What Can Cryptography Do For Transaction Fee Mechanism Design

Ke Wu

CMU-CS-24-115
May 2024

Computer Science Department
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

Thesis Committee:
Elaine Shi, Chair
Ryan O'donnell
Aayush Jain
Tim Roughgarden (Columbia University, A16Z)

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

This research was sponsored 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 author is also supported by CMU Cylab Presidential Fellowship and JP Morgan Chase AI Research PhD Fellowship.

Keywords: Transaction fee mechanism, cryptography

To my beloved mother, Xiaoli, and my husband, Bingkai.

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). Unfortunately, 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, even if the block size is infinite. Second, assuming a finite block size, no non-trivial TFM can simultaneously provide incentive compatibility for any individual user and for any miner-user coalition.

This thesis explores potential relaxations and the theoretical landscape of transaction fee mechanisms under these relaxations. We delve into four key directions:

1. *MPC-assisted model*. We introduce a new MPC-assisted model, where the TFM is implemented through a joint multi-party computation (MPC) protocol among miners. While this model does not get rid of the zero-miner revenue limitation, it indeed allows us to overcome some impossibility results pertaining to the original model (henceforth called the *plain model*), leading to non-trivial mechanisms with useful guarantees that are otherwise impossible in the plain model.
2. *Approximate Incentive Compatibility*. Allowing strategic players to gain no more than ϵ-additional utility compared to honest behavior, we design mechanisms with positive miner revenues in both the plain and MPC-assisted models. Despite achieving optimality with respect to the miner revenue, these mechanisms have poorly scalable miner revenue, as we proved with certain impossibility results.
3. *Reasonable-world assumption*. We show that if we make a mildly stronger assumption assuming that we know a lower bound h on the number of honest users and an upper bound d on the number of bids controlled by the coalition, we can circumvent the previous limitations on miner revenue, and design mechanisms that generate optimal miner revenue linear in h.
4. *Miner-user coalition proof*. Here, we consider another flavor of notion capturing incentives of the miner-user coalitions: miner-user coalition proof, which requires that any miner-user coalition is unstable. We show that, under this new notion, we are able to design interesting transaction fee mechanisms for finite block sizes that satisfy incentive compatibility for any individual user and any miner coalition, as well as miner-user coalition proofness.

Acknowledgments

I would like to express my deepest appreciation to my amazing advisor, Elaine Shi, who has offered invaluable support and help throughout my Ph.D. studies. Elaine spent a lot of time guiding me and teaching me how to write well, how to present good talks, and how to be a good faculty. Elaine has been everything I could have hoped for in an advisor. Thank you, always.

I am deeply appreciative of the time and advice from my thesis committee members, Tim Roughgarden, Ryan O'donell, and Aayush Jain. I would like to thank my collaborators Venkatesan Guruswami, Aaron Wagner, T-H. Hubert Chan, Xin Li, Ryan Gabrys, Gilad Asharov, Ilan Komargodski, Pratik Soni, Sri Aravinda Krishnan Thyagarajan, João Ribeiro, Hao Chung, Kuan Cheng, and Zhengzhong Jin. It has been an honor and a pleasure to collaborate with each of you.

Thank you to the cryptography and security community at CMU: Chen-Da Liu Zhang, Rex Fernando, Wei Dong, Ashrujit Ghoshal, Andrew Park, Mingxun Zhou, Afonso Tinoco, Junxi Song, Nikhil Vanjani, Tianyao Gu, Yifan Song, Abhiram Kothapalli, Justin Raizes, Quang Dao, Orestis Chardouvelis, Alper Cakan, Edward Chen.

Thank you to my wonderful friends, Yuchen Shen and Shengyuan Hu. I enjoy every Saturday night we spend together. Mingxuan Xu, Yue Yao, Kaiyang Zhao, Ziyue Qiu, Yuang Chen, and Longxin Xie, thank you for every interesting dinner we have had together and the intriguing discussions we have had. Lisa Masserova, every outing with you has been a delight. I also want to thank Chenyang Yang, Matias Scharager, Tanli Su, and my friends Nathan Yan, Junxiong Wang, Tao Yu, Jingyi Duan from Cornell University.

Finally, to my family and Bingkai, your support has been my rock. Thank you, forever.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The recent success of cryptocurrency and blockchains offer innovative financial services. As these applications continue to gain widespread adoption, the demand for blockchain is rising. However, space on blockchains is a scarce resource. On average, the two largest blockchains, Bitcoin and Ethereum processes roughly 5 and 15 transactions per second. To address space allocation among competing transactions, *transaction fee mechanisms* are introduced [LSZ19, Yao, BEOS19, BCD+, Rou20, Rou21, FMPS21, CS23, GY22, ZCZ22, SCW23, WSC24]. In a transaction fee mechanism (TFM), whenever a new block is proposed, the space in the block is auctioned off to the users who bid to get their transactions included and confirmed in the block. If the block can contain up to k number of transactions, one can think of the transaction fee mechanism as an auction selling k identical items to the users.

Although the transaction fee mechanism can be equivalently thought of as an auction, earlier works [LSZ19, Yao, BEOS19, BCD+, Rou20, Rou21, FMPS21, CS23, GY22, ZCZ22] observe that the transaction fee mechanism design departs significantly from classical auction design due to challenges posed by the decentralized nature of blockchain.

- First, most classical auction design assumes an honest auctioneer who always implements the prescribed mechanism. However, on the blockchain, the transaction fee mechanism is implemented by the miners[1] who may take advantage of profitable deviations and not implement the mechanism honestly.
- The existence of smart contracts in blockchain makes it easy for the miners and users to form binding coalitions. In particular, they can behave strategically to increase their joint profits and split the gain off-chain. Such collusion is typically not handled in classical auctions.
- Most blockchains offer a permissionless environment: anyone can register for pseudonyms and submit fake bids under these extra pseudonyms or drop off their bids. Therefore, the transaction fee mechanism has no information about the *actual* number of users participating in the auction.

In response to these challenges, earlier works [LSZ19, Yao, BEOS19, BCD+, Rou20, Rou21] formulated the following desiderata for a "dream" transaction fee mechanism:

- *User Incentive Compatibility (UIC)*: Each user is incentivized to bid its true value, represent-

[1] Throughout this thesis, we refer to those consensus nodes who propose new blocks as miners, no matter whether the consensus protocol uses proof-of-work or proof-of-stake.

ing the maximum amount the user is willing to pay. This guarantee should hold even if the user decides its strategy after observing all other users' bids.

- *Miner Incentive Compatibility (MIC)*: The miner is incentivized to implement the mechanism honestly, even after observing all users' bids.
- *c-Side-Contract-Proofness (c-SCP)*: If the miner colludes with no more than c number of users and their goal is to maximize their joint utility, then they are incentivized to behave honestly, even after observing all other users' bids.

Roughgarden [Rou20, Rou21] demonstrated that assuming an infinite block size (i.e., no congestion), Ethereum's EIP-1559 [BCD+] satisfies all three properties. Roughly speaking, when there is no congestion, EIP-155p acts as a "burning posted price auction": there is a fixed reserved price r, every user who bids at least r gets confirmed and pays r, all payments are burnt rather than paid to the miners. In practice, EIP-1559 estimates the demand and dynamically adjusts the reserved price to stay in the "infinite block size" regime.

However, congestion does happen. In practice, roughly 2.4% of the blocks experience congestion in Ethereum [CRS24]. Chung and Shi [CS23] explored the landscape of TFM for the finite block size regime and proved the following fundamental impossibility results: 1.) Any TFM that satisfies UIC and SCP must suffer from 0-miner revenue[2], and 2.) No non-trivial TFM satisfies all three properties for a finite block size. In light of the status quo of our understanding, we ask the following question:

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

Chung and Shi [CS23] made an initial effort along this line. Assuming offending bids (e.g., overbid or fake transactions) that have been posted to the public cannot be retracted in the future, they give a mechanism that circumvents the impossibilities and achieve positive miner revenue under finite block size. While this assumption holds for some cryptocurrencies like Bitcoin, it may not be universally true for all cryptocurrencies. Therefore, an important question is what are the other models or relaxations we can explore.

In this thesis, we explored the following four directions, aiming to understand whether they allow us to circumvent the impossibility results of Chung and Shi [CS23] and how these different relaxations (or a mix of several relaxations) alter the landscape of TFMs.

1. *MPC-assisted model [SCW23]*: Having a group of miners jointly run a multi-party computation (MPC) protocol to implement the TFM instead of having a single miner implement the TFM.
2. *Approximate incentive compatibility [SCW23]*: Relaxing the incentive compatibility notion and allowing an ϵ additive slack.
3. *Resonable-world assumption [WSC24]*: Assuming sufficient honesty among the number of users: in particular, we assume that at least h number of users are honest.
4. *Miner-User Coalition Proofness (MUCP)*: Another notion capturing incentive compatibility for miner-user coalitions. Roughly speaking, MUCP guarantees that the miner-user coalition is unstable and, therefore, disincentivizes the miners/users to form a coalition.

[2] We ignore the fixed block reward of the miner in practice since it is irrelevant to our analysis.

Throughout the thesis, we refer to today's model, where a single miner implements the mechanism without the use of any 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 Contributions

In our exploration of various relaxations, we contribute novel insights at both the conceptual and technical levels. From a technical perspective, we develop new tools for characterizing the solution space under certain relaxations. Specifically, we develop techniques for mathematical reasoning of approximate incentive compatibility, and our mechanisms yield interesting techniques which may be used in other auction designs. On the conceptual front, while existing research has demonstrated the synergy between cryptography and game theory [HT04, KN08, ADGH06, OPRV09, AL11, ACH11, GKM+13, GKTZ15, GTZ15, Kat08, DR07, GLR10, CGL+18, WAS22, CCWS21, PS17, KMSW22, FW20, EFW22] (see Section 1.1.5 for further discussions), our results reveal new connections between cryptography and game theory. Our results highlight the potential of cryptography-meets-game-theory as a promising paradigm for addressing challenges arising from the popularity of blockchains and decentralized applications.

Moreover, our investigation delves into the nuanced definition of "sufficient honesty" and Miner-User Coalition Proofness (MUCP). Defining such relaxations reveals subtleties, as we will elaborate on later in Section 1.1.3 and Section 1.1.4. Notably, certain natural notions fail to yield new and interesting results. In presenting these conceptual contributions, we aim to inspire future works in the captivating space of decentralized mechanism design.

We give a summary of our main results below.

1.1.1 MPC-Assisted Model

While blockchain inherently provides a decentralized environment, the original TFM model exhibits some level of centralization: the miner of the current block has dictatorial control over the content, i.e., the miner receives users' bids and decides which bids to include in the block. To seek avenues that allow us to circumvent the previous impossibility result, we ask: *can we also decentralize the implementation of the TFM?*

New model: MPC-assisted Model.

Consider a scenario hereafter 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 an 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 represented by a non-negative real number to $\mathcal{F}_{\text{TFM}}$.
- The ideal functionality $\mathcal{F}_{\text{TFM}}$ executes the prescribed rules of the TFM and then sends to all players the outcome of the TFM: the set of bids that are confirmed, what price each confirmed bid pays, and the total miner revenue.

To ensure budget feasibility and individual rationality, we require that user payments must not exceed their bids, and the total miner revenue should not surpass the total payment. The total miner revenue from the TFM is split among the participating miners.

Throughout this thesis, we assume that there is a separate process to decide the set of miners to run the MPC. For example, this decision can be made through either proof-of-work or proof-of-stake, where the total miner revenue is effectively split among the miners proportional to their mining power or stake, respectively.

In subsequent sections, we will analyze different incentive compatibility notions and mechanisms assuming an ideal MPC-assisted model, i.e., we assume that a trusted party implements the above ideal functionality $\mathcal{F}_{\text{TFM}}$. Later in Chapter 7, we will show how to instantiate $\mathcal{F}_{\text{TFM}}$ with real-world cryptography and how we extend the analysis to achieve computational soundness in the real world.

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 is committed to, the transaction fee mechanism 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).

Further, the MPC-assisted model enables a relaxed notion of incentive compatibility.

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, CS23] considered an *ex post* notion of incentive compatibility. This notion requires that for any strategic coalition or individual, being honest is the best strategy even *after* they observe other users' bids.

However, in the new MPC-assisted model, players must submit their bids to $\mathcal{F}_{\text{TFM}}$ without seeing others' bids. Hence, it also makes sense to consider a *Bayesian* notion of incentive compatibility. Informally, we say that an MPC-assisted TFM satisfies *Bayesian Incentive Compatibility* 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. We say that a scheme satisfies Bayesian UIC, ρ-MIC or (ρ, c)-SCP, respectively, if the coalition $\mathcal{C}$ consists of a single user, at most ρ fraction of miners, or at most ρ fraction of miners as well 1 to c number of users.

Results in the MPC-Assisted Model

Exactly because the MPC-assisted model imposes the above restrictions on the strategy space and enables Bayesian incentive compatibility, we are hopeful that it may allow us to circumvent impossibilities. Table 1.1 below summarizes the landscape of TFMs under the plain and the MPC-assisted model. We have matching upper- and lower bounds with respect to the miner revenue. In the table, ✗ represents impossibility, and results in green background are presented in this thesis.

	Plain model	MPC-assisted model	
Infinite block size	0-miner revenue [CS23, Rou20]	0-miner revenue	
Finite block size	✗ [CS23]	$c = 1$	0-miner revenue
		$c \geq 2$	0-social welfare

Table 1.1: Landscape for TFMs satisfying UIC, MIC and c-SCP under the plain model and the MPC-assisted model.

Notably, the MPC-assisted model, even with Bayesian notions of incentive compatibility, fails to circumvent the zero miner revenue lower bound, even for infinite block size.

Theorem 1.1.1 (Limit on miner revenue in MPC-assisted model)**.** *For any possibly randomized TFM in the MPC-assisted model that satisfies Bayesian UIC, Bayesian MIC, and Bayesian* 1*-SCP, the expected total miner revenue must be* 0*.*

Now, 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.

For $c = 1$, i.e., when there is only one user in the miner-user coalition, we show that the MPC-assisted model indeed circumvents the strong finite-block impossibility of Chung and Shi [CS23]. In particular, the following *posted price auction with random selection* satisfies UIC, MIC, and $(\rho, 1)$-SCP for any $\rho \in (0, 1]$.

MPC-assisted, posted price auction with random selection

Let r be a fixed reserve price and k be the block size. Randomly choose up to k bids at least =t r to confirm. Any confirmed bid pays r. All payments are burnt, and the miner revenue is 0.

Theorem 1.1.2 (MPC-assisted, posted price auction with random selection)**.** *The above MPC-assisted, posted price auction with a random selection satisfies UIC, MIC, and* $(\rho, 1)$*-SCP in the ex post setting for an arbitrary* $\rho \in [0, 1]$*.*

Since Theorem 1.1.2 holds even in the ex post setting, another interpretation is that the enforcement of the honest implementation (i.e., restriction R1, not R2) is what allows us to circumvent the finite-block impossibility when $c = 1$.

However, the above mechanism fails when the coalition may contain $c \geq 2$ users. Imagine the number of users $n = k+1$, and the coalition consists of two users and any fraction of miners. Suppose one of the colluding users has a true value $v \gg r$, and the other has a true value $v' = r$. In this case, the user with true value $v' = r$ should drop out and not submit a bid. This strategy guarantees that the friend with the 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 with positive social welfare 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 1.1.3 (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 positive social welfare can simultaneously achieve Bayesian UIC, Bayesian MIC, and Bayesian SCP for (ρ, c)-sized coalitions, assuming finite block size.*

Note that all our impossibility results hold for Bayesian incentive compatibility while our mechanism satisfies ex post incentive compatibility. This makes both our upper- and lower-bounds stronger.

1.1.2 Approximate Incentive Compatibility

While the MPC-assisted model circumvents the previous impossibility result, it does not give rise to interesting mechanisms: the miner revenue must be 0, and even the social welfare must be 0 for finite block size if we want to tolerate miner-user coalitions containing 2 or more users. Consequently, we need to explore alternative approaches to achieve positive miner revenue and social welfare. We therefore 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?*

The table below summarizes the landscape for approximate incentive compatibility under the plain and MPC-assisted models. We use k to represent the block size and n to represent the number of users. Here, $C_{\mathcal{D}}$ and M are constant terms depending on the scale of the bid distribution of users' true values, and $\Theta(\cdot)$ implies asymptotically matching upper- and lower-bounds.

	Plain model	MPC-assisted model
Infinite block size	$\Theta(n \cdot (\epsilon + C_{\mathcal{D}}\sqrt{\epsilon}))$-miner rev	$\Theta(n \cdot (\epsilon + C_{\mathcal{D}}\sqrt{\epsilon}))$-miner rev
Finite block size	Impossibility: $\widetilde{O}(k^3\epsilon)$-social welfare Feasibility: $\Theta(k^2\epsilon)$-social welfare	$\Theta(k \cdot M)$-social welfare

Table 1.2: Landscape for TFMs satisfying ϵ-UIC, ϵ-MIC and ϵ-SCP against any miner-user coalition with at most c users.

Infinite block size.

Given a slack ϵ and an integer $c \geq 1$, consider a simple posted price auction with a reserve price $r \leq \frac{\epsilon}{c}$: 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 in the plain model. 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 mechanism achieves a miner revenue linear in n, 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 then show that the following randomized TFM achieves scalable 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, the 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 1.1.4. *Consider the hybrid auction which runs either the posted posted price auction with reserve price $\frac{\epsilon}{c}$ or the proportional auction with the reserve price of r, depending on which one has higher expected revenue. The hybrid auction in the plain model 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 $\Omega\left(n \cdot (\min(\sqrt{\frac{r\epsilon}{c}}, r) + \frac{\epsilon}{c})\right)$ expected total miner revenue.*

Next, we prove a matching bound that shows the limitation of how much miner revenue can be attained under approximate incentive compatibility, as stated in the following theorem — this bound holds for Bayesian incentive compatibility, no matter whether the block size is finite or infinite. This makes both our feasibility results and impossibility results stronger.

Theorem 1.1.5 (Limit on miner revenue for infinite block size)**.** *For any possibly randomized TFM (in the MPC-assisted model) that satisfies Bayesian 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 $\mathcal{D} = \mathcal{D}_1 \times \cdots \times \mathcal{D}_n$ must be upper bounded by*

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

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 $i \in [n]$, and $\mathcal{C}_{\mathcal{D}} = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}_{X\sim\mathcal{D}_i}[\sqrt{X}]$.

Finite block size

Next, we discuss the landscape of approximate incentive compatibility for finite block size in both the plain model and the MPC-assisted model.

Plain model. Unfortunately, introducing approximate incentive compatibility does not yield substantial benefits in the plain model for finite block size. We establish a new impossibility result which 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 1.1.6 (Scalability barrier for approximate incentive compatibility in the plain model)**.** *Fix any $\epsilon > 0$, and suppose 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.*

MPC-assisted model Conversely, considering approximate incentive compatibility in the MPC-assisted model can overcome the above scalability barrier. Specifically, we construct an MPC-assisted TFM called the "diluted posted price auction" that can achieve asymptotically optimal social welfare when enough users' bids are large enough.

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 of r.
- The miner gets $\frac{2\epsilon}{c}$ for each confirmed bid.

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 each confirmed bid obtains utility $M/2$. Thus, we can achieve $\Theta(M \cdot k)$ social welfare, which is optimal.

Theorem 1.1.7 (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.*

Our results show that cryptography can help us circumvent fundamental impossibilities of the plain model under finite block size for both strict incentive compatibility and approximate incentive compatibility. On the other hand, cryptography is also not a panacea. For example, the MPC-assisted model does not fundamentally help us improve miner revenue in the infinite block size setting. In addition, the MPC-assisted model does not help with the social welfare for strict incentive compatibility.

1.1.3 Reasonable-World Assumption

While the above two relaxations give us feasible mechanisms that satisfy all three properties as well as positive miner revenue, the miner revenue we can hope for has weak scalability. For strict incentive compatibility, the zero miner revenue limitation holds in a very strong sense: regardless of whether the block size is finite or infinite and even when we consider Bayesian incentive compatibility and the miner colludes with at most $c = 1$ user. For approximate incentive compatibility, the miner revenue cannot enjoy linear scaling w.r.t. the magnitude of the bids (Theorem 1.1.5). Inspired by the philosophy adopted in a line of work at the intersection of cryptography and game theory [HT04, KN08, ADGH06, OPRV09, AL11, ACH11, GKM+13, GKTZ15, GTZ15, Kat08, DR07, GLR10, CGL+18, WAS22, CCWS21, PS17, KMSW22, FW20, EFW22],

where the game theoretic properties hold as long as sufficiently many players are honest, we ask the following natural question:

Can we circumvent the severe limitation on miner revenue, assuming sufficient honesty? Notably, since the game-theoretic guarantees incentivize honest behaviors, this in turn reinforces the "sufficient honesty" assumption.

Phrasing the precise "sufficient honesty" assumption, however, turns out to be technically subtle. One naïve attempt is to assume an honest majority among the players. But this assumption does not work because TFMs must work in an *open* setting where anyone can post a bid, and the mechanism is unaware of the number of users a-priori. In Section 5.5.3, we show that even under such an "honest majority bids" assumption, we would still suffer from an $O(1)$-miner revenue limitation.

Reasonable-world assumption: known lower bound on the number of honest users.

Instead of the "honest majority bids" assumption, we make a subtly different assumption — we assume that there is an a-priori known lower bound h on the number of honest users. Note that this assumption also promises that at least h users will show up. We refer to this as the *known-h model*.

The table below summarizes the landscape for the known h-model. We use ✗ to represent impossibility. Still, k represents the block size. Here, $\Theta_{\mathcal{D}}(\cdot)$ means that the hiding constant term depended on users' true value distribution. In the table, $\Theta(\cdot)$ or $\Theta_{\mathcal{D}}(\cdot)$ imply asymptotically matching upper- and lower-bounds.

		Plain model	MPC-assisted model
Infinite block size	strict	0 miner rev	$\Theta_{\mathcal{D}}(h)$-miner rev
	approx	$\Theta(n \cdot (\epsilon + \sqrt{m\epsilon}))$-miner rev	$\Theta_{\mathcal{D}}(h)$-miner rev
Finite block size	strict	✗	$\Theta_{\mathcal{D}}(\min\{h, k\})$-miner rev $c \geq 2$: 0-user social welfare
	approx		$\Theta_{\mathcal{D}}(\min\{h, k\})$-miner rev $\Theta_{\mathcal{D}}(k)$-user social welfare

Table 1.3: Landscape for TFMs the known-h model. Here, strict and approx stand for strict incentive compatibility and approximate incentive compatibility with an ϵ-slack, respectively.

In the introduction, we will give the informal results assuming every honest user's true value is i.i.d. sampled from some a-priori known distribution $\mathcal{D}$, while the strategic users' true values can be *arbitrary* non-negative numbers. In Chapter 5, we will give the formal results where we only need to assume that honest users' true values are sampled independently and have the same median m.

With the known-h model, we first observe that the zero miner revenue limitation no longer holds. Instead, we can prove an $O(h)$-limit on the miner revenue as stated in the following

theorem.

Theorem 1.1.8 (Informal: limit on miner revenue in the known-h model)**.** *In the known-h model, no MPC-assisted mechanism that simultaneously satisfies UIC, MIC, and SCP (even in the Bayesian setting) can achieve more than $h \cdot \mathbb{E}(\mathcal{D})$ expected miner revenue where $\mathbb{E}(\mathcal{D})$ denotes the expectation of the value distribution $\mathcal{D}$.*

More generally, in the known-h model, if the number of users is n, no MPC-assisted mechanism that simultaneously satisfies ϵ-UIC, ϵ-MIC, and ϵ-SCP (even in the Bayesian setting) can achieve more than $h \cdot \mathbb{E}(\mathcal{D}) + \frac{2(n-h)}{\rho}(\epsilon + C_{\mathcal{D}}\sqrt{\epsilon})$ expected miner revenue, where ρ is an upper bound on the fraction of miners controlled by the strategic coalition, and $C_{\mathcal{D}} = \mathbb{E}_{X\sim\mathcal{D}}[\sqrt{X}]$.

Furthermore, the above limitation holds no matter when the block size is finite or infinite, and even when the miners collude with at most $c = 1$ user.

One informal interpretation of the above theorem is the following: for ϵ incentive compatibility, Theorem 1.1.8 allows us to hope for a mechanism where roughly speaking, from each of h users, the miners can hope to get $\mathbb{E}(\mathcal{D})$ revenue which scales proportionally w.r.t. to the bid distribution $\mathcal{D}$. For each of the remaining users, the miners can potentially get some function that depends on ϵ and the bid distribution $\mathcal{D}$, but the term does not scale linearly w.r.t. the magnitude of the bid distribution for natural distributions.

Theorem 1.1.8 allows us to hope for a TFM in the known-h model that achieves revenue that scales with h as well as the magnitude of the bid distribution $\mathcal{D}$. So, can we indeed design a mechanism with asymptotically optimal mine revenue matching Theorem 1.1.8?

Mechanisms for infinite block size.

For the infinite block size regime, we propose two mechanisms in the MPC-assisted model:

- The first one, called *threshold-based mechanism*, is a simple and practical mechanism that satisfies almost-strict incentive compatibility except for a tiny slack ϵ that is exponentially small in h.
- The second one, called *LP-based mechanism* (since it uses linear programming), is a result of theoretical interest. It achieves *strict* incentive compatibility, but under one extra assumption (besides a-priori knowledge of h), that the number of fake bids injected by the strategic coalition is bounded by some a-priori fixed number d.

Both mechanisms achieve asymptotically optimal miner revenue[3] w.r.t. Theorem 1.1.8. Next, we state the corresponding theorems for the two mechanisms below.

MPC-assisted, threshold-based mechanism // Let m be the median of the distribution $\mathcal{D}$.

- All bids at least m get confirmed and pay m.
- If the number of confirmed bids is at least $h/4$, then the miner revenue is $m \cdot h/4$; Otherwise, the total miner revenue is 0.

Theorem 1.1.9 (Informal: threshold-based mechanism)**.** *Suppose honest users' values are sampled i.i.d. from some distribution $\mathcal{D}$. Then, there exists an MPC-assisted TFM in the known-h*

[3] We achieve asymptotic optimal miner revenue w.r.t. h assuming that the expectation and median of the distribution $\mathcal{D}$ is a constant independent of h.

model that satisfies ex post UIC, Bayesian ϵ-MIC, and Bayesian ϵ-SCP (for any number of colluding users) for $\epsilon = O_{\mathcal{D}}(\exp(-\Omega(h)))$ where $O_{\mathcal{D}}(\cdot)$ hides terms related to the value distribution $\mathcal{D}$. Furthermore, the expected total miner revenue $\Theta(h) \cdot \mathsf{median}(\mathcal{D})$.

