\equiv_c$ denotes computational indistinguishability.

D.1.2 Shamir Secret Sharing

In our final protocol, each user identity will split its bid into m shares, one for each miner, using a t-out-of-m Shamir secret sharing scheme. Henceforth let $\mathbb{F}$ denote some finite field. A t-out-of-m Shamir secret sharing consists of two algorithms, share and $\mathsf{reconstruct}$.

- share takes as an input a secret $s \in \mathbb{F}$, and outputs m shares $(s_1, \ldots, s_m) \in \mathbb{F}^m$ of the secret.
- $\mathsf{reconstruct}$ takes as input a set $I \subseteq [m]$, and the corresponding shares $\{s_i\}_{i \in I}$, and outputs the corresponding secret if and only if $|I| \geq t$. Otherwise, the algorithm returns $\perp$.

A t-out-of-m secret sharing satisfies the following two properties:

- **Correctness:** For any secret s and any set $I \subseteq [m]$ such that $|I| \geq t$, it must be that
$$\Pr[(s_1, \ldots, s_m) \leftarrow \mathsf{share}(s) : \mathsf{reconstruct}(I, \{s_i\}_{i \in I}) = s] = 1.$$
- **Security:** For any two secret s and s', and for all set $I \subseteq [m]$ such that $|I| \leq t - 1$, it must be that
$$\{\{s_i\}_{i \in I} : (s_1, \ldots, s_m) \leftarrow \mathsf{share}(s)\} \equiv \{\{s_i\}_{i \in I} : (s_1, \ldots, s_m) \leftarrow \mathsf{share}(s')\}.$$
where $\equiv$ denotes identically distributed. In addition, Shamir secret sharing also satisfies the following properties. For any set I such that $|I| < t$,
$$\{\{s_i\}_{i \in I} : (s_1, \ldots, s_m) \leftarrow \mathsf{share}(s)\} \equiv \{\{u_i\}_{i \in I} : u_i \text{ uniformly randomly chosen from } \mathbb{F}\}.$$

D.1.3 Honest-Majority Multi-CRS NIZK

In our protocol, user identities will need to rely on zero-knowledge proofs to prove that they have correctly shared their bids. We will use a non-interactive zero-knowledge proof (NIZK). Since we assume the majority of the miners are honest, NIZK can be instantiated without a common reference string (CRS), using an honest-majority multi-CRS NIZK scheme [GO14]. Specifically, every miner $j \in [m]$ acts as a CRS contributor, and posts a CRS denoted crs_j to the broadcast channel. For any miner j' who did not post a CRS, we treat its $\mathsf{crs}_{j'}$ as 0. Given the collection of all CRSes $\{\mathsf{crs}_j\}_{j\in[m]}$, a prover can prove an NP statement given a valid witness. As long as a majority of the miners (i.e., CRS contributors) are honest, the NIZK scheme satisfies completeness, zero-knowledge, and simulation sound extractability, as defined below.

For an NP language L, let $\mathcal{R}_L(\mathsf{stmt}, w)$ denote the NP relation corresponding to the language L, i.e., $\mathsf{stmt} \in L$ if and only if there exists a w such that $\mathcal{R}_L(\mathsf{stmt}, w) = 1$. An honest-majority multi-CRS NIZK with m CRS contributors for an NP language L, parameterized with a security parameter λ, consists of the following algorithms, where part of the definition is taken verbatim from Guo, Pass, and Shi [GPS19].

- $\mathsf{crs} \leftarrow \mathsf{K}(1^\lambda)$: each CRS contributor $j \in [m]$ runs $\mathsf{K}(1^\lambda)$ to generate a CRS crs_j.
- $\tau \leftarrow \mathsf{P}(\{\mathsf{crs}_j\}_{j\in[m]}, \mathsf{stmt}, w)$: given a statement stmt and a witness w such that $\mathcal{R}_L(\mathsf{stmt}, w) = 1$, and the set of all CRSes denoted $\{\mathsf{crs}_j\}_{j\in[m]}$, compute a proof denoted π.
- $\{0, 1\} \leftarrow \mathsf{V}(\{\mathsf{crs}_j\}_{j\in[m]}, \mathsf{stmt}, \pi)$: given a statement stmt, the set of all CRSes $\{\mathsf{crs}_j\}_{j\in[m]}$, and a proof π, the verifier algorithm V outputs either 0 or 1 denoting either reject or accept.
- $(\widetilde{\mathsf{crs}}, \tau) \leftarrow \widetilde{\mathsf{K}}(1^\lambda)$: a simulated CRS generation algorithm that generates a simulated $\widetilde{\mathsf{crs}}$ and a trapdoor τ.
- $\pi \leftarrow \widetilde{\mathsf{P}}(\mathsf{stmt}, \{\widetilde{\mathsf{crs}}_j\}_{j\in[m]}, \{\tau_j\}_{j\in H})$ where $H \subseteq [m]$ and $|H| \geq \lfloor \frac{m}{2} \rfloor + 1$: a simulated prover algorithm produces a proof for the statement stmt without any witness, and the simulated prover has to have access to at least $\lfloor \frac{m}{2} \rfloor + 1$ number of trapdoors.