Essentially, from each of h users, the miners can obtain revenue that scales linearly w.r.t. both the bid magnitude. By contrast, without the known-h assumption, for our choice of ϵ which is exponentially small in h, the miner revenue must be exponentially small in h, as shown in Theorem 1.1.5. Observe also that the miner revenue is asymptotically optimal up to additive factors that are exponentially small in h due to Theorem 1.1.8.

The second mechanism, the LP-based mechanism, assumes that the number of bids controlled by strategic coalition/individual is at most d.

MPC-assisted, LP-based mechanism // Let m be the median of the distribution.

- All bids at least m get confirmed and pay m.
- Let n be the length of the bid vector, let $\mathbf{y} := (y_0, y_1, \ldots, y_n)$ where y_i represents how much the miner gets when the number of confirmed bids is i. Solve the following linear program:

$$\forall i \in [n]: \ 0 \le y_i \le i \cdot m$$

$$\forall 0 \le j \le d: \ \mathbb{E}[\text{miner revenue} | j \text{ bids from strategic coalition confirmed}] = \frac{m \cdot h}{4}.$$

- The total miner revenue is y_s where s is the number of bids confirmed.

Theorem 1.1.10 (Informal: LP-based mechanism)**.** *Suppose honest users' values are sampled i.i.d. from some distribution $\mathcal{D}$. Then, there exists an MPC-assisted TFM that satisfies ex post UIC, Bayesian MIC, and Bayesian SCP where the MIC and SCP guarantees hold as long as the total number of bids d contributed by the strategic coalition satisfies $d \le \frac{1}{8}\sqrt{\frac{h}{2\log h}}$. Further, the expected total miner revenue is $\Theta(h) \cdot \mathsf{median}(\mathcal{D})$.*

In the above theorem, we need the extra assumption that the strategic coalition does not control too many bids. Effectively, this assumes that the coalition cannot inject too many fake bids. Currently, we do not know whether this extra assumption (besides known-h) is needed to overcome the zero miner revenue limitation. We leave this as an interesting open question.

Mechanisms for finite block size.

For the finite block size case, we propose two mechanisms:

- We propose a simple mechanism called *diluted threshold-based mechanism* that achieves *approximate* incentive compatibility. Further, for sufficiently large h, the mechanism achieves asymptotically optimal miner revenue.
- For theoretical interest, we propose another mechanism called *LP-based mechanism with random selection* which achieves *strict* incentive compatibility and asymptotically optimal miner revenue — but under the additional assumption that the coalition cannot inject too many fake bids, and moreover, the miners collude with at most $c = 1$ user. Jumping ahead, the $c = 1$

assumption will be later justified in Theorem 1.1.13.

MPC-assisted, diluted threshold-based Mechanism
// Let k be the block size amd m be the median of $\mathcal{D}$, let M be the maximum value of the distribution $\mathcal{D}$.

- Let $R := \max\left(2c\sqrt{\frac{kM}{\epsilon}}, k\right)$. All bids offering at least m are candidates. If the number of candidates $s \le R$, randomly select $\frac{k}{R} \cdot s$ candidates to confirm; else, randomly select k candidates to confirm. Every confirmed bid pays m.
- If $s \ge \frac{h}{4}$, then the total miner revenue is $\min(\frac{h}{4} \cdot \frac{k}{R}, k) \cdot m$. Otherwise, the miners get nothing.

Theorem 1.1.11 (Informal: diluted threshold-based mechanism)**.** *Suppose the block size is k, and that honest users' values are sampled i.i.d. from some bounded distribution $\mathcal{D}$. Then, there exists an MPC-assisted TFM in the known-h model that satisfies ex post UIC, Bayesian ϵ-MIC, and Bayesian ϵ-SCP (for any number of colluding users) for $\epsilon = O_{\mathcal{D}}(\exp(-\Omega(h)))$. Furthermore, for sufficiently large h, the mechanism achieves expected total miner revenue $\Theta(k) \cdot \mathsf{median}(\mathcal{D})$.*

The second mechanism, LP-based mechanism with random selection, assumes that the strategic coalition/individual controls at most d number of bids.

MPC-assisted, LP-based mechanism with random selection
// Let m be the median of the distribution.

- All bids offering at least m are candidates. If there are more than k candidates, randomly select k of them to confirm; else confirm all candidates. Every confirmed bid pays m.
- Let n be the length of the bid vector, let $\mathbf{y} := (y_0, y_1, \ldots, y_n)$ where y_i represents how much the miner gets when the number of candidates is i. Solve the following linear program:

  $$\forall i \in [n] : \ 0 \le y_i \le \min(i, k) \cdot m$$

  $$\forall 0 \le j \le d : \ \mathbb{E}[\text{miner revenue} | j \text{ bids from strategic coalition confirmed}] = \frac{m \cdot \min(k, h)}{4}.$$

- The total miner revenue is y_s where s is the number of bids confirmed.

Theorem 1.1.12 (Informal: LP-based mechanism with random selection)**.** *Suppose the block size is k, and suppose that honest users' values are sampled i.i.d. from some distribution $\mathcal{D}$. Then, there exists an MPC-assisted TFM that satisfies ex post UIC, Bayesian MIC, and Bayesian SCP, where the MIC and SCP guarantees hold when 1) at most $c = 1$ user colludes when miners, and 2) the total number of bids d contributed by the strategic coalition satisfies $d \le \frac{1}{8}\sqrt{\frac{h}{2\log h}}$. Further, the expected total miner revenue is $\Theta(\min\{h, k\}) \cdot \mathsf{median}(\mathcal{D})$.*

We justify the $c = 1$ assumption in the LP-based mechanism with random selection by proving the following impossibility result: for finite block size, no "interesting" mechanism can simultaneously achieve UIC, MIC, and SCP for $c \ge 2$ even in the MPC-assisted model.

Specifically,

Theorem 1.1.13 (Informal: finite block, $c \geq 2$)**.** *Even in the known-h model, any MPC-assisted TFM that simultaneously satisfies Bayesian UIC, Bayesian MIC, and Bayesian SCP for $c \geq 2$ must suffer from 0 expected social welfare for the users under a bid vector $\mathbf{b} \sim \mathcal{D}^\ell$ where $\ell > h$.*

Necessity of Bayesian equilibrium.

All of our feasibility results, namely, Theorems 1.1.9 to 1.1.12, rely on a Bayesian notion of equilibrium (for the MIC and SCP guarantees). As argued in Section 1.1.1, the Bayesian notion of equilibrium is suitable for the MPC-assisted model since the users cannot observe others' bids before submitting their own.

We show that the reliance on Bayesian notions of equilibrium is necessary (see Section 5.5.2) — had we insisted on an *ex post* notion of equilibrium in the MPC-assisted model, our additional reasonable-world assumptions would not help us overcome the previously known impossibility results. More specifically, we show that any MPC-assisted mechanism that simultaneously achieves ex post UIC and SCP must suffer from zero miner revenue even in the known-h model. Similarly, for approximate but ex post notions of incentive compatibility, the same miner revenue limitation stated in Theorem 1.1.5 still applies even in the known-h model. Further, the above restrictions on miner revenue hold no matter whether the block size is finite or infinite.

Philosophical Discussions about Our Assumptions and Modeling

Known-h assumption. Our assumption about a known upper bound h on the number of honest users has the following justifications:

- First, as mentioned, the zero miner-revenue limitation in earlier works holds in a very strong sense, and even when we make a cryptographic style assumptions such as "a majority of the users or bids are honest" — see Section 5.5.3 for more details.
- Second, TFMs must work in an *open setting* where anyone can post a bid, and the mechanism or players do not know the number of bids in advance. This is also an important reason why TFMs depart from classical mechanism design. The precise assumption we need for circumventing the zero miner revenue limitation is an absolute lower bound h on the number of honest users. Simply assuming that the majority of the bids are honest is not sufficient (see Section 5.5.3).
- Finally, our definitional framework ensures that *honest behavior is an equilibrium*, and thus players are incentivized to behave honestly. This, in turn, reinforces the h-honest users assumption (as long as enough users show up). As mentioned earlier, the same philosophy has been adopted in a line of prior works at the intersection of game theory and cryptography [HT04, KN08, ADGH06, OPRV09, AL11, ACH11, GKM+13, GKTZ15, GTZ15, Kat08, DR07, GLR10, CGL+18, WAS22, CCWS21, PS17, KMSW22, FW20, EFW22].

Limited fake bids. Recall that besides the known-h assumption, our results that achieve strict incentive compatibility require an extra assumption that the number of fake bids injected by the coalition is limited. This assumption is motivated and justified by the following observations. To submit fake bids, the strategic player or coalition needs to have some coin or account with a

non-zero balance. Given that the strategic player has a limited initial budget, it cannot control infinitely many accounts. Moreover, if one posts multiple conflicting transactions double-spending the same coin or units of currency, they can easily be detected and suppressed.

As mentioned, we currently do not know whether this limited fake bids assumption can be removed while still achieving strict incentive compatibility. We pose this as an open question.

Robustness w.r.t. parameter estimation. Among our proposed mechanisms, the ones that achieve approximate incentive compatibility, namely, threshold-based or diluted threshold-based mechanisms are simpler and more practical. Just like how Ethereum's EIP-1559 needs to estimate a suitable base fee, these mechanisms also need to estimate some parameters a-priori. In particular, our (diluted) threshold-based mechanism needs to know an estimate of h and the median of the value distribution $\mathcal{D}$ in advance. Just like Ethereum's EIP-1559, we can estimate these parameters from recent history. For example, one can estimate the total number of bids n from the degree of congestion observed in recent blocks. Now, if we are willing to assume that half of the n anticipated bids are honest (note that our mechanisms incentivize honest behavior), we can get an estimate of h. Similarly, one can estimate the median of the distribution $\mathcal{D}$ from the recent history too.

One important observation is that our threshold-based mechanism and diluted threshold-based mechanisms are quite robust to errors in the estimates. As mentioned later in Remark 5.2.1, if we set the threshold to $\widehat{h}/4$ for some estimated $\widehat{h}$, and let $\widehat{m}$ be the estimated median, then the mechanisms will achieve approximate incentive compatibility as long as $h_{\text{real}} \cdot q_{\text{real}} \geq \widehat{h} \cdot \frac{(1+\delta)}{4}$ for some arbitrarily small constant $\delta > 0$, where h_{real} is the actual number of honest users, and q_{real} is the actual percentile of the estimate $\widehat{m}$. For example, if $h_{\text{real}} = 0.6h$, and $q_{\text{real}} = 40\%$, then our mechanisms still satisfy approximate incentive compatibility for an exponentially small ϵ.

1.1.4 MUCP

The previous results show that strict incentive compatibility significantly constrains users' social welfare when $c \geq 2$, i.e., when the miners may collude with two or more users. Even in the known-h model, where miners receive positive revenue, users' social welfare is still limited to 0 when the number of users exceeds h. Although we can circumvent this limitation with approximate incentive compatibility, the dilution-based mechanisms do not always guarantee optimal user social welfare. This observation prompts the question of whether the notion of SCP might be overly stringent. Therefore, we ask the following question:

Is there a meaningful incentive compatibility notion for miner-user coalitions, under which we can achieve positive user social welfare for finite block size?

Define Miner-User Coalition Proofness

Defining a meaningful notion to capture incentive compatibility for miner-user coalitions is quite subtle. Previous work [Rou20, Rou21] considered another notion called off-chain agreement proofness (OCA-proofness), which, roughly speaking, requires that there exists a bidding strategy for all users to maximize global joint utility. [CRS24] introduces global SCP, which in-

tuitively captures the requirement that strategic users and miners cannot extract benefits from the protocol. Unfortunately, [CRS24] proved certain impossibility results pertaining to OCA-proofness and global SCP.

In this thesis, we consider a notion of a different flavor called miner-user coalition proofness (MUCP). MUCP captures the intuition that any miner-user coalition is unstable, thus disincentivizing the formation of a miner-user coalition. Roughly speaking, MUCP in the plain model (MUCP in the MPC-assisted model can be defined analogously) guarantees that for any coalition $\mathcal{C}$ consisting of the miner $\mathcal{M}$ and a non-empty set of users $\mathcal{U}$, either $\mathcal{M}$ alone has a better strategy or $\mathcal{U}$ alone has a better strategy. Specifically, for any honest users' bid $\mathbf{b}_{-\mathcal{C}}$, for any strategy $S_{\mathcal{C}}$, either $\mathcal{M}$ has a miner-only strategy $S_{\mathcal{M}}$ such that for any strategic bids $\mathbf{b}_{\mathcal{U}}$ from $\mathcal{U}$,

$$\mathsf{util}^{\mathcal{M}}((\mathbf{b}_{\mathcal{U}}, \mathbf{b}_{-\mathcal{C}}), S_{\mathcal{M}}) \geq \mathsf{util}^{\mathcal{M}}(\mathbf{b}_{-\mathcal{C}}, S_{\mathcal{C}}),$$

where the left-hand side represents the miner's utility when the users bid $(\mathbf{b}_{\mathcal{U}}, \mathbf{b}_{-\mathcal{C}})$ and the miner adopts strategy $S_{\mathcal{M}}$, and the right-hand side denotes the miner's utility when the coalition $\mathcal{C}$ adopts strategy $S_{\mathcal{C}}$; or $\mathcal{U}$ has a user-only strategy $S_{\mathcal{U}}$ such that

$$\mathsf{util}^{\mathcal{U}}(\mathbf{b}_{-\mathcal{C}}, S_{\mathcal{U}}) \geq \mathsf{util}^{\mathcal{U}}(\mathbf{b}_{-\mathcal{C}}, S_{\mathcal{C}}),$$

where the left-hand side represents the user's utility when $\mathcal{U}$ adopts $S_{\mathcal{U}}$ while the miner behaves *honestly*, and the right-hand side denotes $\mathcal{U}$'s utility when the coalition $\mathcal{C}$ adopts strategy $S_{\mathcal{C}}$; We say $S_{\mathcal{M}}$ is a *defecting strategy* for $\mathcal{M}$, or $S_{\mathcal{U}}$ is a *defecting strategy* for $\mathcal{U}$ w.r.t. $S_{\mathcal{C}}$.

Philosophical Discussion about MUCP

Unstable strategies. Using unstable strategies to capture incentive compatibility for the miner-user coalition is motivated by several factors. This approach mainly addresses the limitations imposed by SCP, which tends to rule out certain "benign" attacks. To see this, let us revisit why MPC-assisted posted price with random selection does not satisfy 2-SCP (Section 1.1.1). When there is congestion, a coalition $\mathcal{C}$ containing user i with a small true value and a user j with a very big true value may perform the following strategy to increase their joint utility: user i drops off to increase the chance of user j getting confirmed. They can split the joint gain off-chain.

However, one may recognize that this attack does not "steal" utility from users outside the coalition. In fact, every honest user's expected utility increases when coalition $\mathcal{C}$ performs such a dropping-off strategy because every honest user now gets a higher chance of getting confirmed. Moreover, we may not always prevent such kinds of attacks in practice: Users with multiple transactions should be able to decide their bidding strategy, including whether to drop off one bid for the current block, based only on their private true values. In addition, a user-user coalition is usually harder to form in a blockchain environment since users are ephemeral. Rendezvous between them is, therefore, challenging. Hence, we need a reasonable notion which allows user-user coalition to perform certain "benign" attacks.

Another reason is that such attacks do not involve the miners' power. A coalition consisting of only user i and user j is sufficient to perform such dropping-off attacks without the help of miners, i.e., users do not need to collude with the miners to gain benefits. Consequently, this is enough to prevent miners from gaining undue benefits through collusion with users and prevent miners from soliciting advantageous off-chain transfers from different users.

Connection Between c-Dominant-Strategy MIC and c-MUCP The definition of MUCP introduces an interesting asymmetry: when we state that miner coalition $\mathcal{M}$ has a defecting strategy, user coalition $\mathcal{U}$ may engage in strategic bidding, whereas when the user coalition $\mathcal{U}$ adopts a defecting strategy, the miner is expected to behave honestly. However, this definition alone lacks coherence, as it does not guarantee the miner's honesty.

To address this, we consider MUCP in conjunction with the following MIC requirements: even if the miner sees the strategic bid of $\mathcal{U}$, being honest is still the best strategy for the miner. We call this *c-dominant-strategy MIC*. Note that this notion is the same as MIC in an ex post setting but is stronger than Bayesian MIC. We make this modification to account for the scenario where the miners learn information about the true values of users in $\mathcal{U}$ when they attempt to negotiate a smart contract to form a miner-user coalition. Although MUCP guarantees that neither the users nor the miners have incentives to finally form a coalition, the miners may learn information during negotiation. We want to guarantee that even if the miners learn additional information during negotiation, being honest is the best strategy for the miners. Therefore, when discussing MUCP, we always consider it in conjunction with c-dominant-strategy MIC.

Results under MUCP

Table 1.4 below summarizes the landscape of TFMs satisfying UIC, c-dominant-strategy MIC, and c-MUCP for finite block size under the plain and the MPC-assisted model. In the table below, ✗ represents impossibility. $\Theta(\cdot)$ or $\Theta_{\mathcal{D}}(\cdot)$ implies matching upper- and lower bounds, where $\Theta_{\mathcal{D}}$ hides constant terms depending on users' true value distribution $\mathcal{D}$.

	Plain model	MPC-assisted model
Without reasonable-world assumption	✗	0-miner revenue $\Theta_{\mathcal{D}}(k)$-social welfare
With reasonable-world assumption	✗	$\Theta(h)$-miner revenue $\Theta_{\mathcal{D}}(k)$-social welfare

Table 1.4: Landscape for TFMs satisfying UIC, c-dominant-strategy MIC, and c-MUCP for finite block size under the plain model and the MPC-assisted model.

MUCP does not help us improve miner revenue or circumvent impossibility results in the plain model. In fact, we show that UIC + 1-dominant-strategy MIC + 1-MUCP implies 1-SCP. Together with the previous impossibility results under SCP, we have the following impossibility results.

Theorem 1.1.14 (Informal: impossibility under MUCP)**.** *No non-trivial TFM satisfies UIC,* 1*-dominant-strategy MIC, and* 1*-MUCP in the plain model for finite block size, whether with or without the reasonable-world assumption.*

Any (possibly randomized) TFM satisfying Bayesian UIC, 1*-dominant-strategy Bayesian MIC, and* 1*-Bayesian MUCP must suffer from* 0*-miner revenue. In addition, in the known-*h *mode, no MPC-assisted TFM satisfying these three properties achieves expected miner revenue more than* $h \cdot \mathbb{E}[\mathcal{D}]$*, where* $\mathcal{D}$ *is the distribution of users' true values.*

Although MUCP does not help improve miner revenue, it does help us circumvent the limitation on social welfare. Specifically, we show that MPC-assisted, posted price auction with random selection satisfies c-MUCP against miner-user coalitions consisting of ρ fraction of miners and no more than c users for arbitrary $\rho \in (0, 1]$ and $c \geq 1$. In particular, suppose all users' bids are sampled i.i.d. from some distribution $\mathcal{D}$, setting the reserved price to be the median m of the distribution gives us $k \cdot \mathbb{E}_{x \sim \mathcal{D}}[x | x \geq m]$ expected user social welfare, which is asymptotically optimal as long as there are enough users.

Theorem 1.1.15. *MPC-assisted, posted price auction with random selection (Section 1.1.1) satisfies ex post UIC, ex post MIC, and ex post (ρ, c)-MUCP for arbitrary $\rho \in (0, 1]$ and $c \geq 1$.*

By a similar reasoning, we show that MPC-assisted, LP-based mechanism also satisfies Bayesian c-MUCP against miner-user coalitions consisting of ρ fraction of miners and no more than c users for arbitrary $\rho \in (0, 1]$ and $c \geq 1$. As before, this holds with an additional reasonable-world assumption that there is at least h number of honest users and that the strategic coalition controls no more than d number of bids.

Theorem 1.1.16. *Suppose the block size is k, and honest users' true values are sampled i.i.d. from some distribution $\mathcal{D}$. If $d \leq \frac{1}{16}\sqrt{\frac{h}{2 \log h}}$, then the MPC-assisted, LP-based mechanism with random selection (Section 1.1.3) satisfies ex post UIC, c-dominant strategy Bayesian MIC, and Bayesian c-MUCP against miner-user coalition consisting of no more than ρ fraction of miners and no more than c number of users, for arbitrary $\rho \in (0, 1]$ and $c \geq 1$. In addition, the expected user social welfare is $k \cdot \mathbb{E}_{x \sim \mathcal{D}}[x | x \geq m]$ when there are enough users.*

We stress that in both mechanisms, the defecting strategy can be computed efficiently given the honest strategy and the coalition strategy $S_{\mathcal{C}}$.

1.1.5 Additional Related Work

We now review some closely related recent works besides the prior works on transaction mechanism design [LSZ19, Yao, BEOS19, BCD+, Rou20, Rou21, FMPS21, CS23] already mentioned.

SCP v.s. OCA-proof. Roughgarden [Rou21] introduces the concept of off-chain agreement (OCA) to capture the strategic behavior of a miner-user coalition. A mechanism is OCA-proof if there exists a bidding strategy (not necessarily the honest strategy) mapping users' true values to a bidding strategy that maximizes social welfare. This bidding strategy must follow specific criteria: 1) it does not modify the inclusion rule, 2) it satisfies individual rationality, and 3) each user independently submits only one bid without knowledge of other users' true values. As a comparison, SCP guarantees that the honest strategy maximizes the joint utility of *the coalition* rather than global utility. [CRS24] compared these two notions and demonstrated the impossibility of finding a mechanism simultaneously satisfying UIC, MIC, and OCA-proofness in the plain model for finite block size.

It is noteworthy that a mechanism satisfying SCP must also be OCA-proof, but the reverse is not necessarily true. Even if a mechanism consistently maximizes social welfare, a coalition comprising the miner and a subset of users might still gain from deviation by exploiting other honest users. The fact that miners can extract additional gain by exploiting honest users has been demonstrated in practice. For instance, since the miner has dictatorial control over its block, by

reordering the transactions in a block, miners or miner-user coalitions can gain from arbitrage, known as *miner extractable value (MEV)* [BDKJ23, KDC22, BCLL22, QZG22, ZQC+21, Zus, QZLG21, AEC21]. Given these considerations, this thesis primarily focuses on SCP.

TFM in a Bayesian setting. The recent works of Zhao, Chen, and Zhou [ZCZ22] and Gafni and Yaish [GY22] both consider TFM in a Bayesian setting. Although their works did not explicitly define the MPC-assisted model, from a practical standpoint, their results are in fact only relevant in an MPC-assisted (or a similar) model. As explained in Section 1.1.1, plain-model TFMs that achieve *Bayesian* equilibrium also achieve *ex post* equilibrium, since in the plain-model game, the strategic player can decide its actions *after* having observed honest users' bids.

Gafni and Yaish [GY22] suggest a mechanism that satisfies Bayesian UIC, while also satisfying MIC and OCA-proof (short for offchain-agreement-proof) even if the miner knows everyone's bid. Further, their mechanism works in the finite-block setting while achieving asymptotical optimality in social welfare and revenue. We stress that their result does not contradict our zero miner-revenue limitation Theorem 1.1.1 since their OCA-proofness notion (originally defined by Roughgarden [Rou20, Rou21]) is of a different nature from our side-contract-proofness (SCP) notion (originally defined by Chung and Shi [CS23]). Roughly speaking, OCA-proofness requires that a strategic coalition cannot enter an off-chain contract that increases *everyone*'s utility (*not just those in the coalition*) relative to what's achievable on-chain. In comparison, SCP is the notion that directly captures the cryptocurrency community's outpouring concerns about Miner Extractable Value (MEV). In particular, middleman platforms such as Flashbot facilitate the collusion of miners and users, where the coalition plays strategically to profit themselves at the expense of other users. This is why we choose to use the SCP notion rather than OCA-proofness. Moreover, the reason why the cryptocurrency community is developing encrypted mempool techniques (which can be viewed as instantiations of the MPC-assisted model) is also because they care about SCP (i.e., resilience to MEV).

Zhao, Chen, and Zhou [ZCZ22] suggest a mechanism that generates positive miner revenue while achieving Bayesian UIC and Bayesian 1-SCP even for the finite block setting. Their result does not contradict the 0-miner revenue limitation of [SCW23], since Zhao, Chen, Zhou [ZCZ22] consider only a restricted strategy space. In their work, a strategic user or a miner-user coalition can only deviate by bidding untruthfully; the coalition cannot inject fake bids, strategic users cannot drop out, and nor can strategic miners alter the inclusion rule. Due to their restricted strategy space, their results are only relevant under very stringent assumptions: 1) the TFM is implemented in the MPC-assisted (or similar) model; 2) the TFM is fully "permissioned" and allows only a set of pre-registered users to submit bids. In particular, the latter "permissioned" requirement is unrealistic for major decentralized cryptocurrencies today where any user can join and submit transactions.

Auctioneer deviation. Akbarpour and Li [AL20] introduced *credible auctions*, where users only communicate with the auctioneer through private channels. A credible auction requires that the auctioneer has no incentive to perform safe deviations: those deviations that would not be detected by the users. Akbarpour and Li [AL20] presented a trilemma result, establishing that no optimal auction can simultaneously satisfy credibility, efficiency (in terms of communication

rounds), and truthful bidding for users. Although credible auctions consider a similar concept to MIC, there are significant differences between these two concepts. See [CS23] for an in-depth discussion.

User-user collusion. Traditional auctions, such as the Vickrey auction, do not satisfy incentive compatibility against user-user collusion. Consequently, a line of research [GL79, GH05, CM12, kCK09, MM12, DM17] delves into incentive compatible auction designs against user-user collusion. In their elegant work, Goldberg and Hartline [GH05] demonstrated that only the class of posted price auctions and their utility-equivalent variations can successfully resist user-user coalitions.

Interestingly, the study of Transaction Fee Mechanisms (TFM) has not focused primarily on user-user coalitions. As we discussed in the Section 1.1.4, forming a user-user coalition is inherently difficult. Therefore, our primary focus is to mitigate the potential exploitation of miner-user coalitions.