Henceforth, we use $\mathcal{A}^{\mathcal{O}(\cdot)}(x)$ to mean that $\mathcal{A}$ is given oracle access to the oracle $\mathcal{O}(\cdot)$. Next, we give the security properties we want from the NIZK.

Completeness. Completeness says that an honest prover can always produce a proof that verifies, if it knows a valid witness to the statement. Formally, completeness requires that for every λ, for any set of CRSes $\{\mathsf{crs}_j\}_{j\in[m]}$ where every crs_j is in the support of $\mathsf{K}(1^\lambda)$, for every statement stmt and witness w such that $\mathcal{R}_L(\mathsf{stmt}, w) = 1$, with probability 1, the following holds: let $\pi \leftarrow \mathsf{P}(\{\mathsf{crs}_j\}_{j\in[m]}, \mathsf{stmt}, w)$, it must be that $\mathsf{V}(\{\mathsf{crs}_j\}_{j\in[m]}, \mathsf{stmt}, \pi) = 1$.

Zero-knowledge. An honest-majority multi-CRS NIZK system satisfies zero knowledge iff the following properties hold. First, we require that simulated reference strings are indistinguishable from real ones, i.e., for every non-uniform p.p.t. $\mathcal{A}$, there exists a negligible function $\mathsf{negl}(\cdot)$, such that

$$\left|\Pr\left[\mathsf{crs} \leftarrow \mathsf{K}(1^\lambda) : \mathcal{A}(1^\lambda, \mathsf{crs}) = 1\right] - \Pr\left[(\widetilde{\mathsf{crs}}, \tau) \leftarrow \widetilde{\mathsf{K}}(1^\lambda) : \mathcal{A}(1^\lambda, \widetilde{\mathsf{crs}}) = 1\right]\right| \leq \mathsf{negl}(\lambda).$$

Moreover, we require that as long as the majority of the CRSes are honestly generated, then any efficient adversary cannot distinguish an interaction with a real prover using real witnesses to

prove statements and an interaction with a simulated prover who proves statements without using witnesses — even if $\mathcal{A}$ obtains the trapdoors of the simulated CRSes.

More formally, let $\mathcal{A}^{\widetilde{\mathsf{K}}}$ denote and adversary $\mathcal{A}$ who is allowed to call the simulated key generation algorithm $\widetilde{\mathsf{K}}(1^\lambda)$ multiple times. We say that $\mathcal{A}$ is *minority-constrained*, if among the set of CRSes $\{\mathsf{crs}_j\}_{j\in[m]}$ output by $\mathcal{A}$, the majority of them are CRSes returned to Afrom K. We want that for any non-uniform p.p.t.*minority-constrained* adversary $\mathcal{A}$, there exists a negligible function $\mathsf{negl}(\cdot)$ such that

$$\Big|\Pr\Big[(\{\mathsf{crs}_j\}_{j\in[m]},\mathsf{stmt},w)\leftarrow\mathcal{A}^{\widetilde{\mathsf{K}}}(1^\lambda),\pi\leftarrow\mathsf{P}(\{\mathsf{crs}_j\}_{j\in[m]},\mathsf{stmt},w):\mathcal{A}(\pi)=1\text{ and }\mathcal{R}_L(\mathsf{stmt},w)=1\Big]$$
$$-\Pr\Big[(\{\mathsf{crs}_j\}_{j\in[m]},\mathsf{stmt},w)\leftarrow\mathcal{A}^{\widetilde{\mathsf{K}}}(1^\lambda),\pi\leftarrow\widetilde{\mathsf{P}}(\{\mathsf{crs}_j\}_{j\in[m]},\overrightarrow{\tau},\mathsf{stmt}):\mathcal{A}(\pi)=1\text{ and }\mathcal{R}_L(\mathsf{stmt},w)=1\Big]\Big|$$
$$\leq\mathsf{negl}(\lambda)$$

where $\overrightarrow{\tau}$ is the following vector: for every CRS in the set $\{\mathsf{crs}_j\}_{j\in[m]}$ that is output by the simulated key generation algorithm $\widetilde{\mathsf{K}}$, the vector $\overrightarrow{\tau}$ includes its corresponding trapdoor. Note that there are at least $\lfloor\frac{m}{2}\rfloor+1$ entries in $\overrightarrow{\tau}$ since $\mathcal{A}$ is minority-constrained.

Simulation sound extractability. Intuitively, simulation sound extractability requires that even though an $\mathcal{A}$ may adaptively interact with a simulated prover and obtain simulated proofs of false statements, if $\mathcal{A}$ ever produces a fresh proof for some purposed statement stmt, then except with negligible probability, some p.p.t. extractor must be able to extract a valid witness from the proof, using an extraction key that is produced during a simulated setup procedure.

More formally, an honest-majority multi-CRS NIZK system satisfies simulation sound extractability iff there exist p.p.t. algorithms $\widetilde{\mathsf{K}}_0$ and $\mathcal{E}$ such that the following is satisfied:

- $\widetilde{\mathsf{K}}_0(1^\lambda)$ outputs a triple denoted $(\widetilde{\mathsf{crs}},\tau,\mathsf{ek})$ where the first two terms have an output distribution identical to that of $\widetilde{\mathsf{K}}(1^\lambda)$; and
- for any non-uniform p.p.t.*minority-constrained adversary* $\mathcal{A}$, there exists a negligible function $\mathsf{negl}(\cdot)$, such that the following holds:

$$\Pr\left[\begin{array}{ll}\big(\{\mathsf{crs}_j\}_{j\in[m]},\mathsf{stmt},\pi\big)\leftarrow\mathcal{A}^{\mathcal{O}(1^\lambda,\cdot)}: & (\mathsf{stmt},\pi)\text{ not output from }\mathcal{O}(1^\lambda,\cdot),\text{ and}\\ w\leftarrow\mathcal{E}(\{\mathsf{crs}_j\}_{j\in[m]},\overrightarrow{\mathsf{ek}},\mathsf{stmt},\pi) & \mathcal{R}_L(\mathsf{stmt},w)=0\text{ but }\mathsf{V}(\{\mathsf{crs}_j\}_{j\in[m]},\mathsf{stmt},\pi)=1\end{array}\right]$$
$$\leq\mathsf{negl}(\lambda),$$

where $\mathcal{O}(1^\lambda,\cdot)$ is the following oracle:

1. Upon receiving crs generation query gen from $\mathcal{A}$, it runs $(\mathsf{crs},\tau,\mathsf{ek})\leftarrow\widetilde{\mathsf{K}}_0(1^\lambda)$; it then records τ and returns crs and ek to $\mathcal{A}$.
2. Then, at some point, $\mathcal{A}$ outputs $\{\mathsf{crs}_j\}_{j\in[m]}$ — this set of CRSes must be consistent with the CRSes in $\mathcal{A}$'s final output. $\mathcal{A}$ is required to be minority-constrained, meaning that at least $\lfloor\frac{m}{2}\rfloor+1$ number of entries in $\{\mathsf{crs}_j\}_{j\in[m]}$ must be output from $\widetilde{\mathsf{K}}_0$.
   At this moment, define the following notation:
   - $\overrightarrow{\tau}$ is the following vector: for every CRS in the set $\{\mathsf{crs}_j\}_{j\in[m]}$ that is output by $\widetilde{\mathsf{K}}_0$, the vector $\overrightarrow{\tau}$ includes its corresponding trapdoor. Note that $\overrightarrow{\tau}$ must contain at least $\lfloor\frac{m}{2}\rfloor+1$ such trapdoors since $\mathcal{A}$ is minority-constrained.

– Similarly, the notation $\overrightarrow{\mathsf{ek}}$ denotes the following vector: for every CRS in the set $\{\mathsf{crs}_j\}_{j\in[m]}$ that is output by $\widetilde{\mathsf{K}}_0$, the vector $\overrightarrow{\mathsf{ek}}$ includes its corresponding extraction key ek included in the triple.

3. At this moment, $\mathcal{A}$ is allowed to send $(\texttt{prove}, \mathsf{stmt})$ to the oracle multiple times; and for each such invocation, the oracle would call $\widetilde{\pi} \leftarrow \widetilde{\mathsf{P}}(\{\mathsf{crs}_j\}_{j\in[m]}, \overrightarrow{\tau}, \mathsf{stmt})$ and return the resulting $\widetilde{\pi}$ to $\mathcal{A}$.

Groth and Ostrovsky [GO14] showed how to construct a multi-CRS NIZK from standard cryptographic assumptions, resulting in the following theorem.

Theorem D.1 (Multi-CRS NIZK [GO14])**.** *Assume the existence of enhanced trapdoor permutations. Then, there exists a multi-CRS NIZK system that satisfies completeness, zero-knowledge, and simulation sound extractability.*

D.2 Protocol Description

Below we give the final multi-party computation protocol Π_{MPC}. Roughly speaking, the user identities first secret share their bids among the miners and prove in zero-knowledge the correctness of the sharings. Then, the miners run an MPC protocol using the shares they have received as inputs. The MPC protocol will securely compute the rules of the auction, and determine which bids are confirmed and how much each confirmed bid pays. We will use the honest-majority multi-CRS NIZK defined in Appendix D.1.3. Moreover, we will describe our protocol Π_{MPC} assuming that players have access to an ideal functionality $\mathcal{F}_{\text{TFM}}$ which computes the rules of the auction — the formal description of $\mathcal{F}_{\text{TFM}}$ will be provided at the end of Π_{MPC}. The ideal functionality $\mathcal{F}_{\text{TFM}}$ can be realized using standard techniques — in particular, we can use an MPC protocol that secures against minority corruptions providing fairness and guaranteed output [GMW87, RBO89]. Finally, our Π_{MPC} protocol also makes use of a perfectly binding and computationally hiding commitment scheme denoted comm.