Fake transactions. In classical auction design, auctioneer injecting fake transactions is often referred to as *shill bidding* [GMR90, EW09, NB15], and users injecting fake transactions is sometimes termed *false name bids* [YSM01, AHV14, YSM04, Yok07, TMIY12] in the context of combinatorial auctions. Combinatorial auctions, which simultaneously auction multiple items with interdependent values, allow users to bid for combinations of items rather than individual items.

In this thesis, we assume users' transactions have uniform sizes, and the items being auctioned in transaction fee mechanisms are independent. The mechanisms presented here suggest that under this assumption, it appears that users injecting fake bids is less concerning compared to untruthful bidding. However, an intriguing open question remains: what advantages do users gain from injecting fake bids when TFMs are modeled as combinatorial auctions? Within this framework, users bid for different units of space to accommodate transactions of varying sizes, thus introducing additional complexities in exploring the feasibility of TFMs.

Cryptography meets game theory. Prior to the advent of cryptocurrencies, a line of work [HT04, KN08, ADGH06, OPRV09, AL11, ACH11, GKM+13, GKTZ15, GTZ15, Kat08, DR07, GLR10, CGL+18, WAS22, CCWS21, PS17, KMSW22, FW20, EFW22] investigated how cryptography and game theory can help each other. For example, cryptography can help remove the trusted mediator assumption in correlated equilibria [DR07]. Ferreira and Weinberg [FW20] and Essaidi, Ferreira, and Weinberg [EFW22] showed that cryptographic commitments can help us circumvent impossibilities pertaining to credible auctions. On the other hand, adopting game-theoretic fairness can allow us to circumvent lower bounds pertaining to the more stringent cryptographic notions of fairness [HT04, ADGH06, IML05, OPRV09, CGL+18, WAS22].

1.1.6 Organization

We will discuss the above results in detail in the subsequent sections. In Chapter 2, we will formally define models for TFM and incentive compatibilties. We will then discuss the results

of the MPC-assisted model, approximate incentive compatibility, reasonable-world assumption, and MUCP, respectively, in Chapter 3, Chapter 4, Chapter 5, and Chapter 6. The technical overview is given in the corresponding sections. Finally, in Chapter 7, we discuss how to instantiate the MPC-assisted model with real-world cryptography.

Chapter 2

Models and Definitions

A mechanism specifies the strategy space for each individual, an outcome function that maps these strategies to a social decision, and an intended honest strategy. In this thesis, we only consider direct revelation mechanisms, where each individual holds a private type, and the message space is the same as the type space. For the case of transaction fee mechanisms (TFMs), each user's type is its private *true value*, which captures the maximum amount it is willing to pay. We make the following assumptions on *honest* users' true values.

Assumption 2.1. *Let $\mathcal{D}$ denote the joint distribution of the number of honest users and their true values. We assume that $\mathcal{D}$ first samples the number of honest users, and then samples the true value of each honest user independently. Each honest user's true value is NOT necessarily sampled from the same distribution.*

The joint distribution $\mathcal{D}$ may or may not be known a-priori to the mechanism. Jumping ahead, all our impossibility results are proven based on this assumption and that $\mathcal{D}$ is known to the mechanism, while many of our constructions achieve desired incentive compatibility even if the users' true values are arbitrarily chosen.

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 to represent a vector. Throughout this thesis, we use k to denote the block size, i.e., the maximum number of bids that can be confirmed.

Given a distribution $\mathcal{D}$, we use the notation $\mathrm{Supp}(\mathcal{D})$ to denote its support. We use $\mathbb{R}^{\geq 0}$ to denote non-negative real numbers.

2.2 Transaction Fee Mechanism in the Plain Model

A transaction fee mechanism (TFM) in the plain model involves the miner of the current block and some users, the number of which may be unknown in a-priori. Henceforth, we use $\mathcal{C}$ to denote a coalition of strategic players or a strategic individual. In particular, $\mathcal{C}$ can represent a single user, the miner of the current block, or a coalition of the miner and some users.

In the plain model, a TFM describes the following game.

TFM in the plain model

1. Users not in $\mathcal{C}$ each submit a single bid, where each bid is represented by a single non-negative real value. Let $\mathbf{b}_{-\mathcal{C}}$ denote the resulting bid vector.
2. The coalition $\mathcal{C}$ observes $\mathbf{b}_{-\mathcal{C}}$, and then each user in $\mathcal{C}$ submits an arbitrary number of bids, each represented by a single non-negative real value. Let $\mathbf{b}_{\mathcal{C}}$ denote the resulting bid vector from coalition $\mathcal{C}$.
3. The miner of the current 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 outputs 1) which are confirmed, 2) how much each confirmed bid pays, and 3) how much revenue is paid to the miner.

Figure 2.1: Transaction fee mechanism in the plain model

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 rule*: Given the at most k bids included in the block, the confirmation rule decides which ones to confirm.
- *Payment rule*: Given the confirmed bids, 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, whereas the confirmation, payment, and miner revenue rules are implemented by the blockchain. Thus, the latter three rules are guaranteed to be implemented honestly.

Strategy space. Each user has a non-negative private *true value*. An honest user submits a single bid that equals its true value using its identity; an honest miner does not submit any fake bid. Moreover, an honest miner always implements the prescribed inclusion rule. However, strategic players can choose to perform the following strategies:

- Strategic users may choose to submit zero to multiple bids, and these bids need not reflect their true value.
- A strategic miner can post fake bids, and pick an arbitrary set of up to k bids of its own 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.

Notably, strategic players in $\mathcal{C}$ can decide their actions *after* observing the bids of other players not in $\mathcal{C}$. In this thesis, we only consider a single strategic player/coalition when talking about incentive compatibility.

Symmetry. 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.

2.3 Transaction Fee Mechanism in the MPC-Assisted Model

In the MPC-assisted model, instead of having a single miner implement the mechanism, a group of miners is selected to jointly implement the mechanism. This group of miners jointly run a multi-party computation (MPC) protocol that implements the TFM. Figure 2.2 describes the ideal functionality (denoted $\mathcal{F}_{\text{MPC}}$) realized by the MPC protocol. The ideal functionality is parametrized with the allocation, payment, and miner revenue rules of a TFM.

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

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 2.2: Ideal functionality realized by the MPC protocol.

The MPC protocol achieves full security with guaranteed output as long as a 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 implements $\mathcal{F}_{\text{MPC}}$ — this is often referred to as the $\mathcal{F}_{\text{MPC}}$-hybrid model. The instantiation that securely realizes $\mathcal{F}_{\text{MPC}}$ is given in Chapter 7.

A TFM in the MPC-assisted model describes the following game:

TFM in the MPC-assisted model

1. Every player (i.e., user or miner) can take on *zero to multiple* identities, and every identity submits a single bid represented by a non-negative real value to $\mathcal{F}_{\text{MPC}}$ defined in Figure 2.2.
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.

Figure 2.3: Transaction fee mechanism in the MPC-assisted model

Therefore, to specify a TFM in the MPC-assisted model, we only need to specify the allocation rule, the payment rule, and the miner revenue rule, which are *possibly randomized*, polynomial-time algorithms, and the syntax of the rules are evident from $\mathcal{F}_{\text{MPC}}$ in Figure 2.2. 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, 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 submit any bids. However, a strategic miner can take on one or more identities and submit fake bids. Here, unlike the plain model, a strategic miner can no longer choose which bids to include in the block — the allocation rule (i.e., the counterpart of the inclusion and confirmation rules of the plain model) is enforced by $\mathcal{F}_{\text{MPC}}$.

One technicality is whether the distribution of users' identities matters, 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 impossibility results stronger.

2.4 Defining Incentive Compatibility

Utility. Every user i has a private true value $v_i \in \mathbb{R}^{\geq 0}$. If user i's transaction is confirmed and the user pays p_i, then its utility is defined as $v_i - p_i$. Otherwise, its utility is 0. 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* contract (e.g., decentralized smart contracts) to split off their gains off the table.

Ex post incentive compatibility. Roughly speaking, 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. In our formal definitions below, we define the *approximate* case that allows ϵ slack. Strict incentive compatibility can be achieved when $\epsilon = 0$.

Definition 2.4.1 (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 true value v_i of user i, for any bid vector $\mathbf{b}_{-i}$ of all users other than i, no strategy can increase the user's expected utility by more than ϵ compared to the honest behavior.
- *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[1] 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. When the strategic players have some a-priori knowledge about honest users' bid distribution $\mathcal{D}$, we may define the Bayesian notion of incentive compatibility. 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.

[1] 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.

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 of $\mathbf{b}_{-\mathcal{C}}$. 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 2.4.2 (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 $\mathcal{C}$, 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 $\mathbf{b}_{-\mathcal{C}}$ sampled from $\mathcal{D}_{-\mathcal{C}}$, 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,

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

  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 $\mathbf{b}$.
- *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,

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

  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 $\mathcal{C}$),

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

Note that in the plain model, the strategic individual or coalition can divide its strategy *after* having observed the remaining honest users' bids. Therefore, it is not reasonable to consider the Bayesian notion of incentive compatibility in the plain model. 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, while all of our mechanism designs in the MPC-assisted model work for the ex post notions. This makes both our lower and upper bounds stronger.

Chapter 3

Characterization of Strict Incentive Compatibility in MPC-Assisted Model

One may hope that with the Bayesian notion of incentive compatibility and MPC-assisted model, we can achieve positive miner revenue. Unfortunately, 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. When $c = 1$, i.e., the miner colludes with only one user, we can indeed circumvent the previous impossibility result and give a mechanism that achieves all desired properties Section 3.3. For $c \geq 2$, however, it is still impossible to build a dream mechanism Section 3.4. In this chapter, we present the formal results (Table 1.1) for characterizing strict incentive compatibility in the MPC-assisted model. [Ke: n is the number of users and N is the number of bids]

3.1 Technical Overview

To prove the impossibility results, we make the following assumption on users' true value distributions: We assume that honest users' true values are sampled *independently* from some bounded continuous distribution. We use $\mathcal{D}_i$ to denote user i's true value distribution and $\widehat{\mathcal{D}} := \mathcal{D}_1 \times \cdots \times \mathcal{D}_n$ to denote the joint distribution of users' true values. For any fixed user i, we also use $\mathcal{D}_{-i} := \mathcal{D}_1 \times \cdots \times \mathcal{D}_{i-1} \times \mathcal{D}_{i+1} \times \cdots \times \mathcal{D}_n$ to denote the distribution of other users' true values. Without loss of generality, we assume that for any i, distribution $\mathcal{D}_i$ has the same support $[0, M]$ for some M. Jumping ahead, we will use these assumptions of distributions to prove impossibility results, whereas our TFM design achieves incentive compatibility for any distribution.

Technical Overview: 0-Miner Revenue

To prove the limit on the miner revenue, we use the following simplified notations to represent a mechanism since we only care about the probability of each bid being confirmed, the expected payment of each bid, and the miner revenue. Let n denote the number of users and N denote

the number of bids received by the mechanism. Henceforth in this chapter, we use $(\mathbf{x}, \mathbf{p}, \mu)$ to represent a TFM in the MPC-assisted model:

- **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.

For the i-th user, we define

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

The crux of the proof is to show that the miner revenue should not change when we lower one user's bid down to 0. Then applying this argument for n times, we get the desired impossibility result. For simplicity, in the overview, we omit the ρ factor.

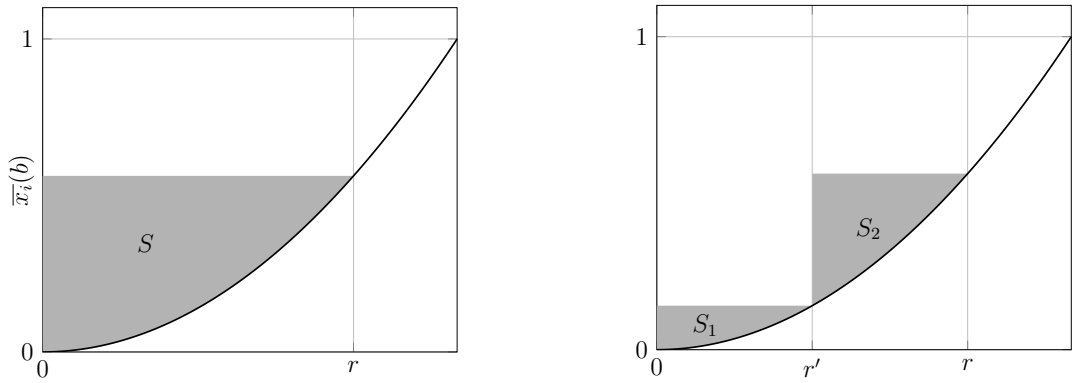
To quantify how much the miner revenue changes when one user lowers its bid to 0, we use the Myerson's Lemma [Mye81]. Myerson's lemma implies that for a Bayesian UIC TFM, it must be that for any user i, the allocation rule $\overline{x_i}(\cdot)$ must be monotone and that the expected payment $\overline{p_i}(\cdot)$ is specified as

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

Imagine that user i's true value is 0, but it bids r instead. In this case, the user's utility loss is represented by the size of the gray area S in Figure 3.1a. Due to SCP, the expected miner revenue increase when user i bids r instead of 0 must be at most S. To further bound the miner revenue change, we split this process into 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 3.1b. By SCP, we conclude that $\overline{\mu_i}(r') - \overline{\mu_i}(0) \leq S_1$. Now, imagine user i's true value is r' but it bids r instead. Using a similar argument, we conclude that $\overline{\mu_i}(r) - \overline{\mu_i}(r') \leq S_2$ (see Figure 3.1b). Summarizing the above, we have that $\overline{\mu_i}(r) - \overline{\mu_i}(0) \leq S_1 + S_2$.

Following a similar argument, if we take infinitely many steps, we get that $\overline{\mu_i}(r) - \overline{\mu_i}(0) = 0$, i.e., the expected miner revenue should not change if we lower one user's bid down to 0. Therefore, we have $\mathbb{E}_{\mathbf{b} \sim \mathcal{D}}[\mu(\mathbf{b})] = \mathbb{E}_{\mathbf{b}_{-n} \sim \mathcal{D}_{-n}}[\mu(\mathbf{b}_{-n}, 0)]$. Due to MIC, if we remove this 0 bid from the bid vector, the miner revenue should not decrease. Otherwise, when there are $n-1$ users whose true values are sampled from $\mathcal{D}_{-n}$, the miner coalition can inject a fake 0 bid to increase its revenue. Repeating this argument for n times, we get our desired impossibility result that $\mathbb{E}_{\mathbf{b} \sim \mathcal{D}}[\mu(\mathbf{b})] = 0$. This holds even if we assume an infinite block size, and the users do not inject fake bids.

In the proof described above, we make use of the assumption that the distribution is continuous, based on which we can take infinitely many steps to bound the change in miner revenue when we lower one bid down to 0. If the value distribution is discrete, then the miner can get some positive revenue upper-bounded by the distance of two neighboring points in the support. See Remark 3.2.6 for a more detailed discussion. The formal proof of 0-miner revenue result is given in Section 3.2.



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

(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 3.1: User's utility change when untruthful bidding.

Achieving Incentive Compatibility in MPC-assisted Model: Finite Block Size

With the power of the MPC-assisted model, we have a mechanism that achieves ex post UIC, ex post ρ-MIC, and ex post $(\rho, 1)$-SCP when the block size is k. This circumvents the previous impossibility result in [CS23]. The mechanism is simply a burning posted price with random selection: The mechanism is parametrized with a fixed reserve price r. It randomly confirms k bids that are at least r to confirm. Each confirmed bid pays r, and all payments are burnt.

The incentive compatibility of this mechanism is easy to see and we give the formal proof in Section 3.3. However, for the above mechanism to achieve (ρ, c)-SCP for $c \geq 2$, we will have to set the posted price to M, the maximum value of the users' true value distributions. Otherwise if $r < M$, there exists a scenario where miners colluding with 2 users have a strategy to increase their joint gain: Imagine that the miners collude with user i and j, where user i's true value is r and user j has a sufficiently large true value. In this scenario, user i may choose not to bid to increase the probability of user j being confirmed.

Necessity of Zero Social Welfare

It turns out that for a TFM to be ex post UIC, ex post MIC, and ex post $(\rho, 2)$-SCP, the social welfare (sum of everyone's utility) must be 0. To prove this, we use the following framework:

1. We first prove in Lemma 3.4.1 that for a TFM to satisfy the above three incentive compatibility, it must be that for any two users i and j, when user j changes its bid, user i's utility should not change. This proof has a similar flavor as the proof that a miner's revenue should not change when one user's bid changes.
2. Next, we prove in Lemma 3.4.2 that for any two users i and j, user i's utility should not change if user j drops off.
3. Step 2 implies that for any user i with true value v_i sampled from $\mathcal{D}_i$, its expected utility $\mathbb{E}_{v_i \sim \mathcal{D}_i}[\mathsf{util}^i(v_i)]$ when it is the only bidder in the world is the same as its expected utility $\mathbb{E}_{(v_i, \mathbf{b}_{-i}) \sim \widehat{\mathcal{D}}}[\mathsf{util}^i(v_i, \mathbf{b}_{-i})]$ when there are other users whose true values are sampled from $\mathcal{D}_{-i}$.

Here, $\mathsf{util}^i(\mathbf{b})$ represents user i's utility when the bid vector is $\mathbf{b}$. This can be proved by repeatedly applying Step 2 to remove other users' bids one by one from the latter world.

4. Now we are ready to prove that each user's utility must be 0. Imagine a world with an arbitrarily large number of bids. There must exist some user j^* whose utility is 0 in this crowded world. Then by Step 3, user j^* gets utility 0 when it is the only bidder in the world. By symmetry of the TFM (see Section 2.2), any user has 0 utility when it is the single bidder. Applying Step 3 again, we have that every user must have utility 0 when the bid vector $\mathbf{b}$ is sampled from the joint distribution $\widehat{\mathcal{D}}$.

One technicality that arises in the full proof (see Section 3.4) is the usage of the weak symmetry assumption. In particular, the proof would have been much easier if we could instead assume *strong symmetry* which, unfortunately, is too stringent. In strong symmetry, we assume that any two users who bid the same amount must receive the same treatment. While it is a good approach for gaining intuition about the proof, it is too stringent since there could well be more bids offering the same value than the block size k — in this case, a non-trivial mechanism would treat them differently, i.e., confirm some while rejecting others. Our actual proof of needs only a *weak symmetry* assumption.

3.2 Necessity of Zero Miner Revenue

Chung and Shi [CS23] 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 $(\mathbf{b}_{-i}, b_i) = \mathbf{b}$.

Lemma 3.2.1 (Myerson's Lemma [Mye81])**.** *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 $\mathbb{E}_{\mathbf{b}_{-i}\sim\mathcal{D}_{-i}}[x_i(\mathbf{b}_{-i}, b_i')] \geq \mathbb{E}_{\mathbf{b}_{-i}\sim\mathcal{D}_{-i}}[x_i(\mathbf{b}_{-i}, b_i)]$.*
2. *The payment rule $\mathbf{p}$ is defined as follows. For any user i, and bid b_i from user i, it must be*

$$\mathop{\mathbb{E}}_{\mathbf{b}_{-i}\sim\mathcal{D}_{-i}}[p_i(\mathbf{b}_{-i}, b_i)] = \mathop{\mathbb{E}}_{\mathbf{b}_{-i}\sim\mathcal{D}_{-i}}\left[b_i \cdot x_i(\mathbf{b}_{-i}, b_i) - \int_0^{b_i} x_i(\mathbf{b}_{-i}, t)dt\right]. \tag{3.1}$$

The following technical lemma was given in [CS23].

Lemma 3.2.2 (Lemma 4.8 in [CS23])**.** *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*

$$r \cdot (\overline{x_i}(r') - \overline{x_i}(r)) \leq \pi(r') - \pi(r) \leq r' \cdot (\overline{x_i}(r') - \overline{x_i}(r)),$$

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

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 3.2.3. *Let $\widehat{\mathcal{D}} = \mathcal{D}_1 \times \cdots \times \mathcal{D}_n$ 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*

$$\overline{\mu_i}(b) = \overline{\mu_i}(0). \tag{3.2}$$

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

$$r \cdot (\overline{x_i}(r') - \overline{x_i}(r)) \leq \widetilde{\pi}(r') - \widetilde{\pi}(r) \leq r' \cdot (\overline{x_i}(r') - \overline{x_i}(r)). \tag{3.3}$$

Therefore, we have the following two inequalities:

$$r \cdot [\overline{x_i}(r') - \overline{x_i}(r)] \leq \widetilde{\pi}(r') - \widetilde{\pi}(r),$$
$$r \cdot [\overline{x_i}(r') - \overline{x_i}(r)] \geq \widetilde{\pi}(r') - \widetilde{\pi}(r).$$

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 Bayesian UIC and satisfies Myerson's Lemma. Therefore, 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 Bayesian UIC, its payment rule itself must also satisfy the same expression (Eq.(3.1)), that is,

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

This implies 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 3.2.3 holds even if users do not inject any fake bids. If, in addition, the mechanism $(\mathbf{x}, \mathbf{p}, \mu)$ is Bayesian MIC, then the total miner revenue must be 0.

Theorem 3.2.4. *Let $\widehat{\mathcal{D}} = \mathcal{D}_1 \times \cdots \times \mathcal{D}_n$ be the joint distribution of the first 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*

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

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

Lemma 3.2.5. *If $(\mathbf{x}, \mathbf{p}, \mu)$ is Bayesian MIC against ρ-sized miner coalition, then $\mathop{\mathbb{E}}_{\mathbf{b}\sim\widehat{\mathcal{D}}}[\mu(\mathbf{b})] \leq \mathop{\mathbb{E}}_{\mathbf{b}'\sim\mathcal{D}_{-n}}[\mu(\mathbf{b}')]$, where $\mathcal{D}_{-n} = \mathcal{D}_1 \times \cdots \times \mathcal{D}_{n-1}$.*

For now, assume Lemma 3.2.5 holds and we explain why Theorem 3.2.4 follows from it. The proof of Lemma 3.2.5 appears right afterward. By induction on n, we have that

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

By Lemma 3.2.3, for any $b^* \in \mathrm{Supp}(\mathcal{D}_1)$, it should be that $\mu(b^*) = \mu(0)$. Therefore,

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

where the last equality comes from the requirement that the miner's revenue cannot exceed the payment of the single identity, who will pay at most what it bids. Theorem 3.2.4 thus follows. □

Proof of Lemma 3.2.5 Since $(\mathbf{x}, \mathbf{p}, \mu)$ is Bayesian MIC against ρ-sized miner coalition, it must be that,

$$\mathop{\mathbb{E}}_{\mathbf{b}\sim\mathcal{D}_{-n}}[\rho\mu(\mathbf{b}, 0)] \leq \mathop{\mathbb{E}}_{\mathbf{b}\sim\mathcal{D}_{-n}}[\rho\mu(\mathbf{b})]. \tag{3.4}$$

Otherwise, the miners can inject a 0 and increase the miner revenue while it does not need to pay anything for injecting the 0-bid. This violates the MIC condition.

Let $f(\cdot)$ denote the p.d.f. of distribution $\mathcal{D}_n$. By the law of total expectation, we have that

$$\begin{aligned}
\mathop{\mathbb{E}}_{\mathbf{b}\sim\widehat{\mathcal{D}}}[\mu(\mathbf{b})] &= \int_0^{+\infty} \mathop{\mathbb{E}}_{\mathbf{b}'\sim\mathcal{D}_{-n}}[\mu(\mathbf{b}', r)] f(r) dr \\
&= \int_0^{+\infty} \mathop{\mathbb{E}}_{\mathbf{b}'\sim\mathcal{D}_{-n}}[\mu(\mathbf{b}', 0)] f(r) dr && \text{By Lemma 3.2.3} \\
&= \mathop{\mathbb{E}}_{\mathbf{b}'\sim\mathcal{D}_{-n}}[\mu(\mathbf{b}', 0)] \leq \mathop{\mathbb{E}}_{\mathbf{b}'\sim\mathcal{D}_{-n}}[\mu(\mathbf{b}')] && \text{By (3.4)}
\end{aligned}$$

Lemma 3.2.5 thus follows.

Remark 3.2.6 (On continuity of the distribution)**.** The formal proof of Lemma 3.2.3 makes use of the fact that the distribution is continuous, i.e., the allocation rule and the payment rule are well-defined in the interval $[0, r]$ to derive the formula of $\widetilde{\Pi}(r)$. If the distribution is instead discrete, then the miner can get some positive revenue. For example, imagine that the support of each $\mathcal{D}_i$ is the set of non-negative integers. Consider the following TFM, which is a posted price auction with reserved price $r = 1$, and the miner gets revenue 1 is there is any bid confirmed. This mechanism satisfies ex post UIC, ex post MIC, and ex post c-SCP. However, such TFMs are tailored to the structure of the actual distribution, so we do not consider these cases in this thesis.

3.3 Feasibility for $c = 1$: Finite Block Size

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, burning 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.

Figure 3.2: MPC-assisted, burning posted price auction with random selection. Here the 0-miner revenue is inevitable based on Theorem 3.2.4

Theorem 3.3.1. *Assuming a finite block size k. The above MPC-assisted, posted price auction with random selection in the MPC-assisted model satisfies ex post UIC, ex post MIC, and ex post $(\rho, 1)$-SCP 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.

3.4 Impossibility for $c \geq 2$: Finite Block Size

Unfortunately, even in the MPC-assisted model, no mechanism with non-trivial utility achieves 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[1], since the miner revenue has to be 0 by Theorem 3.2.4 of Section 3.2. 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 3.4.1. *Given any (possibly random) mechanism in the MPC-assisted model that is Bayesian UIC and Bayesian SCP against $(\rho^*, 2)$-sized coalition for some $\rho^* \in (0, 1]$. Then, for any user i and user j, for any bid b_j and b'_j,*

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

where $\mathbf{b}_{-i,j}$ represents all except user i and user j's bids, and $\mathcal{D}_{-j}$ denotes the joint distribution of all except user j's true value.