During the protocol, miners will keep track of a set $\mathcal{C}$ containing the set of user identities who have misbehaved. The bids of those in $\mathcal{C}$ will be treated as 0. All miners have the same view of $\mathcal{C}$ since $\mathcal{C}$ is determined using only messages sent on the broadcast channel.

Protocol Π_{MPC} instantiating $\mathcal{F}_{\text{MPC}}$

Parameters: Let λ be the security parameter. Let m be the number of miners running the protocol. Let $t = \lceil \frac{m}{2} \rceil$ be the reconstruction threshold of secret sharing. Let ID be the agreed-upon set of user identities that are participating in the protocol. Let $\mathcal{C}$ be an initially empty set.

Building blocks:

- Shamir secret sharing.
- A perfectly binding, computationally hiding commitment scheme comm.
- An honest-majority multi-CRS non-interactive zero-knowledge proof (NIZK) system denoted as $\mathsf{NIZK} := (\mathsf{K}, \mathsf{P}, \mathsf{V})$.

Input: Each user identity $i \in \mathsf{ID}$ has a bid $b_i \in \mathbb{F}$. Each miner has no input.

Sharing phase

1. Each miner j runs $\mathsf{NIZK.K}(1^\lambda)$ and obtains crs_j. Each miner j broadcasts crs_j to all user identities and miners. If a miner j fails to broadcast crs_j, set $\mathsf{crs}_j = \mathbf{0}$. Let $\mathsf{CRS} := \{\mathsf{crs}_j\}_{j\in[m]}$.

2. Each user identity i splits b_i into m secret shares using a t-out-of-m secret sharing scheme. Let $X_{i,j}$ denote the j-th share of b_i. Let $\widehat{X}_{i,j} = \mathsf{comm}(X_{i,j}, r_{i,j})$ where the $r_{i,j}$s are fresh randomness. Broadcast the commitments of shares $\{\widehat{X}_{i,j}\}_{j\in[m]}$ to the miners.

   If a user identity i fails to broadcast all the commitments, each miner adds i to $\mathcal{C}$.

3. Each user identity $i \in \mathsf{ID}$ calls $\pi_i \leftarrow \mathsf{NIZK.P}(\mathsf{CRS}, \mathsf{stmt}_i, w_i)$ with the statement $\mathsf{stmt}_i = (i, \{\widehat{X}_{i,j}\}_{j\in[m]})$ and the witness $w_i = (b_i, \{X_{i,j}, r_{i,j}\}_{j\in[m]})$ to prove that

   - For each $j \in [m]$, $(X_{i,j}, r_{i,j})$ is the correct opening of $\widehat{X}_{i,j}$;
   - $\{X_{i,j}\}_{j\in[m]}$ forms a valid t-out-of-m secret sharing of b_i.

   Each user identity i broadcasts π_i.

4. For each user identity i, if it fails to broadcast π_i, or $\mathsf{NIZK.V}(\mathsf{CRS}, \mathsf{stmt}_i, \pi_i)$ outputs 0, i.e., the verifier algorithm rejects the proof, each miner adds i to $\mathcal{C}$.

5. Each user identity $i \in \mathsf{ID}$ sends $(X_{i,j}, r_{i,j})$ to miner j for all $j \in [m]$.

6. Each miner j does the following: for all $i \in \mathsf{ID} \setminus \mathcal{C}$, if it receives a message $(X_{i,j}, r_{i,j})$ that is a correct opening with respect to $\widehat{X}_{i,j}$, record $(X_{i,j}, r_{i,j})$ and broadcast (ok, i, j). Otherwise, broadcast $(\mathsf{complain}, i, j)$ to complain about user identity i.

7. Each user identity $i \in \mathsf{ID}$ does the following: for all j such that there is a complaint $(\mathsf{complain}, i, j)$ from miner j at Step 6, user identity i broadcasts the corresponding opening $(i, j, X_{i,j}, r_{i,j})$. Every miner records every correct opening $(i, j, X_{i,j}, r_{i,j})$ it hears.

8. If there exists a complaint $(\mathsf{complain}, i, j)$ from miner j in Step 6 such that user identity i has not broadcast the correct opening $(i, j, X_{i,j}, r_{i,j})$, each miner adds i to $\mathcal{C}$.

Computation Phase Miners invoke $\mathcal{F}_{\mathrm{TFM}}$ parameterized with ID, $\mathcal{C}$, the commitments of shares $\{\widehat{X}_{i,j}\}_{i\in\mathsf{ID}\setminus\mathcal{C}, j\in[m]}$, and the transaction fee mechanism. Each miner outputs the output of $\mathcal{F}_{\mathrm{TFM}}$.

Ideal Functionality $\mathcal{F}_{\mathrm{TFM}}$

Parameters: The sets ID and $\mathcal{C}$, as well as commitments of shares $\{\widehat{X}_{i,j}\}_{i\in\mathsf{ID}\setminus\mathcal{C}, j\in[m]}$ and the transaction fee mechanism.

Input: Each miner j has input $\{(X_{i,j}, r_{i,j})\}_{i\in\mathsf{ID}\setminus\mathcal{C}}$, where $(X_{i,j}, r_{i,j})$ is a correct opening of $\widehat{X}_{i,j}$.

Functionality:

1. Each miner sends its input $\{(X_{i,j}, r_{i,j})\}_{i \in \mathsf{ID} \setminus \mathcal{C}}$ to $\mathcal{F}_{\text{TFM}}$.
2. For each $j \in [m]$, the functionality $\mathcal{F}_{\text{TFM}}$ checks if $(X_{i,j}, r_{i,j})$ is an correct opening of $\widehat{X}_{i,j}$ for all $i \in \mathsf{ID} \setminus \mathcal{C}$.
3. For each $i \in \mathsf{ID} \setminus \mathcal{C}$, the functionality reconstructs b_i only using those correct openings. If the reconstruction fails, treat b_i as 0. For each $i \in \mathcal{C}$, set $b_i = 0$.
4. Let $\mathbf{b} = \{b_i\}_{i \in \mathsf{ID}}$ denote all the bids. The functionality then computes the output of the transaction fee mechanism on input $\mathbf{b}$ and sends the output to every miner.

Theorem D.2. *If the commitment scheme* comm *is perfectly binding and computationally hiding, and the honest-majority multi-CRS* NIZK *satisfies completeness, zero-knowledge and simulation sound extractability, then* Π_{MPC} *securely realizes* $\mathcal{F}_{\text{MPC}}$ *(See Figure 1) in the* $\mathcal{F}_{\text{TFM}}$*-hybrid model as long as the number of colluding miners is less than* $\frac{m}{2}$.

D.3 Proof of Theorem D.2

Below we use $\equiv$ to denote identically distributed and $\equiv_c$ to denote computationally indistinguishability. Let $\mathsf{Exp}^{\mathsf{Real}}_{\mathcal{A}}$ denote the joint distribution of the honest parties and the adversary $\mathcal{A}$'s view in the real-world experiment, where the adversary $\mathcal{A}$ who controls a subset of the miners and users interact with honest parties running the real-world protocol Π_{MPC}. Let $\mathsf{Exp}^{\mathsf{Ideal}}_{\mathcal{S}}$ denote the joint distribution of the honest parties and the ideal-world adversary $\mathcal{S}$'s view in the ideal-world experiment, where $\mathcal{S}$ controls the same subset of miners and users, and all parties interact with $\mathcal{F}_{\text{MPC}}$ to compute the outputs. We want to show that $\mathsf{Exp}^{\mathsf{Real}}_{\mathcal{A}}$ and $\mathsf{Exp}^{\mathsf{Ideal}}_{\mathcal{S}}$ are computationally indistinguishable assuming $\mathcal{A}$ is p.p.t.. In the proof, we use $\mathcal{H}_{\text{miner}}$ and $\mathcal{K}_{\text{miner}}$ to denote the set of honest miners and corrupted miners, respectively. Formally, the simulator $\mathcal{S}$ interacting with $\mathcal{F}_{\text{MPC}}$ behaves as follows.

Simulator $\mathcal{S}$ interacting with $\mathcal{F}_{\text{MPC}}$

Sharing Phase

1. Let $\mathcal{C}$ be an empty set.
2. Emulate honest miner $h \in \mathcal{H}_{\text{miner}}$ as follows: run the simulated CRS generation algorithm $\widetilde{\mathsf{K}}_0$ of NIZK and get a triple $(\mathsf{crs}_h, \tau_h, \mathsf{ek}_h)$. Send $\{\mathsf{crs}_h\}$ to $\mathcal{A}$.