Proof. In this proof, we use the following notations for simplicity. For any user i and j, we define the following notations:

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

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

$$\begin{aligned} &\text{Honest utility} = [v \cdot \widehat{x}_i(v, y) - \widehat{p}_i(v, y)] + [y \cdot \widehat{x}_j(v, y) - \widehat{p}_j(v, y)] + \rho\widehat{\mu}(v, y) \\ \geq &\text{Overbid utility} = [v \cdot \widehat{x}_i(v, z) - \widehat{p}_i(v, z)] + [y \cdot \widehat{x}_j(v, z) - \widehat{p}_j(v, z)] + \rho\widehat{\mu}(v, z). \end{aligned}$$

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 $z \geq y$. This will increase the expected joint utility of the coalition, which contradicts $(\rho^*, 2)$-SCP. For the same reason, if user j's true value is z, then

$$\begin{aligned} &\text{Honest utility} = [v \cdot \widehat{x}_i(v, z) - \widehat{p}_i(v, z)] + [z \cdot \widehat{x}_j(v, z) - \widehat{p}_j(v, z)] + \rho\widehat{\mu}(v, z) \\ \geq &\text{Underbid utility} = [v \cdot \widehat{x}_i(v, y) - \widehat{p}_i(v, y)] + [z \cdot \widehat{x}_j(v, y) - \widehat{p}_j(v, y)] + \rho\widehat{\mu}(v, y). \end{aligned}$$

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

$$\begin{aligned} &v[\widehat{x}_i(v, z) - \widehat{x}_i(v, y)] + z[\widehat{x}_j(v, z) - \widehat{x}_j(v, y)] + \rho[\widehat{\mu}(v, z) - \widehat{\mu}(v, y)] \\ \geq &\widehat{p}_i(v, z) - \widehat{p}_i(v, y) + \widehat{p}_j(v, z) - \widehat{p}_j(v, y) \\ \geq &v[\widehat{x}_i(v, z) - \widehat{x}_i(v, y)] + y[\widehat{x}_j(v, z) - \widehat{x}_j(v, y)] + \rho[\widehat{\mu}(v, z) - \widehat{\mu}(v, y)] \end{aligned}$$

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

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

$$v \cdot \frac{\partial}{\partial z}\widehat{x}_i(v,z) + z \cdot \frac{\partial}{\partial z}\widehat{x}_j(v,z) + \rho\frac{\partial}{\partial z}\widehat{\mu}(v,z) = \frac{\partial}{\partial z}\widehat{p}_i(v,z) + \frac{\partial}{\partial z}\widehat{p}_j(v,z). \tag{3.5}$$

Note that Equation (3.5) should hold for at least two different values of $\rho \le \rho^*$. Hence, it must be that $\frac{\partial}{\partial z}\widehat{\mu}(v,z) = 0$. Equation (3.5) thus becomes

$$\begin{aligned} & v[\widehat{x}_i(v,z) - \widehat{x}_i(v,y)] + z[\widehat{x}_j(v,z) - \widehat{x}_j(v,y)] \\ \ge & \widehat{p}_i(v,z) - \widehat{p}_i(v,y) + \widehat{p}_j(v,z) - \widehat{p}_j(v,y) \\ \ge & v[\widehat{x}_i(v,z) - \widehat{x}_i(v,y)] + y[\widehat{x}_j(v,z) - \widehat{x}_j(v,y)]. \end{aligned} \tag{3.6}$$

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

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

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

$$\begin{aligned} \mathop{\mathbb{E}}_{v\sim\mathcal{D}_i}[\text{i-gain}(v, y \to z)] &\le \mathop{\mathbb{E}}_{v\sim\mathcal{D}_i}[\text{j-loss}(v, y \to z)] \\ &\le \mathop{\mathbb{E}}_{v\sim\mathcal{D}_i}[(z-y)(\widehat{x}_j(v,z) - \widehat{x}_j(v,y))]. \end{aligned}$$

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

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

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

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

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

Intuitively, 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 $\mathcal{D}_i$. 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 3.4.2. *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,*

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

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

We first give an intuition of why this lemma is true. By Lemma 3.4.1, we have that

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

Therefore, to prove Lemma 3.4.2, it suffices to prove that

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

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 2.2.

Strong symmetry assumption. On top of our current symmetric assumption defined in Section 2.2, 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.

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,

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

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,

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

Otherwise, user i can inject a 0-bid using an arbitrary identity i', which strictly increases its utility. This contradicts Bayesian UIC. Next, we give the formal proof of Lemma 3.4.2.

Proof. By Lemma 3.4.1, for any $i, j, b_j, \ell \geq 1$, we have

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

Therefore, it suffices to prove that for any i, j, ℓ, we have

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

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

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

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

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

Now, suppose the miner samples a fake identity m. Over the choice of m, either $\Pr(\text{Good identity } m) \geq \frac{1}{2}$ or $\Pr(\text{Bad identity } m) \geq \frac{1}{2}$. If $\Pr(\text{Good identity } m) \geq \frac{1}{2}$, 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 $\Pr(\text{Bad identity } m) \geq \frac{1}{2}$. 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',

$$\mathop{\mathbb{E}}_{\mathbf{b}_{-i,j} \sim \mathcal{D}_{-i,j}} [\mathsf{util}^i(v_i', 0_m, \mathbf{b}_{-i,j})] \leq \mathop{\mathbb{E}}_{\mathbf{b}_{-i,j} \sim \mathcal{D}_{-i,j}} [\mathsf{util}^i(v_i', 0_j, \mathbf{b}_{-i,j})], \tag{3.7}$$

or with probability at least $1/4$ over the choice of m, there exists some v_i' such that

$$\mathop{\mathbb{E}}_{\mathbf{b}_{-i,j} \sim \mathcal{D}_{-i,j}} [\mathsf{util}^i(v_i', 0_m, \mathbf{b}_{-i,j})] > \mathop{\mathbb{E}}_{\mathbf{b}_{-i,j} \sim \mathcal{D}_{-i,j}} [\mathsf{util}^i(v_i', 0_j, \mathbf{b}_{-i,j})].$$

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 (3.7) is satisfied. Recall also, for v_i specifically, we have strict inequality, that is, $\mathop{\mathbb{E}}_{\mathbf{b}_{-i,j} \sim \mathcal{D}_{-i,j}} [\mathsf{util}^i(v_i, 0_m, \mathbf{b}_{-i,j})] < \mathop{\mathbb{E}}_{\mathbf{b}_{-i,j} \sim \mathcal{D}_{-i,j}} [\mathsf{util}^i(v_i, 0_j, \mathbf{b}_{-i,j})]$. Thus, $\mathop{\mathbb{E}}_{\mathbf{b}_{-j} \sim \mathcal{D}_{-j}} [\mathsf{util}^i(0_m, \mathbf{b}_{-j})] < \mathop{\mathbb{E}}_{\mathbf{b}_{-j} \sim \mathcal{D}_{-j}} [\mathsf{util}^i(0_j, \mathbf{b}_{-j})]$.

For every bad identity m that additionally satisfies Equation (3.7), there must exist some $i' \neq i$ and $i \neq j$, and some $b_{i'} > 0$, such that

$$\mathop{\mathbb{E}}_{\mathbf{b}_{-j,i'} \sim \mathcal{D}_{-j,i'}} [\mathsf{util}^{i'}(0_m, b_{i'}, \mathbf{b}_{-j,i'})] > \mathop{\mathbb{E}}_{\mathbf{b}_{-j,i'} \sim \mathcal{D}_{-j,i'}} [\mathsf{util}^{i'}(0_j, b_{i'}, \mathbf{b}_{-j,i'})] \tag{3.8}$$

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 $b_{i'}$, $\mathop{\mathbb{E}}_{\mathbf{b}_{j,i'} \sim \mathcal{D}_{-j,i'}} [\mathsf{util}^{i'}(0_m, b_{i'}, \mathbf{b}_{-j,i'})] \leq \mathop{\mathbb{E}}_{\mathbf{b}_{j,i'} \sim \mathcal{D}_{-j,i'}} [\mathsf{util}^{i'}(0_j, b_{i'}, \mathbf{b}_{-j,i'})]$. Therefore, it must be that for any $i' \neq i$ and $i \neq j$, $\mathop{\mathbb{E}}_{\mathbf{b}_{-j} \sim \mathcal{D}_{-j}} [\mathsf{util}^{i'}(0_m, \mathbf{b}_{-j})] \leq \mathop{\mathbb{E}}_{\mathbf{b}_{-j} \sim \mathcal{D}_{-j}} [\mathsf{util}^{i'}(0_j, \mathbf{b}_{-j})]$. Therefore, we have that

$$\mathop{\mathbb{E}}_{\mathbf{b}_{-j} \sim \mathcal{D}_{-j}} [\mathsf{USW}(0_m, \mathbf{b}_{-j})] < \mathop{\mathbb{E}}_{\mathbf{b}_{-j} \sim \mathcal{D}_{-j}} [\mathsf{USW}(0_j, \mathbf{b}_{-j})] \tag{3.9}$$

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 2.2, it must be that $\mathop{\mathbb{E}}_{\mathbf{b}_{-j} \sim \mathcal{D}_{-j}} [\mathsf{USW}(0_m, \mathbf{b}_{-j})] = \mathop{\mathbb{E}}_{\mathbf{b}_{-j} \sim \mathcal{D}_{-j}} [\mathsf{USW}(0_j, \mathbf{b}_{-j})]$, which contradicts Equation (3.9).

Let ℓ denote the length of $|\mathbf{b}_{-j}|$. Let i' be a user such that Equation (3.8) happens with a 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 (3.7). 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 (3.8) 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 $c = 2$. □

Lemma 3.4.3. *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,*

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

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,

$$\mathop{\mathbb{E}}_{(v_i, \mathbf{b}_{-i,j}) \sim \mathcal{D}_j} [\mathsf{util}^i(v_i, b_j, \mathbf{b}_{-i,j})] = \mathop{\mathbb{E}}_{(v_i, \mathbf{b}_{-i,j}) \sim \mathcal{D}_{-j}} [\mathsf{util}^i(v_i, \mathbf{b}_{-i,j})]. \tag{3.10}$$

To see this, by Lemma 3.4.2, we have

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

Next, we are going to show that

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

To see why this holds, note that the first equality follows from Lemma 3.4.1. The inequality comes from 2-SCP: Since by MIC, it must be that

$$\mathop{\mathbb{E}}_{\mathbf{b}_{-i,j}\sim\mathcal{D}_{-i,j}}[\mu(v_i, 0_j, \mathbf{b}_{-i,j})] \leq \mathop{\mathbb{E}}_{\mathbf{b}_{-i,j}\sim\mathcal{D}_{-i,j}}[\mu(v_i, \mathbf{b}_{-i,j})].$$

Therefore, it must be that for any v_i in the support of $\mathcal{D}_i$, we have $\mathop{\mathbb{E}}_{\mathbf{b}_{-i,j}\sim\mathcal{D}_{-i,j}}[\mathsf{util}^i(v_i, 0_j, \mathbf{b}_{-i,j})] \geq \mathop{\mathbb{E}}_{\mathbf{b}_{-i,j}\sim\mathcal{D}_{-i,j}}[\mathsf{util}^i(v_i, \mathbf{b}_{-i,j})]$. 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 (3.10) thus follows by taking expectation over $v_i \sim \mathcal{D}_i$.

Let $\mathcal{D}$ denote the joint distribution of users' true values and $f_i(\cdot)$ denote the p.d.f. of $\mathcal{D}_i$. By definition of expectation,

$$\begin{aligned}
&\mathop{\mathbb{E}}_{(v_i,\mathbf{b}_{-i})\sim\widehat{\mathcal{D}}}[\mathsf{util}^i(v_i, \mathbf{b}_{-i})] = \int_0^\infty \mathop{\mathbb{E}}_{(v_i,\mathbf{b}_{-i,1})\sim\mathcal{D}_{-1}}[\mathsf{util}^i(v_i, z_1, \mathbf{b}_{-i,1})] f_1(z_1) dz_1 \\
&= \int_0^\infty \mathop{\mathbb{E}}_{(v_i,\mathbf{b}_{-i,1})\sim\mathcal{D}_{-1}}[\mathsf{util}^i(v_i, \mathbf{b}_{-i,1})] f_1(z_1) dz_1 && \text{By Equation (3.10)} \\
&= \mathop{\mathbb{E}}_{(v_i,\mathbf{b}_{-i,1})\sim\mathcal{D}_{-1}}[\mathsf{util}^i(v_i, \mathbf{b}_{-i,1})].
\end{aligned}$$

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 3.4.4. *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 3.2.4, 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 each user i's bid is sampled independently from $\mathcal{D}_i$. There must exist a user j^* whose probability of being confirmed is at most k/K, and thus its expected utility is at most $M \cdot k/K$ where k is the block size and M is supreme values in the support of $\mathcal{D}_{j^*}$. Thus, $\mathop{\mathbb{E}}_{\mathbf{b}\sim\widehat{\mathcal{D}}}[\mathsf{util}^{j^*}(\mathbf{b})] = 0$ by taking K to be arbitrarily large.

By Lemma 3.4.3, it must be that $\mathop{\mathbb{E}}_{v_{j^*}\sim\mathcal{D}_{j^*}} \mathsf{util}^{j^*}(v_{j^*}) = 0$. Since users' honest utility must be non-negative, it must be that $\mathsf{util}^{j^*}(v) = 0$ for all $v \notin \mathsf{Bad}$ for some zero-measure set Bad. Since each $\mathcal{D}_i$ has the same support $[0, M]$, it must ve the for any user i, $\mathop{\mathbb{E}}_{v_i\sim\mathcal{D}_i} \mathsf{util}^i(v_i) = 0$. Therefore, by Lemma 3.4.3, it must be that for any user i, we have $\mathop{\mathbb{E}}_{\mathbf{b}\sim\widehat{\mathcal{D}}}[\mathsf{util}^i(\mathbf{b})] = 0$. □

We assume that the users' true value distributions have the same support because otherwise, there are artificially designed examples of TFMs that achieve positive social welfare while achieving UIC, MIC, and (ρ, c)-SCP for $c \geq 2$. For example, imagine that all users except for

some user i's true value are i.i.d. sampled from a uniform distribution between $[0, 1]$, while user i's true value is sampled from a uniform distribution between $[1, 2]$. Now we can run an MPC-assisted, burning posted price auction with random selection, where the reserved price r is set to 1.5. This mechanism satisfies UIC, MIC, and (ρ, c)-SCP for any c. To see this, note that all users and miners except for user i get at most 0-utility. No matter how other users change their bids or drop off, they cannot increase user i's utility. Therefore, this mechanism satisfies all three properties while achieving positive social welfare. However, such mechanisms are artificially designed and depend heavily on the structure of the underlying distributions, so we do not focus on such scenarios in this thesis.

Chapter 4

Characterization of Approximate Incentive Compatibility

As we have seen that for strict incentive compatibility, no mechanism can achieve positive miner revenue (even in the MPC-assisted model and infinite block size). Moreover, there is no dream mechanism that satisfies all desired properties with $c \geq 2$ even in the MPC-assisted model. Therefore, we consider another dimension of relaxation: approximate incentive compatibility. *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?*

In this section, we present the characterization of approximate incentive compatibility in both the MPC-assisted model and the plain model(Table 1.2).

4.1 Technical Overview

As before, we assume that honest users' true values are sampled *independently* from some bounded continuous distribution. We use $\mathcal{D}_i$ to denote user i's true value distribution and $\mathcal{D} := \mathcal{D}_1 \times \cdots \times \mathcal{D}_n$ to denote the joint distribution of users' true values. For any fixed user i, we also use $\mathcal{D}_{-i} := \mathcal{D}_1 \times \cdots \times \mathcal{D}_{i-1} \times \mathcal{D}_{i+1} \times \cdots \times \mathcal{D}_n$ to denote the distribution of other users' true values. Without loss of generality, we assume that for any i, distribution $\mathcal{D}_i$ has the same support $[0, M]$ for some M.

4.1.1 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 as well — 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 $\rho = 1$.

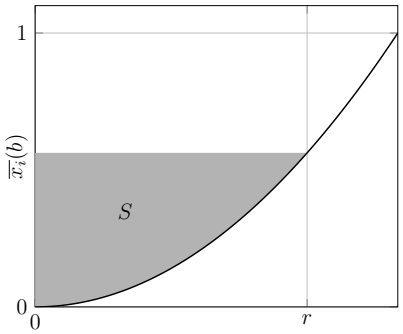
We now show an MPC-assisted mechanism simultaneously satisfies ϵ-UIC, ϵ-MIC and ϵ-SCP even for the Bayesian setting and even for $c = 1$ and an arbitrary choice $\rho \in (0, 1]$, then the miner can gain at most $O(n \cdot (\epsilon + \cdot\sqrt{m^* \cdot \epsilon}))$-miner revenue, where n is the number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

is a term that depends on the "scale" of the bid distribution $\mathcal{D}$. The proof is similar to that of Theorem 3.2.4. The crux is to characterize how miner revenue changes when we lower one user's bid to 0 (Lemma 4.2.3). We then apply this argument n times, and lower each user's bid one by one to 0 to get the desired bound.

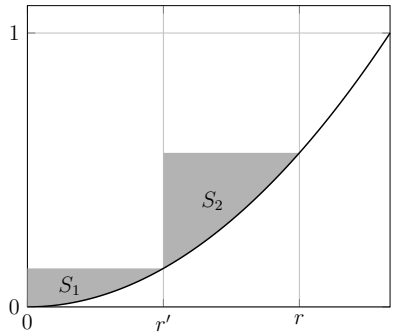
To understand how much the miner revenue changes when one user lowers its bid to 0, we start from a simplified case where a TFM $(\mathbf{x}, \mathbf{p}, \mu)$ is Bayesian *strict*-UIC and Bayesian ϵ-SCP for $c = 1$ and some $\rho \in (0, 1]$. By Myerson's Lemma Lemma 3.2.1, strict-UIC implies that, for any user i, the allocation rule $x_i(\cdot)$ must be non-decreasing, and the expected payment when bidding b must be

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

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 4.1a. 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 4.1b. 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 4.1b). 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.



(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 4.1: User's utility change when untruthful bidding.

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 $\overline{\mu_i}(r) - \overline{\mu_i}(0) \leq 2\sqrt{r\epsilon}$.

The above proof works for strict-UIC and ϵ-SCP. 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 anymore. Therefore, the key idea of

our proof is to give a generalization of Myerson's lemma to account for the ϵ slack in incentive compatibility. The formal proof is given in Section 4.2.

4.1.2 Approximate Incentive Compatible TFM in the Plain Model: Infinite Block Size

We now show that the above limit on miner revenue is asymptotically tight, i.e., we can indeed design a TFM, even in the plain model, that achieves asymptotically optimal miner revenue. The mechanism is called proportional auction 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.

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

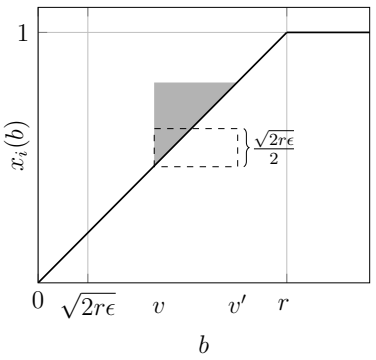
- Every bid $b \geq r$ is confirmed with probability 1 and every candidate bid $b < r$ is confirmed with probability b/r. Each confirmed bid b pays $p = \min\{\frac{b}{2}, \frac{r}{2}\}$.
- For each confirmed bid, the miner gets a pre-determined revenue $r' = \sqrt{\frac{2r\epsilon}{9c}}$ if $p \geq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

We now explain the intuition of why this mechanism satisfies approximate incentive compatibility, and defer the full proof to Section 4.3. 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 Lemma 3.2.1. 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.

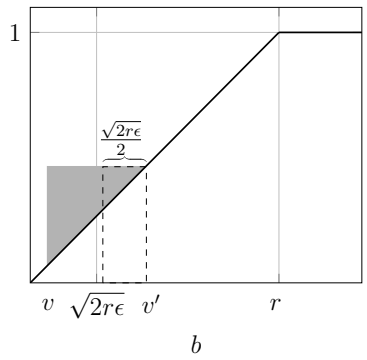
Proving that the mechanism satisfies approximate-SCP is more technical. Here we give an illustrative explanation to show that the joint utility of each user and the miner can increase by at most $\frac{5}{4}\epsilon$. Since underbidding does not increase the user's utility or the miner's revenue, we focus on overbidding. Note that overbidding does not increase the joint utility for a user whose true value is $v \geq r$. 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 4.2a. 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 4.2a. 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 4.2b. 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 4.2b. 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 4.2: Coalition's joint utility change when the miner colludes with one user.

4.1.3 Bound on Social Welfare in Plain Model: Finite Block Size

Unfortunately, although it is indeed possible to overcome the finite-block impossibility with approximate incentive compatibility (see Section 4.4.2), 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. An ϵ-incentive compatible TFM in the plain model must have a social welfare no more than $\widetilde{O}(k^3\epsilon)$, where k is the block size.

We will explain the proof blueprint in this overview and give the formal proof in Section 4.4.1. 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 largest possible true value, the expected miner revenue must be small (see Lemma 4.4.5). 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, then the miner revenue must be small too (see Lemma 4.4.6). Note that showing the first step itself relies on another inductive argument that inducts on the length of the bid vector.

4.1.4 Achieving Optimal and Scalable Social Welfare in the MPC-Assisted Model: Finite Block Size

We now show that if we consider approximate incentive compatibility in the MPC-assisted model, we can overcome the above 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.

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.

The diluted posted price auction satisfies UIC, MIC, and ϵ-SCP against any (ρ, c)-sized coalition. UIC and MIC are easy to see, so we explain why it satisfies approximate SCP. In particular, this is achieved by artificially diluting the probability that a user is confirmed when the number of eligible bids (i.e., offering at least r) is small. With the dilution, we guarantee that a coalition of c users cannot noticeably alter their own probability of getting confirmed, nor their friend's probability. This implies a strategic coalition has little influence over the expected utility of all users in the coalition. The formal proof is given in Section 4.5.

4.2 Bound on the Miner Revenue

We first prove a generalization of Myerson's *payment difference sandwich* for ϵ-UIC. Still, we use $(\mathbf{x}, \mathbf{p}, \mu)$ to represent a mechanism and define for the i-th user,

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

Lemma 4.2.1. *Given any (possibly randomized) MPC-assisted TFM that is Bayesian ϵ-UIC, it must be that for any user i, for any $y \leq z$,*

$$z \cdot [\overline{x_i}(z) - \overline{x_i}(y)] + \epsilon \geq \overline{p_i}(z) - \overline{p_i}(y) \geq y \cdot [\overline{x_i}(z) - \overline{x_i}(y)] - \epsilon. \tag{4.1}$$

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 4.2.2. *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 $y \leq z$,*

$$\overline{\mu_i}(z) - \overline{\mu_i}(y) \leq \frac{1}{\rho}(\epsilon_u + \epsilon_s + S(y, z)), \tag{4.2}$$

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

$$\begin{aligned}
\Delta &= y \cdot \overline{x_i}(y) - \overline{p_i}(y) - [y \cdot \overline{x_i}(z) - \overline{p_i}(z)] \\
&= -y \cdot [\overline{x_i}(z) - \overline{x_i}(y)] + (\overline{p_i}(z) - \overline{p_i}(y)) \\
&\leq -y \cdot [\overline{x_i}(z) - \overline{x_i}(y)] + z \cdot [\overline{x_i}(z) - \overline{x_i}(y)] + \epsilon_u \qquad \text{By Bayesian } \epsilon_u\text{-UIC and (4.1)} \\
&= (z - y) \cdot [\overline{x_i}(z) - \overline{x_i}(y)] + \epsilon_u = S(y, z) + \epsilon_u.
\end{aligned}$$

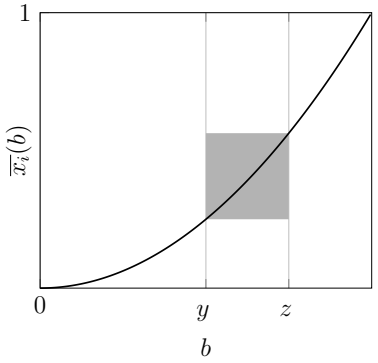
A graphical description of $S(y, z)$ is shown in Figure 4.3 — note that $S(y, z)$ can be *negative* since the allocation rule $\overline{x_i}(\cdot)$ may not be monotone under approximate UIC.

By Bayesian ϵ_s-SCP, it must be that $\rho\overline{\mu_i}(z) - \rho\overline{\mu_i}(y) \leq \Delta + \epsilon_s$; 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. □

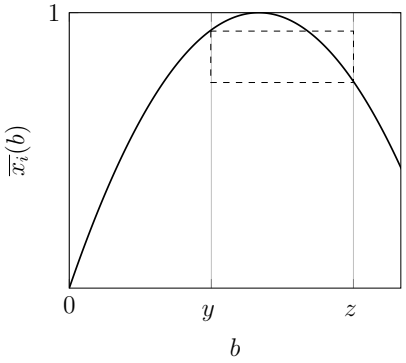
Lemma 4.2.3. *Fix any $\rho \in (0, 1]$. For any (possibly randomized) MPC-assisted TFM that is Bayesian ϵ_u-UIC and Bayesian ϵ_s-SCP against a $(\rho, 1)$-sized coalition, for any user i, for any value r, it must be that*

$$\overline{\mu_i}(r) - \overline{\mu_i}(0) \leq \begin{cases} \frac{2}{\rho}(\epsilon_s + \epsilon_u), & \text{if } r \leq \epsilon_s + \epsilon_u \\ \frac{2}{\rho}(\sqrt{r(\epsilon_s + \epsilon_u)}), & \text{if } r > \epsilon_s + \epsilon_u. \end{cases} \tag{4.3}$$

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



(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 4.3: User's utility change

Case 1: If $r \leq \epsilon'$. In this case, by Lemma 4.2.2, we have that

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

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

$$\begin{aligned}
\overline{\mu_i}(r) - \overline{\mu_i}(0) &= \sum_{l=0}^{L} [\overline{\mu_i}(r_{l+1}) - \overline{\mu_i}(r_l)] \\
&\leq \sum_{l=0}^{L} \frac{1}{\rho}[\epsilon' + S(r_l, r_{l+1})] && \text{By Lemma 4.2.2} \\
&= \frac{L\epsilon'}{\rho} + \frac{1}{\rho}\sum_{l=0}^{L} (r_{l+1} - r_l) \cdot [\overline{x_i}(r_{l+1}) - \overline{x_i}(r_l)] \\
&\leq \frac{L\epsilon'}{\rho} + \frac{1}{\rho}\sqrt{r\epsilon'}\sum_{l=0}^{L} [\overline{x_i}(r_{l+1}) - \overline{x_i}(r_l)] && \text{By the choice of } r_l \\
&\leq \frac{L\epsilon'}{\rho} + \frac{1}{\rho}\sqrt{r\epsilon'} && \text{By } \overline{x_i}(r) \leq 1
\end{aligned}$$

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

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

□

Now, we want to bound the miner revenue by lowering each user's bid to 0 one by one, and apply Lemma 4.2.3 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 (4.5) in the proof of Theorem 4.2.4 below. This ensures that in any step of the induction, any honest user's bid is sampled from $\mathcal{D}$.