   At the end of this step, define the following notation: Let $\overrightarrow{\tau}$ be the vector of $\{\tau_h\}_{h \in \mathcal{H}_{\text{miner}}}$, and $\overrightarrow{\mathsf{ek}}$ be the vector of $\{\mathsf{ek}_h\}_{h \in \mathcal{H}_{\text{miner}}}$.
3. For each corrupted miner $k \in \mathcal{K}_{\text{miner}}$, wait for its crs_k. If a corrupted miner k fails to send crs_k, set $\mathsf{crs}_k = \mathbf{0}$. Let $\mathsf{CRS} = \{\mathsf{crs}_j\}_{j \in [m]}$ be the set of all CRSes generated by miners.
4. Emulate honest user identity i as follows: For every corrupted miner $k \in \mathcal{K}_{\text{miner}}$, let the share $X_{i,k}$ be a uniformly random element in the finite field $\mathbb{F}$. For every honest miner $h \in \mathcal{H}_{\text{miner}}$, let the share $X_{i,h} = 0$.
5. Emulate honest user identity i as follows: commit to the shares $\widehat{X}_{i,j} = \mathsf{comm}(X_{i,j}, r_{i,j})$

using fresh randomness $r_{i,j}$ for each miner $j \in [m]$. Send the commitments $\{\widehat{X}_{i,j}\}_{j\in[m]}$ to $\mathcal{A}$.

6. For each corrupted user identity $\ell \in \mathsf{ID}$, wait for its commitments $\{\widehat{X}_{\ell,j}\}_{j\in[m]}$. If a corrupted user identity ℓ fails to send all the commitments, add ℓ to set $\mathcal{C}$.

7. Emulate honest user identity i as follows: call $\pi_i \leftarrow \widetilde{\mathsf{NIZK.P}}(\mathsf{CRS}, \overrightarrow{\tau}, \mathsf{stmt}_i)$, where $\mathsf{stmt}_i := (i, \{\widehat{X}_{i,j}\}_{j\in[m]})$. Send π_i to $\mathcal{A}$.

8. For each corrupted user identity ℓ, wait for π_ℓ. If a corrupted identity ℓ fails to send a proof π_ℓ, or that $\mathsf{NIZK.V}(\mathsf{CRS}, \mathsf{stmt}_\ell, \pi_\ell) = 0$ for $\mathsf{stmt}_\ell := (\ell, \{\widehat{X}_{\ell,j}\}_{j\in[m]})$, add ℓ to $\mathcal{C}$.

9. For each corrupted user identity $\ell \in \mathsf{ID} \setminus \mathcal{C}$, the simulator $\mathcal{S}$ calls the extraction algorithm $\mathcal{E}$ of NIZK and gets $w_\ell \leftarrow \mathcal{E}(\mathsf{CRS}, \overrightarrow{\mathsf{ek}}, \mathsf{stmt}_\ell, \pi_\ell)$. If there exists an ℓ such that w_ℓ is not a valid witness of stmt_ℓ, the simulator $\mathcal{S}$ aborts.

10. Emulate each honest identity $i \in \mathsf{ID}$ to send the shares for each corrupted miners $\{(X_{i,k}, r_{i,k})\}_{k\in\mathcal{K}_{\text{miner}}}$ to $\mathcal{A}$.

11. Receive the shares $\{(X_{\ell,h}, r_{\ell,h})\}_{h\in\mathcal{H}_{\text{miner}}}$ for honest miners from each corrupted identities $\ell \in \mathsf{ID}$.

12. Emulate honest miner h as follows: for each corrupted user identity $\ell \in \mathsf{ID}$, it checks whether $(X_{\ell,h}, r_{\ell,h})$ it received is a correct opening of $\widehat{X}_{\ell,h}$. If yes, send (ok, h, ℓ) to $\mathcal{A}$. Otherwise, send $(\mathsf{complain}, h, \ell)$ to $\mathcal{A}$. Meanwhile, send (ok, h, i) for each honest user identity $i \in \mathsf{ID}$ to $\mathcal{A}$.

13. Emulate honest user identity i as follows: If it received $(\mathsf{complain}, k, i)$ from a corrupted miner k, send $(i, k, X_{i,k}, r_{i,k})$ to $\mathcal{A}$.

14. For each corrupted user identity $\ell \in \mathsf{ID}$, if there exists a complaint $(\mathsf{complain}, h, \ell)$ from an honest miner h, wait for ℓ's opening $(\ell, h, X_{\ell,h}, r_{\ell,h})$.

15. For each corrupted user identity $\ell \in \mathsf{ID}$: if there exists a miner j that broadcast a complaint $(\mathsf{complain}, \ell, j)$ but ℓ did not broadcast the correct opening $(\ell, j, X_{\ell,j}, r_{\ell,j})$, then add ℓ to $\mathcal{C}$.

Computation Phase Note that by this point, if the simulator did not abort, for each corrupted user identity $\ell \in \mathsf{ID} \setminus \mathcal{C}$, the simulator $\mathcal{S}$ has extracted a valid witness $w_\ell = (b_\ell, \{X_{\ell,j}, r_{\ell,j}\}_{j\in[m]})$. The simulator sets $b_\ell = 0$ for $\ell \in \mathcal{C}$. It then sends b_ℓ for all corrupted user identities $\ell \in \mathsf{ID}$ to the ideal functionality $\mathcal{F}_{\text{MPC}}$.