Theorem 4.2.4 (Limit on miner revenue for approximate incentive compatibility)**.** *We use $\mathcal{D}^{(n)} = \mathcal{D}_1 \times \cdots \times \mathcal{D}_n$ to denote the joint distribution of the first n users' true values, where $\mathcal{D}_i$ is user i's true value distribution. 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 for any n,*

$$\mathop{\mathbb{E}}_{\mathbf{b}\sim\mathcal{D}^{(n)}}[\mu(\mathbf{b})] \le \frac{2n}{\rho}\left(\epsilon + C_{\mathcal{D}}\sqrt{\epsilon}\right), \tag{4.4}$$

where $\epsilon = \epsilon_s + \epsilon_u + \epsilon_m$, and $C_{\mathcal{D}} = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}_{X\sim\mathcal{D}_i}[\sqrt{X}]$ 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,

$$\mathop{\mathbb{E}}_{\mathbf{b}\sim\mathcal{D}^{(\ell)}}[\rho\mu(\mathbf{b}, 0)] \le \mathop{\mathbb{E}}_{\mathbf{b}\sim\mathcal{D}^{(\ell)}}[\rho\mu(\mathbf{b})] + \epsilon_m. \tag{4.5}$$

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,

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

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 4.2.3, it must be that

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

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

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

By (4.5), we have that $\mathop{\mathbb{E}}_{\mathbf{b}'\sim\mathcal{D}^{(n-1)}}[\mu(\mathbf{b}',0)] \le \mathop{\mathbb{E}}_{\mathbf{b}'\sim\mathcal{D}^{(n-1)}}[\mu(\mathbf{b}')] + \frac{\epsilon_m}{\rho}$. Therefore,

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

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

It is easy to see that the same miner revenue limit of Theorem 4.2.4 also holds in the plain model, as stated in the following corollary.

Corollary 4.2.5. *We use $\mathcal{D}$ to denote the joint distribution of the n users' true values. 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 for any n,*

$$\mathop{\mathbb{E}}_{\mathbf{b}\sim\mathcal{D}}[\mu(\mathbf{b})] \le 2n\left(\epsilon + C_{\mathcal{D}}\sqrt{\epsilon}\right), \tag{4.6}$$

where $\epsilon = \epsilon_s + \epsilon_u + \epsilon_m$, and $\mathcal{C}_{\mathcal{D}} = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}_{X\sim\mathcal{D}_i}[\sqrt{X}]$ is a term that depends on the "scale" of the distribution $\mathcal{D}$.

Proof. Follows directly from Theorem 4.2.4 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. □

4.3 Achieving Optimal Revenue: Proportional Auction

We now present the mechanism that achieves approximate incentive compatibility while achieving asymptotically optimal miner revenue, even in the plain model.

Theorem 4.3.1. *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. 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

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

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

Proportional Auction

Parameters: the slack ϵ, the reserved price r where $r \geq 2\epsilon$.
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 $b \geq r$, 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 $b \geq \sqrt{2r\epsilon}$[a], then miner is paid $\frac{\sqrt{2r\epsilon}}{2}$.

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

Figure 4.4: Proportional auction: achieving asymptotically optimal social welfare even in the plain model.

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

$$\frac{\Delta}{r}\sqrt{\frac{r\epsilon}{2}} - \frac{\Delta^2}{2r}. \tag{4.7}$$

The maximum of Eq.(4.7) 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

$$\frac{v_i}{r}\sqrt{\frac{r\epsilon}{2}} + \frac{\Delta}{r}\sqrt{\frac{r\epsilon}{2}} - \frac{\Delta^2}{2r}. \tag{4.8}$$

Because the maximum of Eq.(4.7) is $\frac{\epsilon}{4}$, the maximum of Eq.(4.8) 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. □

4.4 Characterizing Social Welfare in the Plain Model: Finite Block Size

4.4.1 Bound on Social Welfare

Theorem 4.4.1. *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 4.4.1 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 4.4.2. *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 4.4.1, 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 compared to the honest strategy. This contradicts ϵ-SCP. Note that since users' true values 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 4.4.1 is dedicated to proving Theorem 4.4.1.

Individual User's Influence on Miner Revenue is Bounded

Before proving Theorem 4.4.1, 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 4.4.3. *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 4.2.1 and Equation (4.9) 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 4.2.1 implies that for any $\mathbf{b}_{-i}$, for any $b \leq b'$,

$$b' \cdot [x_i(\mathbf{b}_{-i}, b') - x_i(\mathbf{b}_{-i}, b)] + \epsilon \geq p_i(\mathbf{b}_{-i}, b') - p_i(\mathbf{b}_{-i}, b) \geq b \cdot [x_i(\mathbf{b}_{-i}, b') - x_i(\mathbf{b}_{-i}, b)] - \epsilon. \quad (4.9)$$

Lemma 4.2.2 implies that for any $\mathbf{b}_{-i}$, for any $b \leq b'$,

$$\mu(\mathbf{b}_{-i}, b') - \mu(\mathbf{b}_{-i}, b) \leq 2\epsilon + (b' - b) \cdot [x_i(\mathbf{b}_{-i}, b') - x_i(\mathbf{b}_{-i}, b)]. \quad (4.10)$$

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 $v \leq 2\epsilon$, the lemma directly follows from (4.10). 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 $b \leq b'$, 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 $u_i(b) \leq 0$ for some b. However, we have that guarantee that at any point, $\int_0^b u_i(t)dt$ is non-negative.

By Equation (4.9), we know that for any $b \leq b'$, we have $p_i(b') - p_i(b) \leq b'[x_i(b') - x_i(b)] + \epsilon$, i.e.,

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

which is equivalent to

$$\xi(b, b') := (b' - b) \cdot x_i(b) - \int_b^{b'} u_i(t)dt \leq \epsilon. \tag{4.11}$$

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[1] for u_i, between b and b'. For example, the blue area in Figure 4.5a 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 (4.10), the miner revenue

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

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

$$\begin{aligned}
&\mu(v) - \mu(0) = \mu(v) - \mu(b_L) + \sum_{l=1}^{L} \mu(b_l) - \mu(b_{l-1}) \\
\leq &2\epsilon + (v - b_L) \cdot [x_i(v) - x_i(b_L)] + \sum_{l=1}^{L} (S(b_{l-1}, b_l) + 2\epsilon) && \text{By (4.10)} \\
\leq &4\epsilon + 2L\epsilon + \sum_{l=1}^{L} S(b_{l-1}, b_l). && \text{By } v - b_L \leq 2\epsilon
\end{aligned}$$

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

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

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

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

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

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 4.5b. 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 following calculation says that this difference is upper-bounded by Formally,

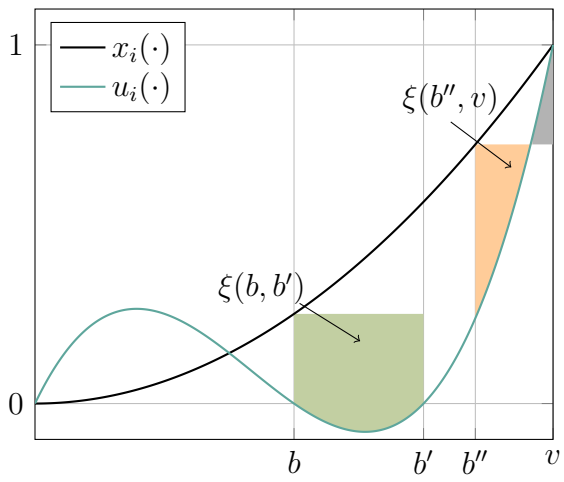
$$\begin{aligned}
&\sum_{l=1}^{L} S(b_{l-1}, b_l) - \int_0^v u_i(t)dt = \sum_{l=1}^{L} (b_{l+1} - b_l) x_i(b_l) - \int_0^v u_i(t)dt \\
\leq &\sum_{l=1}^{L} \left\{ (b_{l+1} - b_l) \cdot x_i(b_l) - \int_{b_l}^{b_{l+1}} u_i(t)dt \right\} && \text{By } \int_0^{b_1} u_i(t) \geq 0 \\
= &\sum_{l=1}^{L} \xi(b_l, b_l + 1) \leq \sum_{l=1}^{L} \epsilon = L\epsilon. && \text{By (4.11)}
\end{aligned}$$

[Ke: Add a label in figure]

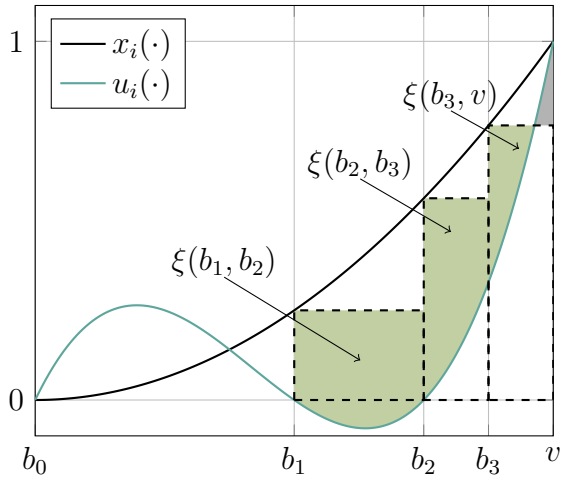
Putting it together, the change in miner revenue $\mu(v) - \mu(0)$ is upper bounded by

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

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



(a) The green 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 4.5: Graphical explanation of the proof to Lemma 4.4.3

Because the miner can inject a bid 0 for free, Lemma 4.4.3 implies the following corollary, which says that if we remove a bid, the miner revenue should not be affected by too much.

Corollary 4.4.4. *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,*

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

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

$$\mu(\mathbf{b}_{-i}, 0) - \mu(\mathbf{b}_{-i}) \le \epsilon. \tag{4.12}$$

The corollary is now directly implied by Equation (4.12) and Lemma 4.4.3. □

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 4.4.5). Then, we generalize to the case when everyone's bids need not be the same (see Lemma 4.4.6).

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 $\mu(\mathbf{b})$ denotes 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 4.4.5. *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

$$\mu(\mathbf{m}_t) + \mathsf{util}^i(\mathbf{m}_t) \leq \mu(\mathbf{m}_K) + 2\epsilon. \tag{4.13}$$

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 (4.13) must hold for any $t \leq K$ and any user $i \in [t]$.

According to Equation (4.13), we have $\mu(\mathbf{m}_{t^*}) + \mathsf{util}^i(\mathbf{m}_{t^*}) \leq \mu(\mathbf{m}_K) + 2\epsilon$ for any user i. As we have shown, it must be $\mu(\mathbf{m}_{t^*}) \geq \mu(\mathbf{m}_K) - \epsilon$. Combining these two inequalities, we have

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

Since the utility of user i is bounded, by applying Corollary 4.4.4, it must be

$$\mu(\mathbf{m}_{t^*}) - \mu(\mathbf{m}_{t^*-1}) \leq \mathsf{util}^i(\mathbf{m}_{t^*}) + 3\epsilon \log \frac{M}{\epsilon} + 5\epsilon \leq 8\epsilon + 3\epsilon \log \frac{M}{\epsilon}. \tag{4.14}$$

Consequently, we have

$$\begin{aligned}
&\mathsf{util}^i(\mathbf{m}_{t^*-1}) \leq \mu(\mathbf{m}_K) + 2\epsilon - \mu(\mathbf{m}_{t^*-1}) && \text{By (4.13)} \\
\leq &\mu(\mathbf{m}_K) + 2\epsilon - \mu(\mathbf{m}_{t^*}) + 8\epsilon + 3\epsilon \log \frac{M}{\epsilon} && \text{By (4.14)} \\
\leq &\mu(\mathbf{m}_K) + 2\epsilon - \mu(\mathbf{m}_K) + \epsilon + 8\epsilon + 3\epsilon \log \frac{M}{\epsilon} && \text{By } \mu(\mathbf{m}_{t^*}) \geq \mu(\mathbf{m}_K) - \epsilon \\
= &11\epsilon + 3\epsilon \log \frac{M}{\epsilon}.
\end{aligned}$$

Then, we can apply Corollary 4.4.4 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

$$\mu(\mathbf{m}_{t^*-r}) - \mu(\mathbf{m}_{t^*-r-1}) \le (8r+8)\epsilon + (3r+3) \cdot \epsilon \log \frac{M}{\epsilon}. \tag{4.15}$$

Since $M \ge 2\epsilon$, we have $\epsilon \log \frac{M}{\epsilon} \ge \epsilon$. By Eq.(4.15), 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 4.4.6. *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 $M \ge 2\epsilon$. 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 $t \ge 1$, 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 $t + |\mathbf{L}| \le k$, the miner revenue $\widetilde{\mu}(\mathbf{m}_t, \mathbf{L}) \le \tau + 2|\mathbf{L}|\epsilon$, 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 4.4.5.

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 $\widetilde{\mu}(\mathbf{b}) \leq 12k^2\epsilon \log \frac{M}{\epsilon} + 2k\epsilon$. □

Completing the Proof of Theorem 4.4.1

We now complete the proof of Theorem 4.4.1. 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 4.4.7), which then leads to our conclusion that the total social welfare must be small.

Lemma 4.4.7. *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 $i \leq |\mathbf{b}|$, 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 4.4.1. 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 4.4.6. 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 $M \geq 2\epsilon$ and we define $\widehat{util}^i(\mathbf{b})$ to be the utility of user i conditioned on being confirmed when the world consists of the bid vector $\mathbf{b}$. By Lemma 4.4.6, the miner can have at most $\left(12k^2\epsilon \log \frac{M}{\epsilon} + 2k\epsilon\right)$-miner revenue. By Lemma 4.4.7, for any i, $\widehat{util}^i(\mathbf{b}) \leq 12k^2\epsilon \log \frac{M}{\epsilon} + (2k+1)\epsilon$. Let γ_i be the probability that user i is included in the block given the bid vector $\mathbf{b}$. Observe that $\sum_i \gamma_i \leq k$ for any $\mathbf{b}$. Therefore, the expected total utility of all users is upper-bounded by

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

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.

4.4.2 Achieving Approximate Incentive Compatibility in Plain Model: Finite Block Size

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 4.4.1, 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 $t \cdot \epsilon$.

Figure 4.6: Staircase mechanism: achieving approximate incentive compatibility in the plain model for finite block size.

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 4.7, which explains why the mechanism is called "staircase".

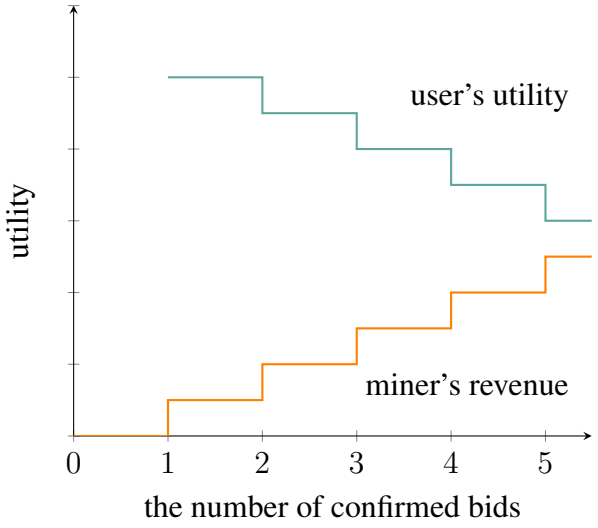


Figure 4.7: The miner's revenue and any user's utility as the functions of the number of the confirmed bids in the block.

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 4.4.8. *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 $\mathbf{b}$.** 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 $\mathbf{b}$.** 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 $\mathbf{d}$.** 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.

□

4.5 Diluted Posted Price Auction in the MPC-Assisted Model

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 $\frac{\epsilon}{2c}$.

Figure 4.8: Diluted posted price auction: achieving approximate incentive compatible and optimal social welfare in the plain model for finite block size.

Theorem 4.5.1. *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 positive 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

$$(v-r)\frac{k}{\max\{T,\ell'\}} - (v-r)\frac{k}{\max\{T,\ell\}}. \tag{4.16}$$

Note that for $\ell' \geq \ell$, then (4.16) ≤ 0. Therefore, we only need to consider the case where $\ell' < \ell$. If $\ell \leq T$, then (4.16) is 0. If $\ell > T$, then (4.16) is upper bounded by

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

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

$$(4.16) \leq M \cdot \frac{ck}{T(T-c)} \leq \frac{2Mck}{T^2} \leq \frac{\epsilon}{2c}.$$

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.

□

Chapter 5

Reasonable-World Assumption

From previous chapters, we have seen that in both the plain and the MPC-assisted models, the 0 miner-revenue limitation holds regardless of whether the block size is finite or infinite, and even when the miner colludes with at most $c = 1$ user. Although with approximate incentive compatibility, we can indeed circumvent the 0 miner-revenue barrier (Theorem 3.2.4), unfortunately, it cannot scale proportionally as the magnitude of the bids increases.

To get rid of the above limitations, we consider a mildly stronger reasonable-world assumption. We will see that even with this extra assumption, it does not overcome the 0-miner revenue limitation if we insist on ex post incentive compatibility. Therefore, in the rest of this section, we will mostly focus on the characterization of the miner revenue for *Bayesian incentive compatibility* in the MPC-assisted model. Still, we make the assumption that each honest user i's bid is sampled independently from some a-priori known continuous distribution $\mathcal{D}_i$ over the support $[0, M]$ for some M. In addition, throughout this chapter, we assume that for any i, $\mathcal{D}_i$ has the same median m such that $\Pr_{x_i \sim \mathcal{D}_i}[x \geq m] = \frac{1}{2}$.

5.1 (h, ρ, c, d)-Environment

Environment. In this chapter, we consider an (h, ρ, c, d)-environment, where h is the promised lower bound on the number of honest users, $\rho \in [0, 1]$ is the fraction of strategic miners, c is the maximum number of strategic users that collude with miners, and d is the maximum number of bids contributed by the strategic coalition, i.e. fake bids.

We can replace a subset of these variables with a wildcard $*$ if the mechanism achieves incentive compatibility no matter what the variable turns out to be. For example, a TFM that achieves incentive compatibility in an $(*, \rho, c, *)$-environment if it works when the maximum fraction of strategic miners is ρ, and the maximum number of colluding users is at most c, and regardless of of how many honest users there are and how many bids are contributed by strategic individuals or the coalition. In this case, we also say that the mechanism is universal in the parameters h and d. Using this notation, the previous mechanisms in this thesis are universal in the parameters h and d. Similarly, the limitation on miner revenue they prove can also be interpreted as a limitation of mechanisms that are universal in h and d.

Incentive Compatibility. The incentive compatibility notions in an (h, ρ, c, d)-environment are similar to definitions in Section 2.4, with an extra condition that strategic players cannot increase their own utility compared to honest strategy as long as the conditions required by the environment are respected, i.e., there are at least h number of honest users, the strategic players control at most d number of colluding bids, and the strategic coalition controls at most ρ fraction of the miners and at most c number of users.

5.2 Technical Roadmap

5.2.1 Characterization under Infinite Block Size

For the infinite block setting, we can achieve $\Theta(h)$ miner revenue in (h, ρ, c, d)-environments. To aid understanding, we first present a simple parity-based mechanism that works for $h = 1$, and then we present our main results.

Glimpse of hope. First, consider the special case where we are promised that there is at least $h = 1$ honest user. In this case, the following simple parity-based mechanism satisfies ex-post UIC, Bayesian MIC, and Bayesian SCP in $(1, *, *, *)$-environments.

MPC-assisted, parity-based mechanism

// Let m be the median of honest users' true value distribution.

- All bids that are at least m get confirmed and pay m.
- If the number of confirmed bids is odd, then the total miner revenue is m; else the total miner revenue is 0.

In the above mechanism, as long as there is at least one honest bid, the expected miner revenue is always $m/2$ no matter how the coalition behaves. This is because the strategic coalition cannot predict whether the honest bid is bidding at least the median or not. With this key observation, it is not hard to see that the mechanism satisfies Bayesian MIC and Bayesian SCP (for an arbitrary c). Further, ex post UIC follows directly since the mechanism is a simple posted-price auction from a user's perspective.

Observe also the following subtlety: when the number of confirmed bids is odd, it implies that there is at least one confirmed bid. Therefore, the mechanism guarantees that the miner revenue does not exceed the total payment.

Approximate Incentive Compatibility: Threshold-based mechanism.

The parity-based mechanism overcomes the 0 miner-revenue limitation by assuming the existence of at least $h = 1$ honest user. However, the drawback is obvious: the total miner revenue is severely restricted and does not increase w.r.t. the number of bids. A natural question is whether we can achieve $O(h)$ expected miner revenue for general h.

We give an affirmative answer. We first present a simple, practical mechanism called the threshold-based mechanism that achieves almost-strict incentive compatibility except for a tiny

slack ϵ that is exponentially small in h. Then, for theoretical interest, we present another mechanism that achieves strict incentive compatibility but requires an extra assumption on the number of bids contributed by strategic players.

MPC-assisted, threshold-based mechanism

// Let m be the median of honest users' true value distribution.

- All bids that are at least m get confirmed and pay m.
- If the number of confirmed bids is at least $h/4$, then the miner revenue is $m \cdot h/4$; else the total miner revenue is 0.

Due to the standard Chernoff bound, except with $e^{-\Omega(h)}$ probability, the number of confirmed bids among the h (or more) honest bids is at least $h/4$. Therefore, the above mechanism achieves at least $m \cdot h/4 \cdot (1 - e^{-\Omega(h)})$ expected miner revenue. If the number of confirmed honest bids is $h/4$ or higher, then the coalition cannot increase the miner revenue no matter how it behaves. Only when the number of confirmed honest bids is less than $h/4$, is it possible for the coalition to influence the miner revenue by at most $m \cdot h/4$. Therefore, it is not hard to see that the mechanism satisfies ϵ-Bayesian MIC and ϵ-Bayesian SCP in $(h, *, *, *)$-environments, for $\epsilon = m \cdot h \cdot e^{-\Omega(h)}/4$. The formal proof is given in Section 5.3.2

Remark 5.2.1 (On the robustness of parameter estimation)**.** The threshold-based mechanism requires the mechanism to estimate h and the median m of the distribution $\mathcal{D}$. Just like how Ethereum EIP-1559 estimates its base price, we can estimate h and the median from recent history. In particular, from the congestion level in recent blocks, we can estimate a lower bound on the total number of bids. Now, assuming that at least half of them are honest (recall that our mechanism incentivizes honesty), we can get an estimate of h correspondingly. Similarly, we can estimate the median of the bid distribution from past history.

An advantage of the threshold-based mechanism is that it is quite tolerant of errors in the estimation. For example, if the estimated m is actually the 40-percentile of $\mathcal{D}$, and the actual number of honest users is only $0.7h$ where h is the estimate used by the mechanism, the expected number of users bidding at least m is at least $0.4 \cdot 0.7h = 0.28h$. In this case, we can still guarantee that except with exponentially small in h probability, at least $h/4$ users will bid at least m. Thus, the resulting mechanism would still be almost strictly incentive compatible except for a slack ϵ that is exponentially small in h.

Strict Incentive Compatibility: LP-based mechanism.

Although the ϵ slack in the threshold-based mechanism is exponentially small which is not a problem in practice, it is still theoretically interesting to ask whether we can get $\Theta(h)$ total miner revenue but with *strict* incentive compatibility. To achieve this, our idea is to devise a mechanism that is "close in distance" to the aforementioned threshold-based mechanism, but correcting the "error" such that we can achieve strict incentive compatibility. Observe that the earlier threshold-based mechanism only needs an a-priori known lower bound on h, and it is universal in the parameters c and d. To achieve strict incentive compatibility, we additionally assume that the the number of bids contributed by the strategic coalition is upper bounded by some a-priori known parameter d.

Now, consider the following mechanism that relies on linear programming to correct the error in the earlier threshold-based mechanism.

MPC-assisted, LP-based mechanism

Parameters: Let m be the median of honest users' true value distributions.

Mechanism:

- *Allocation and payment rule*: All bids that are at least m get confirmed and pay m.
- Miner revenue rule: Let n be the length of the bid vector, let $\mathbf{y} := (y_0, y_1, \ldots, y_n)$ be any feasible solution to the following linear program:

$$\forall i \in [n] : \ 0 \leq y_i \leq i \cdot m \tag{5.1}$$

$$\forall 0 \leq j \leq d : \ \sum_{i=0}^{n-d} q_i \cdot y_{i+j} = \frac{m \cdot h}{4} \tag{5.2}$$

  where q_i is the probability of observing i heads if we flip $n - d$ independent fair coins.

  The total miner revenue is y_s where s is the number of bids confirmed.

Figure 5.1: LP-based mechanism in the MPC-assisted model

In the above, Equation (5.1) expresses a *budget feasibility* requirement, i.e., the total miner revenue cannot exceed the total user payment. Equation (5.2) expresses a *fixed-revenue requirement* stipulating that the miner revenue must be exactly $m \cdot h/4$ no matter how the strategic individual or coalition behaves (as long as it controls at most d bids). More specifically, Equation (5.2) contains one requirement for each $j \in [0, d]$: conditioned on the fact that among the (at most) d bids controlled by the strategic individual or coalition, exactly j of them are confirmed, the expected miner revenue must be exactly $m \cdot h/4$ where h is an a-priori known lower bound on the number of honest users.

Remark 5.2.2. We know that the actual number of honest users that show up is at least $max(n - d, h)$. So if $n - d > h$, it means that more honest users showed up than the anticipated number h. Observe that on the left-hand side of Equation (5.2), we are tossing coins for $n - d$ honest users' bids. However, it is important that the right-hand-side of Equation (5.2) use the a-priori known h rather than the observed $n - d$; otherwise, injecting extra (but up to $d - c$) fake 0-bids can increase the expected miner revenue, which violates MIC and SCP.

If the LP in the above mechanism indeed has a feasible solution, then we can prove that the resulting mechanism satisfies ex post UIC, Bayesian MIC, and Bayesian SCP in $(h, *, *, d)$-environments. The formal proofs are presented in Section 5.3.3.

The key technical challenge is to answer the question of why the LP has a feasible solution. Intuitively, the earlier threshold-based mechanism gives an "approximate" solution $\widehat{\mathbf{y}} := (\widehat{y}_0, \ldots, \widehat{y}_n)$ to the LP, where $\widehat{y}_i = 0$ for $i \leq (n - d)/4$ and $\widehat{y}_i = \frac{m \cdot h}{4}$ otherwise. With the approximate solution $\widehat{\mathbf{y}}$, the equality constraints in Equation (5.2) may be satisfied with some small error. We want to show that we can adjust the $\widehat{\mathbf{y}} := (y_0, y_1, \ldots, y_n)$ vector slightly such that we can correct the error, and yet without violating the budget feasibility constraints (Equation (5.1)).

To achieve this, we will take a constructive approach. We first guess that a feasible solution is of the form $\mathbf{y} = \widehat{\mathbf{y}} + \mathbf{e}$ where $\mathbf{e}$ is a correction vector that is zero everywhere except in the coordinates $\tau, \tau+1, \ldots, \tau+d$ for some appropriate choice of τ that is close to $(n-d)/2$. Henceforth, let $\boldsymbol{\delta} := \mathbf{e}[\tau : \tau+d] / \left(\frac{m \cdot h}{4}\right)$ be the non-zero coordinates of the correction vector $\mathbf{e}$ scaled by $\frac{m \cdot h}{4}$.

By Equation (5.2), we know that the correction vector $\boldsymbol{\delta}$ must satisfy the following system of linear equations where $t := \frac{n-d}{4}$:

$$\begin{pmatrix} \binom{n-d}{\tau} & \binom{n-d}{\tau+1} & \cdots & \binom{n-d}{\tau+d} \\ \binom{n-d}{\tau-1} & \binom{n-d}{\tau} & \cdots & \binom{n-d}{\tau+d-1} \\ \vdots & \vdots & \ddots & \vdots \\ \binom{n-d}{\tau-d} & \binom{n-d}{\tau-d+1} & \cdots & \binom{n-d}{\tau} \end{pmatrix} \cdot \boldsymbol{\delta} = \begin{pmatrix} \sum_{i=0}^{t} \binom{n-d}{i} \\ \sum_{i=0}^{t-1} \binom{n-d}{i} \\ \vdots \\ \sum_{i=0}^{t-d} \binom{n-d}{i} \end{pmatrix}. \tag{5.3}$$

In Lemma 5.3.8, we prove that as long as $d \leq \frac{1}{8}\sqrt{\frac{h}{2 \log h}}$, and that τ is an appropriate choice close to $n/2$, then the solution $\boldsymbol{\delta}$ to the linear system in Equation (5.3) has a small infinity norm — specifically, $\|\boldsymbol{\delta}\|_\infty \leq 1$ — such that the resulting $\mathbf{y}$ vector will respect the budget feasibility constraints, i.e., Equation (5.1). The actual proof of this bound is somewhat involved and thus deferred to Section 5.3.3. In particular, a key step is to bound the *smallest singular value* of the matrix in Equation (5.3) (henceforth denoted A) appropriately — to achieve this, we first bound A's determinant, and then use an inequality proven by [YG97] which relates the smallest singular value and the determinant.

Limit on miner revenue.

In Section 5.5.1, we prove that $\Theta(h)$ revenue is optimal in (h, ρ, c, d)-environments for strict incentive compatibility; and further, we generalize the bound to approximate incentive compatibility as well (see Theorem 5.5.1). The proof is a generalization of the techniques in Theorem 4.2.4. Specifically, Theorem 4.2.4 proved that any mechanism that satisfies Bayesian UIC, MIC, and SCP in $(*, \rho, 1, *)$-environments must suffer from 0 miner revenue. Recall that in the proof, we gradually remove users' bids one by one and bound the miner revenue change during this process. In (h, ρ, c, d)-environment, because the mechanism is promised a lower bound h on the number of honest users, we can repeat this argument till there are h bids left.

5.2.2 Characterization under Finite Block Setting

Feasibility for $c = 1$: LP-Based Mechanism with Random Selection

The LP-based mechanism confirms any bid that offers to pay at least m. Thus, total number of confirmed bids may be unbounded. Therefore, when the block size k is finite, we cannot directly run the LP-based mechanism. We suggest the following modification to the LP-based mechanism such that it works for the finite-block setting:

MPC-assisted, LP-based mechanism with random selection

// Let k be the block size, let m be the median of honest users' true value distributions

- All bids offering at least m are candidates. If there are more than k candidates, randomly select k of them to confirm; else confirm all candidates. Every confirmed bid pays m.
- Let n be the length of the bid vector, let $\mathbf{y} = (y_0, y_1, \ldots, y_n)$ be any feasible solution to the following linear program:

$$\forall i \in [n] : \ 0 \leq y_i \leq min(i, k) \cdot m \tag{5.4}$$

$$\forall 0 \leq j \leq d : \ \sum_{i=0}^{n-d} q_i \cdot y_{i+j} = \frac{m \cdot min(h, k)}{4} \tag{5.5}$$

  where q_i is the probability of observing i heads if we flip $n - d$ independent fair coins.
- The total miner revenue is y_s where s is the number of *candidates*.

In comparison with the earlier LP-based mechanism, we modify the budget feasibility constraints (Equation (5.4)) to make sure that the total miner revenue is constrained by the actual number of confirmed bids which is now $min(i, k)$ if the number of candidates is i. Further, we modify the expected miner revenue (Equation (5.5)) to be $\frac{m \cdot min(h,k)}{4}$ which takes into account the block size k. In Section 5.4.1, we prove that as long as $c \leq d \leq \frac{1}{8}\sqrt{\frac{h}{2 \log h}}$, the above LP indeed has a feasible solution and the resulting mechanism satisfies ex post UIC, Bayesian MIC, and Bayesian SCP in $(h, *, 1, d)$-environments.

Impossibility for $c \geq 2$: 0-User Social Welfare

Unfortunately, the above mechanism fails for $c \geq 2$. In this case, two users i and j may be in the same coalition. User i can now help user j simply by dropping out and not posting a bid, thus effectively increasing User j's chance of getting confirmed. In the event that user i's true value is very small and user j's true value is sufficiently large, this strategic action can increase the coalition's joint utility.

Interestingly, it turns out that this is no accident. Even with the extra reasonable-world assumption, we prove that for any $h \geq 1$, $\rho \in (0, 1)$, and $d \geq c \geq 2$, no "interesting" mechanism can simultaneously achieve Bayesian UIC, MIC and SCP in (h, ρ, c, d)-environments — any such mechanism must suffer from 0 total social welfare for the users if the actual number of bids received is greater than h (see Theorem 5.4.2 for the formal statement).

The proof blueprint is similar to that of Theorem 3.4.4, except with the difference that we are now guaranteed the existence of at least h number of honest users. Below, consider any TFM that satisfies Bayesian UIC, MIC and Bayesian SCP in (h, ρ, c, d)-environments where $d \geq c \geq 2$.

1. First, in Lemma 5.4.3, using techniques inspired by Goldberg and Hartline [GH05] we prove the following: provided that there are at least h honest users (not including i and j) whose bids are sampled at random from $\mathcal{D}$, then a strategic user i changing its bid should not affect the utility of another user j, if user j's bid is also sampled at random from $\mathcal{D}$.
2. Next, in Lemma 5.4.3, we prove a strategic user i dropping out should not affect another user

j's utility, assuming that at least h bids (excluding user i) sampled at random from $\mathcal{D}$.

3. Next, in Corollary 5.4.4, we show that in a world of at least h random bids (excluding user i) sampled from $\mathcal{D}$, user i's expected utility when its bid is sampled randomly from $\mathcal{D}$ depends only on i's identity, and does not depend on the identities of the other random bids. Therefore, henceforth we can use U_i to denote this expected utility.
4. Next, in Lemma 5.4.5, we show that for any two identities i, j, it must be that $U_i = U_j$, otherwise, it violates the assumption that the mechanism is weakly symmetric (see definition of weak symmetry below).
5. Next, we can show that $U_i = 0$: imagine a world with K bids sampled independently from $\mathcal{D}$ whose support is bounded. There must exist some user whose confirmation probability is upper bounded by k/K. This user's expected utility must be arbitrarily small when K is arbitrarily large. With a little more work, we can show that if the world consists of more than h bids sampled independently at random from $\mathcal{D}$, it must be that every user's expected utility is 0.

As with Theorem 3.4.4, one technicality that arises in the full proof (see Section 5.4.2) is the usage of the weak symmetry assumption. Our actual proof of the zero user social welfare result relies only on the *weak symmetry* assumption (see Section 2.2).

Approximate Incentive Compatibility: Diluted Threshold-Based Mechanism

Because of the above limitation on user social welfare, we relax the notion into approximate incentive compatibility, and ask if we can achieve optimal miner revenue in the finite block setting. Consider the following TFM.

MPC-assisted, diluted threshold-based Mechanism

// Let k be the block size, let m be the median of honest users' true value distributions.

- Let $R := max\left(2c\sqrt{\frac{kM}{\epsilon}}, k\right)$. All bids offering at least m are candidates. If the number of candidates $s \leq R$, randomly select $\frac{k}{R} \cdot s$ candidates to confirm; else, randomly select k candidates to confirm. Every confirmed bid pays m.
- If $s \geq \frac{h}{4}$, then the total miner revenue is $min(\frac{h}{4} \cdot \frac{k}{R}, k) \cdot m$. Otherwise, the miners get nothing.

Intuitively, here are modifying the earlier threshold-based mechanism to 1) make it compatible with finite block size, and 2) make sure that up to c users dropping out can only minimally increase their friend's probability of getting confirmed. Similar to Figure 4.8, dilution guarantees that a coalition of c users cannot noticeably alter their own probability of getting confirmed, nor their friend's probability. This implies a strategic coalition has little influence over the expected utility of all users in the coalition. Moreover, we guarantee that a strategic coalition has very little influence on the miner revenue as well: similar to the threshold-based mechanism, except with $\exp(-\Omega(h))$ probability, the miner revenue is an a-priori fixed amount, that is, $min(\frac{h}{4} \cdot \frac{k}{R}, k) \cdot m$. Summarizing the above, we can show that the mechanism satisfies ex post UIC, Bayesian ϵ-MIC, and Bayesian ϵ-SCP in $(h, *, c, *)$-environments, as long as $\epsilon \geq m \cdot \frac{h}{2} \cdot e^{-\frac{h}{16}}$.

Finally, for sufficiently large $h \geq max(4k, 8c\sqrt{\frac{kM}{\epsilon}})$, the mechanism achieves $k \cdot m$ total miner revenue and $k \cdot C_{\mathcal{D}}$ user social welfare where $C_{\mathcal{D}}$ is some constant depending on the distribution. For example, suppose we are willing to tolerate $\epsilon = 0.01T$, then we just need $h \geq max(4k, 80c \cdot \sqrt{k})$ to achieve asymptotic optimality in miner revenue and social welfare. The full proof is deferred to Section 5.4.3.

5.3 Feasibility under Infinite Block Size

5.3.1 MPC-Assisted, Parity-Based Mechanism

MPC-assisted, parity-based mechanism

Parameters: Let m be the median of honest users' true value distributions.

Mechanism:

- *Allocation and payment rule*: All bids that are at least m get confirmed and pay m.
- *Miner revenue rule*: If the number of confirmed bids is odd, then the total miner revenue is m; else the total miner revenue is 0.

Figure 5.2: Parity-based mechanism in the MPC-assisted model.

Theorem 5.3.1. *Fix any $h \geq 1$. The MPC-assisted, parity-based mechanism satisfies ex post UIC, Bayesian MIC, and Bayesian SCP in an $(h, *, *, *)$-environment.*

Proof. We prove the three properties separately.

UIC. Since the confirmation and the payment of each bid are independent of other bids, injecting fake bids or dropping out does not help to increase each user's utility. For a fixed user i, suppose its true value is v. When it bids b, its expected utility is $v - m$ if $b \geq m$. By direct calculation, the expected utility is maximized when $b = v$ no matter what other bids are. Thus, the MPC-assisted, parity-based mechanism satisfies UIC.

MIC. MIC follows directly from the fact that as long as there is one honest bid sampled from $\mathcal{D}$, the expectation of the parity is $\frac{1}{2}$, no matter what other bids are. Therefore, the expected total miner revenue is always $\frac{1}{2}m$.

SCP. By the same reasoning as in MIC, no matter how the coalition deviates, they cannot increase the total miner revenue. Henceforth, we only need to argue that no matter how the coalition deviates, it cannot increase the joint users' utility. This follows directly from UIC and the fact that each bid is confirmed independently from other bids. □

MPC-assisted, threshold-based mechanism

Parameters: lower bound h on the number of honest users, the distribution median m.

Mechanism:

- *Allocation rule.* Given a bid vector $\mathbf{b} = (b_1, \ldots, b_\ell)$, for each bid b_i, confirm b_i if $b_i \geq m$.
- *Payment rule.* Each confirmed bid pays m.
- *Miner revenue rule.* Let s be the number of confirmed bids. If $s \geq \frac{h}{4}$, miner gets $\frac{h}{4} \cdot m$. Otherwise, the miner gets nothing.

Figure 5.3: Threshold-based mechanism in the MPC-assisted model

5.3.2 MPC-Assisted, Threshold-Based Mechanism

Recall that we assume that each honest user i's true values are drawn independently from $\mathcal{D}_i$, and that m is the median of the true value distributions such that $\Pr_{x \sim \mathcal{D}_i}[x \geq m] = \frac{1}{2}$.

Theorem 5.3.2. *Fix any $h \geq 1$. The MPC-assisted, threshold-based mechanism satisfies ex post UIC, Bayesian ϵ-MIC and Bayesian ϵ-SCP in an $(h, *, *, *)$-environment, where $\epsilon = \frac{h}{4} \cdot m \cdot e^{-\frac{h}{16}}$.*

Proof. First, UIC follows from the same reasoning Theorem 5.3.1. We will focus on MIC and SCP in the rest of the proof.

ϵ-MIC. Recall that the only strategy that a strategic miner can apply is injecting some fake bids. Because injecting fake bids smaller than m does not influence the colluding miners' utility, we only consider injecting bids at least m. Let X denote the random variable representing the number of honest bids at least m. The only situation where the colluding miners can increase their expected gain by injecting fake bids is when $X < \frac{h}{4}$. By the following Chernoff Bound,

Lemma 5.3.3 (Chernoff bound, Corollary A.1.14 [AS16])**.** *Let $X_1, \ldots, X_n$ be independent Bernoulli random variables. Let $mu = \mathbb{E}\left[\sum_{i=1}^{n} X_i\right]$. Then, for any $\epsilon \in (0, 1)$, it holds that*

$$\Pr\left[\sum_{i=1}^{n} X_i \leq (1-\epsilon)mu\right] \leq e^{-\epsilon^2 mu/2}.$$

We have

$$\Pr\left[X < \frac{h}{4}\right] \leq e^{-\frac{h}{16}}.$$

Therefore, the colluding miners can gain at most $\frac{h}{4} \cdot m \cdot e^{-\frac{h}{16}}$ more expected revenue by injecting fake bids.

ϵ-SCP. Since the confirmation and the payment of each bid are independent of other bids, and the mechanism is strict UIC, the coalition cannot increase colluding users' utilities. Therefore, by deviating from the mechanism, the coalition can only try to increase the expected total miner revenue. By a similar argument as MIC, the coalition can only increase the expected total miner revenue when $X < \frac{h}{4}$, which happens with a probability no more than $e^{-\frac{h}{16}}$. It follows that no

matter how the coalition deviates, the expected miner's revenue can increase by at most $\rho \cdot \frac{h}{4} \cdot m \cdot e^{-\frac{h}{16}}$. □

5.3.3 MPC-Assisted, LP-Based Mechanism

We first introduce some linear algebra tools needed for analyzing the LP-based mechanism.

Preliminaries: Linear Algebra Tools

Throughout this section, all our indexing for vectors and matrices starts from 0. Given a vector $\mathbf{b} = (b_0, b_1, \ldots, b_n)$ and two integers i, j such that $i \leq j$, we define $\mathbf{b}[i : j]$ to be the subvector $(b_i, \ldots, b_j)$. We use $A = (a_{ij}) \in \mathbb{R}^{n,m}$ to denote a matrix in which the entry of the i-th row and j-th column is a_{ij}. Let A^T denote the transpose of A, and A^{-1} denote the inverse of A if A is non-singular.

Norm. Define the *infinity-norm* $\|\mathbf{b}\|_\infty$ of a vector $\mathbf{b}$ to be $\|\mathbf{b}\|_\infty = max\{|b_i| : 0 \leq i \leq n\}$. For a square $n \times n$ matrix $A = (a_{ij})$, define the following matrix norms:

- *Infinity norm*: $\|A\|_\infty = \sup_{\|x\|_\infty = 1} \|Ax\|_\infty = max_i \sum_{j=1}^{n} |a_{ij}|$.
- *ℓ_2-norm*: $\|A\|_2 = \sup_{\|x\|_2 = 1} \|Ax\|_2$.
- *Frobenius norm*: $\|A\|_F = \left(\sum_{i,j=0}^{n-1} a_{ij}^2\right)^{1/2}$.

It is easy to check that $\|A\|_\infty \leq \|A\|_2$, and that $\|Ax\|_\infty \leq \|A\|_\infty \|x\|_\infty$.

Singular value. For a square $n \times n$ matrix A, the singular values are the square roots of the eigenvalues of $A^T A$.

Fact 5.3.4. *Let $A \in \mathbb{R}^{n \times n}$ be non-singular. Let $\lambda_1 \geq \cdots \geq \lambda_n$ be the singular values of A. Then $\|A^{-1}\|_2 = \frac{1}{\lambda_n}$.*

Lemma 5.3.5 (Yu and Gu [YG97])**.** *Let $A \in \mathbb{R}^{n \times n}$ be non-singular and λ be the smallest singular value of A. Then*

$$\lambda \geq |\det(A)| \cdot \left(\frac{n-1}{\|A\|_F^2}\right)^{(n-1)/2} > 0.$$

Determinant. The determinant of a matrix $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ is $\det(A) = \sum_{\sigma \in S_n} \mathsf{sgn}(\sigma) \prod_{i=1}^{n} a_{i,\sigma_i}$, where S_n is the set of all permutations σ over the set $\{0, \ldots, n-1\}$. For each permutation $\sigma \in S_n$, let σ_i denote the value of the i-th position after reordering by σ. The signature $\mathsf{sgn}(\sigma)$ of a permutation σ is $+1$ if the permutation can be obtained by an even number of swaps between two entries and -1 otherwise.

Proofs for the LP-Based Mechanism

We now prove that the MPC-assisted LP-based mechanism satisfies strict incentive compatibility in an $(h, *, c, d)$-environment. Suppose that each honest user i's true value is sampled independently from $\mathcal{D}_i$, where m denotes the median of the distribution such that $\Pr_{x \sim \mathcal{D}_i}[x \geq m] = \frac{1}{2}$. Moreover, $C_{\mathcal{D}} = \frac{1}{|H|} \sum_{i \in H} \mathop{\mathbb{E}}_{v_i \sim \mathcal{D}_i}[v_i - mmidv_i > m]$, where H denotes the set of actual honest players. Here, $C_{\mathcal{D}}$ is a constant that depends on the distributions of honest users' true values.

Theorem 5.3.6. *Suppose that the block size is infinite. Fix any*[1] $h \geq 2$*, and any* $d \leq \frac{1}{8}\sqrt{\frac{h}{2 \log h}}$*, the MPC-assisted, LP-based mechanism guarantees ex post UIC, Bayesian MIC, and Bayesian SCP in an* $(h, *, *, d)$*-environment, and meanwhile, the mechanism achieves* $\Theta(h \cdot m)$ *expected miner revenue, and at least* $\widetilde{h} \cdot \Theta(C_{\mathcal{D}})$ *expected social welfare for the users, where* $\widetilde{h} \geq h$ *is the actual number of honest users that show up.*

We prove Theorem 5.3.6 in two steps. First, we show that if the linear program defined in Equations (5.1) and (5.2) has a feasible solution, then the resulting mechanism satisfies incentive compatibility, as formally stated below:

Lemma 5.3.7. *When the linear program defined in Equations* (5.1) *and* (5.2) *has a feasible solution, the LP-based mechanism satisfies ex post UIC, Bayesian MIC, and Bayesian SCP in an* $(h, *, *, d)$*-environment. Moreover, the expected miner revenue is* $\frac{h \cdot m}{4}$*, and the user social welfare is* $\Theta(\widetilde{h} \cdot C_{\mathcal{D}})$.

Proof. First, it is easy to see that the expected total miner revenue is $\frac{h \cdot m}{4}$, as guaranteed by the linear program Equations (5.1) and (5.2). Moreover, since the expected utility of a user with true value v is $v - m$ if $v \geq m$, the expected user social welfare is at least

$$\sum_{i \in H} \mathop{\mathbb{E}}_{v_i \sim \mathcal{D}_i}[v_i - mmidv_i > m] = \widetilde{h} \cdot C_{\mathcal{D}}.$$

where H is the set of all honest users.

Next, we prove that the mechanism is strict incentive compatible if the linear program has a solution. UIC is easy to see. Next, we only prove SCP since MIC follows from the same reasoning.

SCP. Since the confirmation and the payment of each bid are independent of other bids, and the mechanism is strict UIC, the coalition cannot increase colluding users' expected utilities. Therefore, we only need to show that the coalition cannot increase the expected total miner revenue by deviating from the mechanism. Intuitively, the linear program Equations (5.1) and (5.2) ensures that for arbitrary d bids, the total miner revenue taking an expectation over the remaining $n - d$ bids always remains $\frac{h \cdot m}{4}$.

Formally, let $\widetilde{h}$ denote the number of *real honest bids* and $\mathbf{b}_{-\mathcal{C}}$ denote the random variable of honest users' bids sampled from $\mathcal{D}_H$. Then $\widetilde{h} \geq n - d = \gamma$. Let $\mathcal{D}_{H_\gamma}$ denote the joint distribution of the last γ number of honest users, and $\mathcal{D}_{-H_\gamma}$ denote the joint distribution of the first $\widetilde{h} - \gamma$

[1] For the special case $h = 1$, we can just use the parity-based mechanism of Figure 5.2.

number of honest users. Then, for any bid $\mathbf{b}_{\mathcal{C}}$ controlled by the coalition, the expected total miner revenue is

$$\mathop{\mathbb{E}}_{\mathbf{b}_{-\mathcal{C}}\sim\mathcal{D}_H}[mu(\mathbf{b}_{-\mathcal{C}},\mathbf{b}_{\mathcal{C}})] = \int_{\mathbf{t}\sim\mathcal{D}_{-H_\gamma}} \mathop{\mathbb{E}}_{\mathbf{b}\sim\mathcal{D}_{H_\gamma}}[mu(\mathbf{b},\mathbf{t},\mathbf{b}_{\mathcal{C}})]\, f(\mathbf{t})d\mathbf{t}, \tag{5.6}$$

where $f(\cdot)$ is the p.d.f. for $\mathcal{D}_{-H_\gamma}$. For any fixed $(\mathbf{t},\mathbf{b}_{\mathcal{C}})$, let I denote the number of bids that are larger than or equal to m in $(\mathbf{t},\mathbf{b}_{\mathcal{C}})$. Since the probability of an honest bid being at least m is exactly $\frac{1}{2}$,

$$\mathop{\mathbb{E}}_{\mathbf{b}\sim\mathcal{D}_{H_\gamma}}[mu(\mathbf{b},\mathbf{t},\mathbf{b}_{\mathcal{C}})] = \sum_{i=0}^{\gamma}\frac{1}{2^\gamma}\binom{\gamma}{i}y_{i+I},$$

which is exactly $\frac{h\cdot m}{4}$ as guaranteed by Equation (5.2). Substituting back into (5.6), for any bid $\mathbf{b}_{\mathcal{C}}$, we have that

$$\mathop{\mathbb{E}}_{\mathbf{b}_{-\mathcal{C}}\sim\mathcal{D}_H}[mu(\mathbf{b}_{-\mathcal{C}},\mathbf{b}_{\mathcal{C}})] = \int_{\mathbf{t}\sim\mathcal{D}_{-H_\gamma}} \frac{h\cdot m}{4}\cdot f(\mathbf{t})d\mathbf{t} = \frac{h\cdot m}{4}.$$

Therefore, for any d bids controlled by the coalition, the expected miner revenue remains $\frac{h\cdot m}{4}$. □

In the main body, we focus on proving the more challenging step, that is, as long as $d \le \frac{1}{8}\sqrt{\frac{h}{2\log h}}$, the linear program indeed has a feasible solution, formally stated below.

Lemma 5.3.8. *For $h \ge 2$ and $d \le \frac{1}{8}\sqrt{\frac{h}{2\log h}}$, the linear program specified by Equations (5.1) and (5.2) is guaranteed to have a feasible solution.*

Proof. We will give a constructive solution to the linear program Equations (5.1) and (5.2). Let $\gamma := n - d$ denote the number of bids that are sampled randomly from $\mathcal{D}$. Let $t = \lfloor\frac{\gamma}{4}\rfloor$, and $\overline{\mu}$ be our target expected miner revenue $\frac{m\cdot h}{4}$. We start from an "approximate" solution $\widehat{\mathbf{y}} = (\widehat{y}_0,\ldots,\widehat{y}_n) \in \mathbb{R}^{n+1}$ such that $\widehat{y}_i = 0$ for any $i \le t$, and $\widehat{y}_i = \overline{\mu}$ for any $i > t$. Our goal is to find a correction $\mathbf{e} = (e_0,\ldots,e_n) \in \mathbb{R}^{n+1}$ that is zero everywhere except for the indices $i \in [z+d, z+2d]$ for some $z \ge \frac{\gamma}{2}$ such that $\widehat{\mathbf{y}} + \mathbf{e}$ is a feasible solution to the linear program Equations (5.1) and (5.2). Henceforth, let $\boldsymbol{\delta} := \mathbf{e}[z+d, z+2d]/\overline{\mu}$ be the non-zero coordinates of the correction, scaled by $\overline{\mu}$. Then $\boldsymbol{\delta}$ must satisfy the linear system $matA(z)\boldsymbol{\delta} = \boldsymbol{\Delta}$, where $A(z)$ and $\boldsymbol{\Delta}$ are defined as follows:

$$matA(z) = \begin{pmatrix} \binom{\gamma}{z+d} & \binom{\gamma}{z+d+1} & \cdots & \binom{\gamma}{z+2d} \\ \binom{\gamma}{z+d-1} & \binom{\gamma}{z+d} & \cdots & \binom{\gamma}{z+2d-1} \\ \vdots & \vdots & \ddots & \vdots \\ \binom{\gamma}{z} & \binom{\gamma}{z+1} & \cdots & \binom{\gamma}{z+d} \end{pmatrix}, \qquad \boldsymbol{\Delta} = \begin{pmatrix} \sum_{i=0}^{t}\binom{\gamma}{i} \\ \sum_{i=0}^{t-1}\binom{\gamma}{i} \\ \vdots \\ \sum_{i=0}^{t-d}\binom{\gamma}{i} \end{pmatrix}.$$

If there exists a $z^* \in [\lceil\frac{\gamma}{2}\rceil, \lceil\frac{\gamma}{2}\rceil + 2d^2]$ such that this linear system $matA(z^*)\boldsymbol{\delta} = \boldsymbol{\Delta}$ has a solution $\boldsymbol{\delta}$, then choosing $\mathbf{e}$ such that $\mathbf{e}[z^*+d : z^*+2d] = \overline{\mu}\cdot\boldsymbol{\delta}$ gives a solution $\widehat{\mathbf{y}} + \mathbf{e}$ that satisfies Equation (5.2).