After the simulator $\mathcal{S}$ receives the output from $\mathcal{F}_{\text{MPC}}$, it sends the output of the mechanism to $\mathcal{A}$ on behalf of $\mathcal{F}_{\text{TFM}}$.

We construct the following sequence of hybrid experiments.

$\underline{\mathsf{Hyb}_0}$. This experiment is identical to a real execution of Π_{MPC}, except that now the adversary $\mathcal{A}$ interacts with a fictitious simulator $\mathcal{S}'$ which internally emulates the execution of all honest players. Moreover, the simulator $\mathcal{S}'$ also emulates $\mathcal{F}_{\text{TFM}}$. We use Hyb_0 to denote the joint distribution of honest players' outputs and the adversary's view in this experiment.

By definition, $\mathsf{Exp}^{\mathsf{Real}}_{\mathcal{A}} \equiv \mathsf{Hyb}_0$.

Hyb_1. This experiment is almost identical to the experiment in Hyb_0, except the following modifications:

- Instead of calling $\mathsf{NIZK.K}$ to generate the CRS, the simulator $\mathcal{S}'$ calls the simulated CRS generation algorithm $\widetilde{\mathsf{K}}_0$, such that for each honest miner $h \in \mathcal{H}_{\text{miner}}$, the simulator gets $(\widetilde{\mathsf{crs}}_h, \tau_h, \mathsf{ek}_h)$. The simulator uses $\widetilde{\mathsf{crs}}_h$ as miner h's NIZK CRS, and keeps the trapdoor τ_h and extraction key ek_h to itself.
- Whenever the simulator $\mathcal{S}'$ needs to compute a proof on behalf of an honest user identity i, it calls the simulated prover algorithm $\widetilde{\mathsf{P}}$ supplying the trapdoor $\overrightarrow{\tau} := \{\tau_h\}_{h \in \mathcal{H}_{\text{miner}}}$ to compute a simulated proof without using the witness.

We use Hyb_1 to denote the joint distribution of honest players' outputs and the adversary's view in this experiment.

Claim D.3. *Assuming that* NIZK *satisfies zero-knowledge, then* $\mathsf{Hyb}_0 \equiv_c \mathsf{Hyb}_1$.

Proof. The proof can be done via a sequence of hybrid experiments. First, one by one for each honest miner, we replace the real generation algorithm K with the simulated generation algorithm $\widetilde{\mathsf{K}}$. Next, one by one for each NIZK proof of an honest user identity, we replace the proof with a simulated proof computed using $\widetilde{P}$ without using the witness. Since the number of corrupted miners is less than half, the adversary is minority-constrained (as defined in Appendix D.1.3), the adjacent hybrids in each step are indistinguishable by a straightforward reduction to the zero-knowledge property of NIZK. □

Hyb_2. This experiment is almost identical to the experiment in Hyb_1, except that whenever $\mathcal{A}$ supplies a correct NIZK proof π_ℓ on behalf of a corrupted user identity ℓ for statement stmt_ℓ, the simulator $\mathcal{S}'$ calls the NIZK's extraction algorithm $\mathcal{E}(\mathsf{CRS}, \overrightarrow{\mathsf{ek}}, \mathsf{stmt}_\ell, \pi_\ell)$ to extract the witness w_ℓ. If w_ℓ is not a valid witness yet $\mathsf{NIZK.V}(\mathsf{CRS}, \mathsf{stmt}_\ell, \pi_\ell) = 1$, the simulator $\mathcal{S}'$ aborts. We use Hyb_2 to denote the joint distribution of honest players' outputs and the adversary's view in this experiment.

Claim D.4. *Assuming that* NIZK *satisfies simulation sound extractability, then* $\mathsf{Hyb}_1 \equiv_c \mathsf{Hyb}_2$.

Proof. Given that the simulator $\mathcal{S}'$ does not abort, the two experiments are identical. Since the adversary controls less than half corrupted miners, by the simulation sound extractability property of NIZK, the probability of $\mathcal{S}'$ aborting in Hyb_2 is negligible. Specifically, for applying the simulation sound extractability, all NIZK statements in the protocol are tagged with the user identity (identity of the prover), thus no statement can be reused. Therefore, $\mathsf{Hyb}_1 \equiv_c \mathsf{Hyb}_2$. □

Hyb_3. This experiment is almost identical to the experiment of Hyb_2, except for the following difference:

- In the sharing phase, for each honest user identity i, instead of committing to the m shares $\{X_{i,j}\}_{j \in [m]}$ of the t-out-of-m secret sharing scheme, the simulator $\mathcal{S}'$ commits to $X_{i,k}$ for corrupted miner $k \in \mathcal{K}_{\text{miner}}$, and commits to 0 for honest miner $h \in \mathcal{H}_{\text{miner}}$.
- $\mathcal{S}'$ uses the simulated prover algorithm $\widetilde{\mathsf{P}}$ of NIZK to vouch for honest user identities.
- Upon receiving the openings, it sends (ok, h, i) for all honest user identities $i \in \mathsf{ID}$ and all honest miners $h \in \mathcal{H}_{\text{miner}}$, without actually checking the openings of the commitments.