Claim 5.3.9. *There exists a $z^* \in [\lceil \frac{\gamma}{2} \rceil, \lceil \frac{\gamma}{2} \rceil + 2d^2]$ such that the matrix $matA(z^*)$ is non-singular, and*

$$\|matA(z^*)^{-1}\|_\infty \leq \frac{(z^*+2d)^{2d(d+1)}}{\binom{\gamma}{z^*}} \cdot \left(\frac{d+1}{\sqrt{d}}\right)^d. \tag{5.7}$$

When choosing this z^*, we have a unique solution $\boldsymbol{\delta} = matA(z^*)^{-1}\boldsymbol{\Delta}$. Moreover, under the given parameter range, the solution $\boldsymbol{\delta}$ has bounded infinity norm:

Claim 5.3.10. *For $h \geq 2$ and $d \leq \frac{1}{8}\sqrt{\frac{h}{2\log h}}$, we have $\|\boldsymbol{\delta}\|_\infty \leq 1$.*

For now, we assume that Claim 5.3.9 and Claim 5.3.10 are true, and we show how they lead to Lemma 5.3.8. The proofs of the two claims appear right afterward. To prove Lemma 5.3.8, it suffices to show that $\widehat{\mathbf{y}} + \mathbf{e}$ indeed satisfies the budget feasibility specified by Equation (5.1). Since for all $i \notin [z^*+d, z^*+2d]$, we have $\widehat{y}_i + e_i = \widehat{y}_i \leq i \cdot m$, so we only need to show that for the correction position $z^*+d, \ldots, z^*+2d$, the budget feasibility is satisfied. Substituting $\|\boldsymbol{\delta}\|_\infty \leq 1$, for each $i \in [z^*+d, z^*+2d]$, we have $|e_i| \leq \overline{\mu}$. This implies that $\widehat{y}_i + e_i \geq \overline{\mu} - \overline{\mu} = 0$. Moreover,

$$\widehat{y}_i + e_i \leq 2\overline{\mu} \leq \frac{\gamma}{2} \cdot m \leq i \cdot m.$$

Lemma 5.3.8 thus follows. □

Proof of Claim 5.3.9. We separate the proof in two parts: we first show that there exists a $z^* \in [\lceil \frac{\gamma}{2} \rceil, \lceil \frac{\gamma}{2} \rceil + 2d^2]$ such that $matA(z^*)$ is non-singular; then we show that the infinity norm of the inverse of $A(z^*)$ satisfies Equation (5.7).

Non-singularity. We show that there exists z^* in the given range such that $\det(matA(z^*)) \neq 0$. Define

$$B(z) = \frac{A(z)}{\binom{\gamma}{z}} \cdot \prod_{i=1}^{2d}(z+i).$$

Since

$$\begin{aligned}\frac{\binom{\gamma}{z+j}}{\binom{\gamma}{z}} \cdot \prod_{i=1}^{2d}(z+i) &= \frac{(\gamma-z-j+1)\ldots(\gamma-z)}{(z+1)\ldots(z+j)} \cdot \prod_{i=1}^{2d}(z+i) \\ &= \prod_{i=1}^{j}(\gamma-z-j+i) \cdot \prod_{i=j+1}^{2d}(z+i),\end{aligned}$$

$B(z)$ is equal to the following matrix:

$$\begin{pmatrix} \prod_{i=1}^{d}(\gamma-z-d+i)\prod_{i=d+1}^{2d}(z+i) & \prod_{i=1}^{d+1}(\gamma-z-d-1+i)\prod_{i=d+2}^{2d}(z+i) & \ldots & \prod_{i=1}^{2d}(\gamma-z-2d+i) \\ \prod_{i=1}^{d-1}(\gamma-z-d+1+i)\prod_{i=d}^{2d}(z+i) & \prod_{i=1}^{d}(\gamma-z-d+i)\prod_{i=d+1}^{2d}(z+i) & \ldots & \prod_{i=1}^{2d-1}(\gamma-z-d+1+i)(z+2d) \\ \vdots & \vdots & \ddots & \vdots \\ \prod_{i=1}^{2d}(z+i) & (\gamma-z)\prod_{i=2}^{2d}(z+i) & \ldots & \prod_{i=1}^{d}(\gamma-z-d+i)\prod_{i=d+1}^{2d}(z+i) \end{pmatrix}$$

It is sufficient to show that there exists a $z^* \in [\lceil \frac{\gamma}{2} \rceil, \lceil \frac{\gamma}{2} \rceil + 2d^2]$ such that $\det(B(z^*)) \neq 0$. To show this, note that the determinant of $B(z)$ is a polynomial $q(z)$ of z with a degree at most $2d^2$. As long as $q(z)$ is not a zero polynomial, $q(z)$ has at most $2d^2$ roots. That means, there must exist a $z^* \in [\lceil \frac{\gamma}{2} \rceil, \lceil \frac{\gamma}{2} \rceil + 2d^2]$ such that $q(z^*) \neq 0$. The non-singularity of $A(z^*)$ thus follows.

Hence, it suffices to show that $q(z)$ is not a zero polynomial.

Indeed, when $z = \gamma - d$, the matrix $B(z)$ becomes the following lower triangle matrix, which has a positive determinant.

$$\begin{pmatrix} \prod_{i=1}^{c} i \cdot \prod_{i=c+1}^{2c}(z+i) & 0 & \dots & 0 \\ \prod_{i=1}^{c-1}(i+1) \cdot \prod_{i=c}^{2c}(z+i) & \prod_{i=1}^{c} i \cdot \prod_{i=c+1}^{2c}(z+i) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \prod_{i=1}^{2c}(z+i) & c \cdot \prod_{i=2}^{2c}(z+i) & \dots & \prod_{i=1}^{c} i \cdot \prod_{i=c+1}^{2c}(z+i) \end{pmatrix}$$

This implies that $q(z)$ is not a zero polynomial.

Infinity norm. For simplicity, we use $A := A(z^*)$ in this part. By Fact 5.3.4, $\|matA^{-1}\|_2 = \frac{1}{\lambda}$, where λ is the smallest singular value of $matA$. By Lemma 5.3.5, the smallest singular value λ satisfies

$$\lambda \geq |\det(matA)| \cdot \left(\frac{d}{\|matA\|_F^2} \right)^{\frac{d}{2}}.$$

By the definition of Frobenius norm and the fact that the largest term in $matA$ is $\binom{\gamma}{z^*}$,

$$\|matA\|_F^2 = \sum_{i=0}^{c} \sum_{j=0}^{d} a_{ij}^2 \leq (d+1)^2 \cdot \binom{\gamma}{z^*}^2.$$

We only need to bound the determinant of $matA$. Let $matA' = (a'_{i,j})_{(d+1)\times(d+1)}$ where $a'_{i,j} = \frac{a_{i,j}}{\binom{\gamma}{z^*}}$. Then we have that $|\det(matA)| = \binom{\gamma}{z^*}^{(d+1)} \cdot |\det(matA')|$, where

$$matA' = \begin{pmatrix} \frac{(\gamma-z^*-d+1)\dots(\gamma-z^*)}{(z^*+1)\dots(z^*+d)} & \frac{(\gamma-z^*-d)\dots(\gamma-z^*)}{(z^*+1)\dots(z^*+d+1)} & \dots & \frac{(\gamma-z^*-2d+1)\dots(\gamma-z^*)}{(z^*+1)\dots(z^*+2d)} \\ \frac{(\gamma-z^*-d+2)\dots(\gamma-z^*)}{(z^*+1)\dots(z^*+d-1)} & \frac{(\gamma-z^*-d+1)\dots(\gamma-z^*)}{(z^*+1)\dots(z^*+d)} & \dots & \frac{(\gamma-z^*-2d+2)\dots(\gamma-z^*)}{(z^*+1)\dots(z^*+2d-1)} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \frac{\gamma-z^*}{z^*+1} & \dots & \frac{(\gamma-z^*-d+1)\dots(\gamma-z^*)}{(z^*+1)\dots(z^*+d)} \end{pmatrix}$$

By the definition of determinant, we have

$$\det(matA') = \sum_{\sigma \in S_{d+1}} \mathsf{sgn}(\sigma) \prod_{i=1}^{d+1} a'_{i,\sigma_i}.$$

For each $\sigma \in S_{d+1}$, let q_σ denote the product $\prod_{i=1}^{d+1} a'_{i,\sigma_i}$. Since all the entries in $matA'$ are rational numbers, for each σ, the product q_σ is also a rational number. Note that the denominator of each entry is a factor of $\prod_{i=1}^{2d}(z^*+i)$, we can write each $q_\sigma = \frac{p_\sigma}{\prod_{i=1}^{2d}(z^*+i)^{(d+1)}}$ for an integer p_σ. Thus,

$$\begin{aligned}
|\det(matA')| &= \left|\sum_{\sigma\in S_{d+1}} \mathsf{sgn}(\sigma)\prod_{i=1}^{d+1} a'_{i,\sigma_i}\right| \\
&= \left|\sum_{\sigma\in S_{d+1}} \mathsf{sgn}(\sigma)\frac{p_\sigma}{\prod_{i=1}^{2d}(z^*+i)^{(d+1)}}\right| \\
&= \frac{|\sum_{\sigma\in S_{d+1}} \mathsf{sgn}(\sigma)p_\sigma|}{\prod_{i=1}^{2d}(z^*+i)^{(d+1)}} \geq \frac{1}{\prod_{i=1}^{d}(z^*+i)^{(d+1)}},
\end{aligned}$$

where the last step follows from the fact that $matA'$ is non-singular; thus the absolute value of the nominator is at least 1. Therefore,

$$\begin{aligned}
\lambda &\geq |\det(matA)| \cdot \left(\frac{d}{\|matA\|_F^2}\right)^{\frac{d}{2}} \\
&\geq \binom{\gamma}{z^*}^{(d+1)} \cdot \frac{1}{\prod_{i=1}^{2d}(z^*+i)^{(d+1)}} \cdot \left(\frac{d}{(d+1)^2\binom{\gamma}{z^*}^2}\right)^{\frac{d}{2}} \\
&\geq \binom{\gamma}{z^*} \cdot \frac{1}{(z^*+2d)^{2d(d+1)}} \cdot \left(\frac{\sqrt{d}}{d+1}\right)^{d}.
\end{aligned}$$

The claim thus follows from the fact that $\|matA^{-1}\|_\infty \leq \|matA^{-1}\|_2 = \frac{1}{\lambda}$. □

Proof of Claim 5.3.10. Since $matA(z^*)$ is non-singular, we have $\boldsymbol{\delta} = matA(z^*)^{-1}\boldsymbol{\Delta}$. By properties of matrix norms, we have that

$$\|\boldsymbol{\delta}\|_\infty \leq \|matA(z^*)^{-1}\|_\infty \cdot \|\boldsymbol{\Delta}\|_\infty. \tag{5.8}$$

By Claim 5.3.9 and note that $\|\boldsymbol{\Delta}\|_\infty \leq t \cdot \binom{\gamma}{t}$, we have

$$\|\boldsymbol{\delta}\|_\infty \leq \|matA(z^*)^{-1}\|_\infty \cdot \|\boldsymbol{\Delta}\|_\infty \leq \frac{(z^*+2d)^{2d(d+1)}}{\binom{\gamma}{z^*}} \cdot \left(\frac{d+1}{\sqrt{d}}\right)^{d} \cdot t \cdot \binom{\gamma}{t}.$$

Because $z^*+2d \leq \gamma$, $\frac{d+1}{\sqrt{d}} \leq \gamma$ and $t \leq \gamma$, we have

$$\|\boldsymbol{\delta}\|_\infty \leq \frac{(z^*+2d)^{2d(d+1)}}{\binom{\gamma}{z^*}} \cdot \left(\frac{d+1}{\sqrt{d}}\right)^{d} \cdot t \cdot \binom{\gamma}{t} \leq \gamma^{2d^2+3d+1} \cdot \frac{\binom{\gamma}{t}}{\binom{\gamma}{z^*}}. \tag{5.9}$$

By the assumption that $d \leq \frac{1}{8}\sqrt{\frac{h}{2\log h}}$, we have

$$(2d^2+3d+1)\cdot\log(h-d) + \frac{\log\frac{6}{5}}{4}\cdot d \leq 8d^2\cdot\log h \leq \frac{h}{4}\cdot\log\frac{6}{5}.$$

Re-arrange the inequality and notice that $h - d \leq \gamma$, we have

$$2d^2 + 3d + 1 \leq \frac{(h-d)\log\frac{6}{5}}{4\log(h-d)} \leq \frac{\gamma\log\frac{6}{5}}{4\log\gamma}, \tag{5.10}$$

therefore,

$$\gamma^{2d^2+3d+1} \leq \left(\frac{6}{5}\right)^{\frac{\gamma}{4}}. \tag{5.11}$$

Next, note that for any integers a, b such that $a < b$, we have $\frac{\binom{\gamma}{a}}{\binom{\gamma}{b}} = \frac{(a+1)(a+2)\cdots b}{(\gamma-b+1)(\gamma-b+2)\cdots(\gamma-a)} \leq (\frac{b}{\gamma-a})^{b-a}$. Because $z^* \in [\lceil\frac{\gamma}{2}\rceil, \lceil\frac{\gamma}{2}\rceil + 2d^2]$, we have $\binom{\gamma}{z^*} \geq \binom{\gamma}{\lceil\frac{\gamma}{2}\rceil+2d^2}$. Thus,

$$\begin{aligned}\frac{\binom{\gamma}{t}}{\binom{\gamma}{z^*}} &\leq \frac{\binom{\gamma}{t}}{\binom{\gamma}{\lceil\frac{\gamma}{2}\rceil+2d^2}} \leq \left(\frac{\lceil\frac{\gamma}{2}\rceil + 2d^2}{\gamma - t}\right)^{\lceil\frac{\gamma}{2}\rceil+2d^2-t} \\ &\leq \left(\frac{\frac{\gamma}{2} + 2d^2 + 1}{\frac{3}{4}\gamma}\right)^{\lceil\frac{\gamma}{2}\rceil+2d^2-t}.\end{aligned} \tag{5.12}$$

Because $\gamma \geq h - d$, for any $h \geq 2$ and $d \leq \frac{1}{8}\sqrt{\frac{h}{2\log h}}$, it must be $\frac{\log\frac{6}{5}}{\log\gamma} < \frac{1}{2}$. By Equation (5.10), we have

$$2d^2 + 1 \leq 2d^2 + 3d + 1 \leq \frac{\gamma\log\frac{6}{5}}{4\log\gamma} \leq \frac{\gamma}{8}.$$

By $2d^2 + 1 \leq \frac{\gamma}{8}$ and Equation (5.12), we have

$$\frac{\binom{\gamma}{t}}{\binom{\gamma}{z^*}} \leq \left(\frac{\frac{\gamma}{2} + 2d^2 + 1}{\frac{3}{4}\gamma}\right)^{\lceil\frac{\gamma}{2}\rceil+2d^2-t} \leq \left(\frac{5}{6}\right)^{\lceil\frac{\gamma}{2}\rceil+2d^2-t} \leq \left(\frac{5}{6}\right)^{\frac{\gamma}{4}}. \tag{5.13}$$

Combining Equations (5.9), (5.11), and (5.13), we have that $\|\boldsymbol{\delta}\|_\infty \leq 1$. □

5.4 Characterization for Finite Block Size

In this section, we give a characterization for finite block size. We will first focus on strict incentive compatibility. In an (h, ρ, c, d)-environment, we can indeed circumvent the 0-miner revenue impossibility result Theorem 3.2.4. However, it turns out that for $c = 1$ and $c \geq 2$, the mechanisms are different. Specifically, for $c \geq 2$, each user's utility has to be 0. Therefore, we separately give the mechanisms for $c = 1$ and $c \geq 2$.

5.4.1 Strict Incentive Compatibility: Feasibility for $c = 1$

For $c = 1$, the mechanism is simply the LP-based mechanism in Section 5.3.3 with a random selection process. Still, we assume that honest users' values are sampled independently, and the median m of the distribution satisfies that $\Pr[x \geq m] = \frac{1}{2}$. For convenience, we repeat the MPC-assisted, LP-based mechanism with random selection below.

MPC-assisted, LP-based mechanism with random selection

Parameters: the block size k, the environment parameter $(h, *, 1, d)$, the distribution median m.

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

Mechanism:

- *Allocation Rule.* Let $\widetilde{\mathbf{b}} = (\widetilde{b}_1, \ldots, \widetilde{b}_s)$ denote the bids that are at least m. If $s \leq k$, confirm all bids in $\widetilde{\mathbf{b}}$. Otherwise, randomly select k bids from $\widetilde{\mathbf{b}}$ to confirm.
- *Payment rule.* Each confirmed bid pays m.
- *Miner revenue rule.* Let $\mathbf{y} := (y_0, y_1, \ldots, y_n)$ be any feasible solution to the following linear program:

$$\forall i \in [n] : \ 0 \leq y_i \leq min(i, k) \cdot m \tag{5.14}$$

$$\forall 0 \leq j \leq d : \ \sum_{i=0}^{n-d} q_i \cdot y_{i+j} = \frac{m \cdot min(h, k)}{4} \tag{5.15}$$

where $q_i = \frac{1}{2^{n-d}} \binom{n-d}{i}$ is the probability of observing i heads if we flip $n - d$ independent fair coins. The total miner revenue is y_t where t is the number of confirmed bids in the block.

Figure 5.4: LP-based mechanism with random selection in the MPC-assisted model.

Theorem 5.4.1. *Suppose the block size is k. Fix any[2] $h \geq 2$, and any $d \leq \frac{1}{8}\sqrt{\frac{h}{2 \log h}}$. The MPC-assisted, LP-based mechanism with random selection is ex post UIC, Bayesian MIC, and Bayesian SCP in an $(h, *, 1, d)$-environment. Moreover, the expected miner revenue is $\Theta(min\{h, k\})$.*

Proof. First, Equation (5.14) guarantees that total miner revenue is at most the total payment of the confirmed users, so the mechanism satisfies budget feasibility.

Next, we show that when the linear program Equations (5.14) and (5.15) has a solution, the mechanism satisfies all three incentive-compatible properties.

- **UIC:** By the same reasoning as in Theorem 5.3.1, overbidding or underbidding does not increase the user's utility. Injecting bids cannot increase the user's utility either: it may only decrease the probability that the user gets confirmed. Moreover, dropping out can only give the user zero utility. Therefore, a user cannot increase its utility by deviating.
- **SCP:** By the same reasoning as in the proof of Lemma 5.3.7, the linear program Equation (5.15) guarantees that no matter how the coalition chooses the d bids it controls, the expected total miner revenue remains unchanged. Meanwhile, the coalition cannot increase the colluding user's utility by UIC. Therefore, this mechanism is SCP.
- **MIC:** Follows by the same reasoning as SCP.

[2]For the special case $h = 1$, we can just use the parity-based mechanism of Figure 5.2 with the random selection.

It remains to show that the linear program indeed has a feasible solution. We will give a constructive solution. Let $\widetilde{\mathbf{y}} = (\widetilde{y}_1, \ldots, \widetilde{y}_n)$ denote the constructive solution given in the proof of Lemma 5.3.8 that satisfies

$$\forall 0 \leq j \leq d: \quad \sum_{i=0}^{n-d} q_i \cdot \widetilde{y}_{i+j} = \frac{m \cdot h}{4}.$$

In the proof of Lemma 5.3.8, $\widetilde{\mathbf{y}}$ satisfies that $0 \leq \widetilde{y}_i \leq min(i, h) \cdot m$ for any $0 \leq i \leq n$. There are two possible cases.

- $h \leq k$. We have $0 \leq \widetilde{y}_i \leq min(i, h) \cdot m \leq min(i, k) \cdot m$. Thus, $\widetilde{\mathbf{y}}$ is a feasible solution to the linear program in this case.
- $h > k$. Let $\mathbf{y} = (y_0, \ldots, y_n) = \frac{k}{h} \cdot \widetilde{\mathbf{y}}$. Then y_i satisfies that $0 \leq y_i \leq \frac{k}{h} min(i, h) \cdot m \leq min(i, k) \cdot m$. Moreover, for any $0 \leq j \leq d$,

$$\sum_{i=0}^{n-d} q_i \cdot y_{i+j} = \frac{k}{h} \cdot \sum_{i=0}^{n-d} q_i \cdot \widetilde{y}_{i+j} = \frac{m \cdot k}{4}.$$

  Thus, $\mathbf{y} = \frac{k}{h} \cdot \widetilde{\mathbf{y}}$ is a feasible solution to the linear program if $h > k$.

□

5.4.2 Zero Social Welfare for Users When $c \geq 2$

Unfortunately, the above MPC-assisted, LP-based mechanism with random selection only works for $c = 1$. When $c \geq 2$, although deviating cannot increase the expected total miner revenue, the coalition can increase a colluding user's utility. Imagine that the coalition consists of some colluding miners and two users i and j, where user i has true value m and user j has a large true value. Then user i may choose to drop out to increase the probability of user j getting confirmed. This strictly increases the expected joint utility of the coalition.

Therefore, to construct a Bayesian SCP mechanism in an (h, ρ, c, d)-environment for $d \geq c \geq 2$, we need to make sure that deviating cannot increase a colluding user's utility. Indeed, for some (contrived) distributions, we can construct a mechanism that generates optimal miner revenue and achieves UIC, MIC, and SCP in an (h, ρ, c, d)-environment for $d \geq c \geq 2$. However, the total social welfare for all users is 0. For example, imagine that honest users' true values are drawn i.i.d. from $\text{Bernoulli}(\frac{1}{2})$. Now, if we run the MPC-assisted, LP-based mechanism with random selection (see Section 5.4.1) and set $m = 1$, the resulting mechanism achieves ex post UIC, Bayesian MIC, and Bayesian SCP in $(h, *, c, d)$-environments, even when $c \geq 2$ (as long as the condition $d \leq \frac{1}{8}\sqrt{\frac{h}{2 \log h}}$ is satisfied). This is because setting $m = 1$ makes sure that every user's utility is always 0. Thus, no matter how the coalition deviates, it cannot increase the strategic users' joint utility. Moreover, as long as the linear program Equations (5.14) and (5.15) has a feasible solution, the coalition cannot increase the expected total miner revenue either. The mechanism achieves $\Theta(m)$ expected miner revenue but unfortunately, the total user social welfare is always 0. It turns out that this zero user social welfare limitation is. In the proof below, for a set S of players, we use $\mathcal{D}_S$ to denote the joint distribution of their true values.

Theorem 5.4.2. *Suppose that the block size is finite, and fix any $h \geq 1$, any $d \geq c \geq 2$, and any $\rho \in (0,1)$. Then, any MPC-assisted TFM that simultaneously satisfies Bayesian UIC, MIC, and SCP in an (h, ρ, c, d)-model must suffer from 0 social welfare for the users when there actually are more than h honest bids. Equivalently, for any set S of $\ell > h$ number of players,*

$$\mathop{\mathbb{E}}_{\mathbf{b} \sim \mathcal{D}_S}[\mathsf{USW}(\mathbf{b})] = 0, \tag{5.16}$$

In the above, $\mathsf{USW}(\mathbf{b})$ denotes the expected total user social welfare under the bid vector $\mathbf{b}$ where the expectation is taken over the randomness of the mechanism.

The proof is similar to the proof of Theorem 3.4.4. Although Theorem 3.4.4 considers a universal MPC-assisted mechanism, the proof also holds for MPC-assisted TFM in an (h, ρ, c, d)-environment for $d \geq c \geq 2$. Henceforth, we use $\mathsf{util}^i(\mathbf{b})$ to denote the utility of identity i when the input bid vector is $\mathbf{b}$. In the proof, we use v_{id} (b_{id}) to denote a bid v (b) coming from identity id.

Lemma 5.4.3. *Fix any $h \geq 1$, any $d \geq c \geq 2$, any $\rho \in (0,1)$. Given any (possibly random) MPC-assisted mechanism that is Bayesian UIC, MIC and SCP in an (h, ρ, c, d)-environment, for any identity i and identity j, for any bid b_j and b'_j, for any set S of ℓ number of users not including i and j,*

$$\mathop{\mathbb{E}}_{(v_i, \mathbf{b}_{-i,j}) \sim \mathcal{D}_{S \cup \{i\}}}[\mathsf{util}^i(v_i, b_j, \mathbf{b}_{-i,j})] = \mathop{\mathbb{E}}_{(v_i, \mathbf{b}_{-i,j}) \sim \mathcal{D}_{S \cup \{i\}}}[\mathsf{util}^i(v_i, b'_j, \mathbf{b}_{-i,j})], \tag{5.17}$$

where $\mathbf{b}_{-i,j}$ represents all except identity i and j's bids and $\mathcal{D}^{(\ell+1)}$ denote the joint distribution of $\ell + 1$ number of users. Moreover, it must be that

$$\mathop{\mathbb{E}}_{(v_i, \mathbf{b}_{-i,j}) \sim \mathcal{D}_{S \cup \{i\}}}[\mathsf{util}^i(v_i, b_j, \mathbf{b}_{-i,j})] = \mathop{\mathbb{E}}_{(v_i, \mathbf{b}_{-i}) \sim \mathcal{D}_{S \cup \{i\}}}[\mathsf{util}^i(v_i, \mathbf{b}_{-i})]. \tag{5.18}$$

Proof. The proof to this lemma is the same as Lemma 3.4.1, except that now we need to guarantee that at least h bids are sampled randomly from $\mathcal{D}$. □

Corollary 5.4.4. *Fix any $h \geq 1$, any $d \geq c \geq 2$, any $\rho \in (0,1)$. Given any (possibly random) MPC-assisted mechanism that is Bayesian UIC, MIC and SCP in an (h, ρ, c, d)-environment, for any two sets H and H' consisting of at least h identities, let $\mathbf{b}_H$ ($\mathbf{b}_{H'}$) denote the bids from identities in H (H'). For any $i \notin H \cup H'$, it must be that*

$$\mathop{\mathbb{E}}_{(v_i, \mathbf{b}_H) \sim \mathcal{D}_{H \cup \{i\}}}[\mathsf{util}^i(v_i, \mathbf{b}_H)] = \mathop{\mathbb{E}}_{(v_i, \mathbf{b}_{H'}) \sim \mathcal{D}_{H' \cup \{i\}}}[\mathsf{util}^i(v_i, \mathbf{b}_{H'})],$$

where v_i denotes that identity i bids v.