We use Hyb_3 to denote the joint distribution of honest players' outputs and the adversary's view in this experiment.

Claim D.5. *Assuming that the commitment scheme* comm *is computationally hiding, then* $\mathsf{Hyb}_2 \equiv_c \mathsf{Hyb}_3$.

Proof. The proof can be done via a sequence of hybrid experiments, where one by one for each honest user identity i, we replace the commitments $\{\widehat{X}_{i,h}\}_{h\in\mathcal{H}_{\text{miner}}}$ of the shares $X_{i,h}$ with commitments of 0. The adjacent hybrids in each step are indistinguishable by a direct reduction to the computational hiding property of comm. □

Recall that $\mathsf{Exp}^{\mathsf{Ideal}}_{\mathcal{S}}$ denotes the honest players' outputs computed by $\mathcal{F}_{\text{MPC}}$ and the view simulated by $\mathcal{S}$ which interacts with $\mathcal{F}_{\text{MPC}}$.

Claim D.6. *If the commitment scheme* comm *is perfect binding and that the* t*-out-of-*m *secret sharing scheme is secure, then* $\mathsf{Hyb}_3 \equiv \mathsf{Exp}^{\mathsf{Ideal}}_{\mathcal{S}}$.

Proof. The only differences in Hyb_3 and $\mathsf{Exp}^{\mathsf{Ideal}}_{\mathcal{S}}$ are:

1. In $\mathsf{Exp}^{\mathsf{Ideal}}_{\mathcal{S}}$, the simulator is generating honest-to-corrupt shares at random; whereas in Hyb_3, the honest-to-corrupt shares are generated honestly. By the security of Shamir secret sharing, the two approaches result in the same distribution since the adversary controls fewer than $m/2$ miners.

2. In Hyb_3, if the experiment did not abort, then the simulator sends the shares actually opened by corrupt user identities to $\mathcal{F}_{\text{TFM}}$. By contrast, in $\mathsf{Exp}^{\mathsf{Ideal}}_{\mathcal{S}}$, the simulator uses the shares output by the NIZK's extractor $\mathcal{E}$ instead. Since the commitment is perfectly binding, the two approaches result in the same outcome as long as the simulator did not abort.

Therefore, the two hybrids are identically distributed. □

By the hybrid lemma, we have that $\mathsf{Exp}^{\mathsf{Real}}_{\mathcal{A}} \equiv_c \mathsf{Exp}^{\mathsf{Ideal}}_{\mathcal{S}}$. Therefore, the protocol Π_{MPC} securely realizes $\mathcal{F}_{\text{MPC}}$ in the $\mathcal{F}_{\text{TFM}}$-hybrid model as long as the adversary controls only a minority number of miners.

D.4 MPC Protocol in the Presence of Majority-Miner Coalitions

So far, we have focused on instantiating the MPC protocol when the coalition controls only minority of the miners. As we explained in Remark 1.5, our game-theoretic analyses also naturally extend to the case when the coalition may control majority of the miners.

In this case, we can modify our MPC protocol as follows to achieve security with abort under corrupt majority. First, instead of threshold secret sharing, the user identities may use additive secret sharing to share their bids among the miners. As before, each user identity will broadcast commitments of all shares of its bid, and then it gives the corresponding opening to every miner. There is no more need to prove that the committed values are internally consistent secret shares. If a miner did not receive the correct opening from a user identity, it can broadcast a complaint in which case the corresponding user identity must reveal the correct opening or it will get kicked out. During the reconstruction phase, if any miner fails to open, then the protocol just aborts and no output is produced, i.e., no block will be mined. Finally, $\mathcal{F}_{\text{TFM}}$ should also be instantiated with a corrupt majority MPC protocol.

E Efficient Instantiations of our MPC-Assisted Mechanisms

The MPC-assisted mechanisms proposed in our paper, including posted price with random selection and the diluted posted price mechanism, achieve incentive compatibility in the ex post setting. This means that instantiating these mechanisms in practice actually does not require the use of generic MPC. We can use the following efficient protocols:

- Instead of having the user identities verifiably secret share their bids with the miners, they can simply post the bids in the clear over a broadcast channel. In practice, we can use any consensus mechanism to realize the broadcast channel, such that the miners agree on the set of all bids posted. In particular, we can use the underlying blockchain itself to reach this consensus — importantly, if we do this, we stress that the initial set of bids agreed upon need not be permanently stored by the blockchain, i.e., here we are using the blockchain for (transient) consensus but not for storage.

- Once the miners agree on the initial set of bids, they can then run any coin toss protocol to decide a randomness seed, which can be used to generate the random coins and perform the random selection needed by the mechanisms.