Proof. Let $S = H' \setminus H$. Without loss of generality, we assume that S consists of identities $1, \ldots, |S|$. By the definition of the total expectation, we have

$$
\begin{aligned}
&\mathop{\mathbb{E}}_{(v_i,\mathbf{b}_S,\mathbf{b}_H)\sim\mathcal{D}_{S\cup H\cup\{i\}}}[\mathsf{util}^i(v_i,\mathbf{b}_S,\mathbf{b}_H)]\\
&=\int_0^\infty \mathop{\mathbb{E}}_{(v_i,\mathbf{b}_{S\setminus\{1\}},\mathbf{b}_H)\sim\mathcal{D}_{S\cup H}}[\mathsf{util}^i(v_i,z_1,\mathbf{b}_{S\setminus\{1\}},\mathbf{b}_H)]f(z_1)dz_1\\
&=\mathop{\mathbb{E}}_{(v_i,\mathbf{b}_{S\setminus\{1\}},\mathbf{b}_H)\sim\mathcal{D}_{S\cup H}}[\mathsf{util}^i(v_i,b_1,\mathbf{b}_{S\setminus\{1\}},\mathbf{b}_H)]\int_0^\infty f(z_1)dz_1 \qquad \text{By Equation (5.17)}\\
&=\mathop{\mathbb{E}}_{(v_i,\mathbf{b}_{S\setminus\{1\}},\mathbf{b}_H)\sim\mathcal{D}_{S\cup H}}[\mathsf{util}^i(v_i,b_1,\mathbf{b}_{S\setminus\{1\}},\mathbf{b}_H)]\\
&=\mathop{\mathbb{E}}_{(v_i,\mathbf{b}_{S\setminus\{1\}},\mathbf{b}_H)\sim\mathcal{D}_{S\cup H}}[\mathsf{util}^i(v_i,\mathbf{b}_{S\setminus\{1\}},\mathbf{b}_H)] \qquad \text{By Equation (5.18)}\\
&=\cdots=\mathop{\mathbb{E}}_{(v_i,\mathbf{b}_H)\sim\mathcal{D}_{H\cup\{i\}}}[\mathsf{util}^i(v_i,\mathbf{b}_H)].
\end{aligned}
$$

By the same reasoning, consider $S' = H \setminus H'$. Then it must be that

$$
\mathop{\mathbb{E}}_{(v_i,\mathbf{b}_{S'},\mathbf{b}_{H'})\sim\mathcal{D}_{S'\cup H'\cup\{i\}}}[\mathsf{util}^i(v_i,\mathbf{b}_{S'},\mathbf{b}_{H'})] = \mathop{\mathbb{E}}_{(v_i,\mathbf{b}_{H'})\sim\mathcal{D}_{H'\cup\{i\}}}[\mathsf{util}^i(v_i,\mathbf{b}_{H'})].
$$

Note that $S' \cup H' = S \cup H = H' \cup H$. Hence,

$$
\mathop{\mathbb{E}}_{(v_i,\mathbf{b}_{S'},\mathbf{b}_{H'})\sim\mathcal{D}_{S'\cup H'\cup\{i\}}}[\mathsf{util}^i(v_i,\mathbf{b}_{S'},\mathbf{b}_{H'})] = \mathop{\mathbb{E}}_{(v_i,\mathbf{b}_S,\mathbf{b}_H)\sim\mathcal{D}_{S\cup H\cup\{i\}}}[\mathsf{util}^i(v_i,\mathbf{b}_S,\mathbf{b}_H)].
$$

Combining the equalities, we have

$$
\begin{aligned}
&\mathop{\mathbb{E}}_{(v_i,\mathbf{b}_H)\sim\mathcal{D}_{H\cup\{i\}}}[\mathsf{util}^i(v_i,\mathbf{b}_H)]\\
&=\mathop{\mathbb{E}}_{(v_i,\mathbf{b}_S,\mathbf{b}_H)\sim\mathcal{D}_{S\cup H\cup\{i\}}}[\mathsf{util}^i(v_i,\mathbf{b}_S,\mathbf{b}_H)]\\
&=\mathop{\mathbb{E}}_{(v_i,\mathbf{b}_{S'},\mathbf{b}_{H'})\sim\mathcal{D}_{S'\cup H'\cup\{i\}}}[\mathsf{util}^i(v_i,\mathbf{b}_{S'},\mathbf{b}_{H'})]\\
&=\mathop{\mathbb{E}}_{(v_i,\mathbf{b}_{H'})\sim\mathcal{D}_{H'\cup\{i\}}}[\mathsf{util}^i(v_i,\mathbf{b}_{H'})]
\end{aligned}
$$

The corollary thus follows. □

This corollary implies that when identity i's bid is sampled from $\mathcal{D}_i$ in a world with h or more random bids, its expected utility only depends on its identity. Henceforth, fix an arbitrary set $\widetilde{H}$ of $h+1$ number of players. For any $i \in H$, we use the following notation to denote this utility (where the notation v_i means identity i is bidding the value v):

$$
U_i := \mathop{\mathbb{E}}_{(v_i,\mathbf{b}_H)\sim\mathcal{D}_{H\cup\{i\}}}[\mathsf{util}^i(v_i,\mathbf{b}_H)]. \tag{5.19}
$$

Lemma 5.4.5. *Fix any $h \geq 1$, any $d \geq c \geq 2$, any $\rho \in (0,1)$. Given any (possibly random) MPC-assisted mechanism that is Bayesian UIC, MIC and SCP in an (h,ρ,c,d)-environment, for any user i, j, it must be that*

$$
U_i = U_j. \tag{5.20}
$$

Proof. By our symmetric assumption, it must be that

$$\mathop{\mathbb{E}}_{\mathbf{b}_H \sim \mathcal{D}_H} [\mathsf{USW}(v_i, \mathbf{b}_H)] = \mathop{\mathbb{E}}_{\mathbf{b}_H \sim \mathcal{D}_H} [\mathsf{USW}(v_j, \mathbf{b}_H)], \tag{5.21}$$

where v_i (v_j) denotes that identity i (j) bids v, and $\mathsf{USW}(\mathbf{b})$ denotes the expected social welfare for all users when the input bid vector is $\mathbf{b}$. For any identity $l \in H$, for any v_i from identity i, let $H' = H \setminus \{l\}$. It must be

$$\begin{aligned}\mathop{\mathbb{E}}_{(b_l, \mathbf{b}_{H'}) \sim \mathcal{D}_H} [\mathsf{util}^l(b_l, v_i, \mathbf{b}_{H'})] &= \mathop{\mathbb{E}}_{(b_l, \mathbf{b}_{H'}) \sim \mathcal{D}_H} [\mathsf{util}^l(b_l, \mathbf{b}_{H'})] && \text{By Equation (5.18)} \\ &= \mathop{\mathbb{E}}_{(b_l, \mathbf{b}_{H'}) \sim \mathcal{D}_H} [\mathsf{util}^l(b_l, v_j, \mathbf{b}_{H'})]\end{aligned}$$

Thus, for any value v, the sum of the expected utility of every user in H is

$$\sum_{l \in H} \mathop{\mathbb{E}}_{\mathbf{b}_H \sim \mathcal{D}_H} [\mathsf{util}^l(v_i, \mathbf{b}_H)] = \sum_{l \in H} \mathop{\mathbb{E}}_{\mathbf{b}_H \sim \mathcal{D}_H} [\mathsf{util}^l(v_j, \mathbf{b}_H)].$$

Combining this with Equation (5.21), it must be that for any v_i and v_j (which denote that identity i and j bid value v, respectively),

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

The lemma follows by taking expectations over v on both sides. □

Lemma 5.4.6. *Fix any $h \geq 1$, any $d \geq c \geq 2$, any $\rho \in (0, 1)$, and suppose that the distribution $\mathcal{D}$ has bounded support. Given any (possibly random) MPC-assisted mechanism that is Bayesian UIC, MIC and SCP in an (h, ρ, c, d)-environment, for any identity i, it must be that*

$$U_i = 0.$$

Proof. The proof is similar to Theorem 3.4.4. Consider a crowded world with many users and all of their bids are sampled independently at random from $\mathcal{D}$. Let K be the total number of users. By Corollary 5.4.4 and Lemma 5.4.5, every user's expected utility is the same where the expectation is taken over the random coins for sampling all bids as well as random coins of the mechanism. On the other hand, since there are K total bids, there must exist a user whose confirmation probability is at most k/K, and thus its expected utility is at most $M \cdot k/K$ where k is the block size. The lemma follows by taking K to be arbitrarily large. □

Proof of Theorem 5.4.2 Fix any set H of size at least $h + 1$. Then

$$\mathop{\mathbb{E}}_{\mathbf{b} \in \mathcal{D}_H} [\mathsf{USW}(\mathbf{b})] = \sum_{i \in H} \mathop{\mathbb{E}}_{\mathbf{b} \in \mathcal{D}_H} \mathsf{util}^i(\mathbf{b}).$$

By Lemma 5.4.6, for each identity i in H,

$$\mathop{\mathbb{E}}_{\mathbf{b} \in \mathcal{D}_H} \mathsf{util}^i(\mathbf{b}) = U_i = 0.$$

Therefore, the user social welfare is 0.

5.4.3 Feasibility for Approximate IC: Diluted Threshold-Based Mechanism

Although there is no interesting mechanism for strict incentive compatibility when $c \geq 2$, there are meaningful mechanisms if we allow approximate incentive compatibility. Still, we assume that honest users' values are sampled independently and m is the median of the distribution. Without loss of generality, we assume the maximum of the true values $M \geq \epsilon$.

MPC-assisted, diluted threshold-based Mechanism

Parameters: the block size k, the environment parameter $(h, *, c, *)$, the approximation parameter ϵ, the distribution median m, and the upper bound M of the distribution.
Input: a bid vector $\mathbf{b} = (b_1, \ldots, b_n)$.
Mechanism:

- *Confirmation rule.* Let $R := max\left(2c\sqrt{\frac{kM}{\epsilon}}, k\right)$. Given a bid vector $\mathbf{b}$, let $\widetilde{\mathbf{b}} = (\widetilde{b}_1, \ldots, \widetilde{b}_s)$ denote the bids that are at least m. If $s \leq R$, randomly select $\frac{k}{R} \cdot s$ bids from $\widetilde{\mathbf{b}}$ to confirm; otherwise, randomly select k bids from $\widetilde{\mathbf{b}}$ to confirm.
- *Payment rule.* Every confirmed bid pays m.
- *Miner revenue rule.* If $s \geq \frac{h}{4}$, the total miner revenue is $\overline{\mu} := m \cdot min\left(\frac{h}{4} \cdot \frac{k}{R}, k\right)$. Otherwise, the total miner revenue is 0.

Figure 5.5: Diluted threshold-based mechanism in the MPC-assisted model.

Theorem 5.4.7. *Suppose the block size is k. For any $h \geq 1$, $c \geq 1$, and $\epsilon \geq m \cdot \frac{h}{2} \cdot e^{-\frac{h}{16}}$, the diluted threshold posted price auction satisfies strict ex post UIC, Bayesian ϵ-MIC, and Bayesian ϵ-SCP in an $(h, *, c, *)$-environment. Moreover, the expected total miner revenue is $m \cdot min\left(\frac{h\sqrt{k\epsilon}}{8c\sqrt{M}}, \frac{h}{4}, k\right)$, where M is the upper bound of users' true values.*

Proof. We first show that the budget feasibility is satisfied. Since the mechanism confirms $min\left(s \cdot \frac{k}{R}, k\right)$ number of bids that are at least m, the total payment is $m \cdot min\left(s \cdot \frac{k}{R}, k\right)$. When $s \geq \frac{h}{4}$, the total miner revenue is at most $m \cdot \frac{h}{4} \cdot \frac{k}{R} \leq m \cdot s \cdot \frac{k}{R}$, which is no more than the total payment of the users. Next, we prove UIC, MIC, and SCP separately.

UIC. Since the mechanism is posted price auction from each user's perspective, each user's best response is to follow the protocol honestly, as in the proof of Theorem 5.4.1.

MIC. By the same reasoning as in Theorem 5.3.2, by injecting fake bids, the miner can only increase its expected miner revenue if the number of bids that are at least m from honest users is less than $\frac{h}{4}$. This happens with a probability at most $e^{-\frac{h}{16}}$. Thus, the expected total miner revenue increases by no more than

$$\overline{\mu} \cdot e^{-\frac{h}{16}} \leq m \cdot e^{-\frac{h}{16}} \cdot \frac{h}{4} \leq \frac{\epsilon}{2}.$$

SCP. By the same reasoning as in MIC, the expected increase of the miner revenue is at most $\epsilon/2$ by any deviation. Thus, to show that the mechanism is Bayesian ϵ-SCP, it suffices to show that the coalition cannot increase the joint utility of the "users" in the coalition by more than $\frac{\epsilon}{2}$.

Because injecting bids smaller than m does not change the confirmation probability and the payment of each confirmed bid is fixed, injecting bids smaller than m does not increase the users' utilities. On the other hand, injecting bids at least m will only decrease the probability of each colluding user getting confirmed, which does not increase the users' utilities.

Now, it suffices to show that overbidding and underbidding do not increase the coalition's joint utility since dropping out is equivalent to underbidding to some value less than m. Let s be the number of bids $\geq m$ when every user bids truthfully. Each bid is confirmed with probability $\frac{k}{R}$ if $s \leq R$, and $\frac{k}{s}$ if $s > R$. Let s' be the number of bids $\geq m$ when the colluding users bid strategically.

The colluding users can be partitioned into four groups:

- S_1: Those whose true values are less than m but overbid to values larger than or equal to m;
- S_2: Those whose true values are less than m and bid values less than m;
- S_3: Those whose true values are at least m but underbids to values less than m;
- S_4: Those whose true values are at least m and still bid values at least m.

When the coalition bids strategically, only the utilities of the users in S_4 increase compared to the honest case. Consider a colluding user in S_4 with the true value $v \geq m$. Its utility increases by at most

$$(v-m) \cdot \frac{k}{max\{s', R\}} - (v-m) \cdot \frac{k}{max\{s, R\}}. \tag{5.22}$$

Note that Equation (5.22) is positive only when $s' < s$ and $s > R$. In this case, Equation (5.22) can be upper bounded by

$$\begin{aligned}
&(v-m)\left[\frac{k}{max\{s', R\}} - \frac{k}{s}\right] \leq (M-m)\left[\frac{k}{s'} - \frac{k}{s}\right] \\
\leq &(M-m) \cdot \frac{ck}{s(s-c)} && \text{By } s' \geq s - c. \\
\leq &M \cdot \frac{ck}{R(R-c)}.
\end{aligned}$$

Since by the choice of R, $R(R-c) \geq \frac{1}{2}R^2$, we have

$$\text{Equation (5.22)} \leq M \cdot \frac{2ck}{R^2} \leq \frac{\epsilon}{2c}.$$

Therefore, by bidding untruthfully, each user's utility can increase by at most $\frac{\epsilon}{2c}$. Therefore, the joint utility of the users in the coalition by more than $\frac{\epsilon}{2}$. □

5.5 Bounds on Miner Revenue

In this section, we prove bounds on the miner revenue under different settings. Henceforth, let $\mu(\mathbf{b})$ denote the expected total miner revenue when the input bid vector is $\mathbf{b}$, where the expectation is taken over the mechanism's randomness.

5.5.1 Bounds on Miner Revenu in (h, ρ, c, d)-Environment

In this section, we prove limits on miner revenue under the (h, ρ, c, d)-environment.

Theorem 5.5.1. *Fix any $h \geq 1$, $d \geq c \geq 1$, and $\rho \in (0, 1)$. Given any MPC-assisted mechanism that is Bayesian ϵ_u-UIC, Bayesian ϵ_m-MIC and Bayesian ϵ_s-SCP in an (h, ρ, c, d)-environment, for all $n \geq h$, it must be that*

$$\mathop{\mathbb{E}}_{\mathbf{b}\sim\mathcal{D}}[\mu(\mathbf{b})] \leq h \cdot E_{\mathcal{D}} + \frac{2(n-h)}{\rho}\left(\epsilon + C_{\mathcal{D}}\sqrt{\epsilon}\right), \tag{5.23}$$

where $\epsilon = \epsilon_u + \epsilon_m + \epsilon_s$, $\mathcal{D} = \mathcal{D}_1 \times \cdots \times \mathcal{D}_n$ is the joint distribution of n users' true values, $E_{\mathcal{D}} = \frac{1}{h}\sum_{i=1}^{h} \mathbb{E}_{X\sim\mathcal{D}_i}[X]$ and $C_{\mathcal{D}} = \frac{1}{n-h}\sum_{i=h+1}^{n} \mathbb{E}_{X\sim\mathcal{D}_i}[\sqrt{X}]$ are the terms that depend on the "scale" of the distribution $\mathcal{D}$. As a special case, for strict incentive compatibility where $\epsilon = 0$, we have that

$$\mathop{\mathbb{E}}_{\mathbf{b}\sim\mathcal{D}^n}[\mu(\mathbf{b})] \leq h \cdot \mathbb{E}(\mathcal{D}). \tag{5.24}$$

Proof. The proof is almost the same as Theorem 4.2.4. Recall that in Theorem 4.2.4, we have

$$\mathop{\mathbb{E}}_{\mathbf{b}\sim\mathcal{D}^{(n)}}[\mu(\mathbf{b})] \leq \mathop{\mathbb{E}}_{\mathbf{b}'\sim\mathcal{D}^{(n-1)}}[\mu(\mathbf{b}')] + \frac{2\epsilon}{\rho} + \frac{2\sqrt{\epsilon}}{\rho}\mathbb{E}_{X\sim\mathcal{D}_n}[\sqrt{X}]. \tag{5.25}$$

The only difference here is that we cannot apply the above step for n times as the environment guarantees the existence of at least h honest users. Therefore, we can only apply Equation (5.25) for $n - h$ times and get the desired result. □

5.5.2 Necessity of Bayesian Incentive Compatibility

If we insist on ex post incentive compatibility, the additional reasonable-world assumption will not help us overcome the limit on miner revenue for mechanisms that are universal in h (Theorem 4.2.4).

Theorem 5.5.2. *Fix any $h \geq 1$, $d \geq c \geq 1$ and $\rho \in (0, 1)$. Given any (possibly randomized) MPC-assisted TFM that is ex post ϵ_u-UIC and ex post ϵ_s-SCP in an (h, ρ, c, d)-environment, it must be that for any $\mathbf{b} = (b_1, \ldots, b_n)$ of length $n > h$,*

$$\mu(\mathbf{b}) \leq \frac{2n\epsilon}{\rho} + \frac{2\sqrt{\epsilon}}{\rho}\sum_{i=1}^{n}\sqrt{b_i}, \tag{5.26}$$

where $\epsilon = \epsilon_s + \epsilon_u$.

As a special case, for strict incentive compatibility where $\epsilon = 0$, we have that for any bid $\mathbf{b}$,

$$\mu(\mathbf{b}) = 0.$$

Proof. Recall that Lemma 4.2.3 still holds in the ex post setting. For any $\mathbf{b} = (b_1, \ldots, b_n)$ of length $n > h$, it must be that

$$\begin{aligned}
&\mu(\mathbf{b}) = \mu(b_1, b_2, \ldots, b_n) \\
\leq &\mu(b_1, \ldots, b_{n-1}, 0) + \frac{2}{\rho}\epsilon + \frac{2}{\rho}\sqrt{b_n \epsilon} && \text{By Lemma 4.2.3} \\
\leq &\mu(b_1, \ldots, b_{n-2}, 0, 0) + \frac{4}{\rho}\epsilon + \frac{2}{\rho}\sqrt{b_n \epsilon} + \frac{2}{\rho}\sqrt{b_{n-1}\epsilon} \\
\leq &\cdots \leq \mu(0 \ldots, 0) + \frac{2n\epsilon}{\rho} + \frac{2\sqrt{\epsilon}}{\rho}\sum_{i=1}^{n}\sqrt{b_i} \\
\leq &\frac{2n\epsilon}{\rho} + \frac{2\sqrt{\epsilon}}{\rho}\sum_{i=1}^{n}\sqrt{b_i}.
\end{aligned}$$

□

5.5.3 Bounds on Miner Revenue if Assuming Honest Majority of Bids

As mentioned in Section 1.1.3, we consider a "sufficient honesty" assumption but the precise statement of the assumption matters. In particular, had we assumed that the majority of bids are submitted by honest users (referred to as the "honest majority bids" assumption), then we would not be able to overcome the severe limitation on miner revenue. The theorem below states that under the honest majority bids assumption, we should still suffer from a constant miner revenue limitation.

Theorem 5.5.3. *Assuming that a majority number of bids are submitted by honest users. Again, we use $\mathcal{D} = \mathcal{D}_1 \times \cdots \times \mathcal{D}_n$ to denote the joint distribution of n users' true values. If a mechanism is Bayesian UIC, Bayesian MIC, and Bayesian SCP (even for $c = 1$) under the "honest majority bids" assumption, it must be that $\mathop{\mathbb{E}}_{\mathbf{b}\sim\mathcal{D}}[\mu(\mathbf{b})] \leq \mathbb{E}[\mathcal{D}_1] + \mathbb{E}[\mathcal{D}_2]$.*

Proof. The proof is based on the following lemma.

Lemma 5.5.4. *Assuming that a majority of bids are submitted by honest users. Suppose that a TFM is Bayesian UIC, Bayesian MIC, and Bayesian SCP (even for $c = 1$) under the "honest majority bids" assumption. Then, as long as the number of bids $n \geq 3$, it must be that $\mathop{\mathbb{E}}_{\mathbf{b}\sim\mathcal{D}}[\mu(\mathbf{b})] \leq \mathop{\mathbb{E}}_{\mathbf{b}'\sim\mathcal{D}_{-n}}[\mu(\mathbf{b}')]$, where $\mathcal{D}_{-n}$ denotes the distribution of all but user n's true values.*

For now assume that Lemma 5.5.4 holds, and we will show how the theorem follows. The proof of Lemma 5.5.4 appears afterwards. By induction on n, for any $n \geq 3$,

$$\mathop{\mathbb{E}}_{\mathbf{b}\sim\mathcal{D}}[\mu(\mathbf{b})] \leq \mathop{\mathbb{E}}_{\mathbf{b}'\sim\mathcal{D}_1\times\mathcal{D}_2}[\mu(\mathbf{b}')] \leq \mathbb{E}[\mathcal{D}_1] + \mathbb{E}[\mathcal{D}_2],$$

where the last inequality follows from the fact that the miner revenue must be upper bounded by the bids. □

Proof of Lemma 5.5.4. The proof of this is similar to Theorem 3.2.4. Suppose that the number of bids is some arbitrary n. For any $n \geq 3$, we have

$$\begin{aligned}
\mathop{\mathbb{E}}_{\mathbf{b}\sim\mathcal{D}}[\mu(\mathbf{b})] &= \int_0^{+\infty} \mathop{\mathbb{E}}_{\mathbf{b}'\sim\mathcal{D}_{-n}}[\mu(\mathbf{b}', r)] f(r) dr \\
&= \int_0^{+\infty} \mathop{\mathbb{E}}_{\mathbf{b}'\sim\mathcal{D}_{-n}}[\mu(\mathbf{b}', 0)] f(r) dr && \text{By Lemma 3.2.3} \\
&= \mathop{\mathbb{E}}_{\mathbf{b}'\sim\mathcal{D}_{-n}}[\mu(\mathbf{b}', 0)].
\end{aligned}$$

By Bayesian MIC, it must be that $\mathop{\mathbb{E}}_{\mathbf{b}'\sim\mathcal{D}_{-n}}[\mu(\mathbf{b}', 0)] \leq \mathop{\mathbb{E}}_{\mathbf{b}'\sim\mathcal{D}_{-n}}[\mu(\mathbf{b}')]$. Otherwise, when there are $n-1$ honest bids, the miners can inject a 0 and increase the miner revenue while it does not need to pay anything for injecting the 0-bid. This violates MIC. Notice that since $n - 1 \geq 2$, the miners injecting one bid does not violate the "honest majority bids" assumption. □

Chapter 6

Characterization of Miner-User Coalition Proofness

So far we have explored several different relaxations. There are severe limitations on user social welfare for strict incentive compatibility for finite block size: the user social welfare must be 0 when $c \geq 2$ and when there are enough users. This limitation holds no matter whether we assume reasonable-world assumptions or not. The fundamental reason behind this impossibility result is that a coalition with 2 or more users may choose a strategy that drops off some bids with small true values to increase the chance that those with large true values get confirmed. This prompts the following question:

> *Is this attack really unacceptable? Is there a meaningful notion for the miner-user coalition that tolerates such kinds of attacks?*

6.1 Define Miner-User Coalition Proofness

We define a notion called miner-user coalition proofness (MUCP) that captures the intuition that any miner-user coalition is unstable. Roughly speaking, for any miner-user coalition and any coalition's strategy $S_{\mathcal{C}}$, either the users have a better strategy such that the users can do better, or the miners have a better strategy such that the miners can do better. Recall that in the plain model, the miner's strategy, or a miner-user coalition's strategy can be represented by a pair $(I^*, \mathbf{b}^*)$, where I^* denotes the strategic inclusion rule and $\mathbf{b}^*$ denotes the (possibly empty) strategic bid vector. In the MPC-assisted model, miners' strategy, or a miner-user coalition's strategy can be simply represented by $\mathbf{b}^*$, which denotes the (possibly empty) strategic bid vector.

We assume that there is a "negotiation" phase before the transaction fee mechanism where the miners $\mathcal{M}$ may attempt to sign a binding side contract/form a coalition with a set of users $\mathcal{U}$. During this process, they may decide their bidding strategy, inclusion strategy for the miners if in the plain model, and off-chain payments. Therefore, a user i in this miner-user coalition now gets a utility $(\sum_{b_j \text{ submitted by } i} x_j v_j - p_j) +$ off-chain payment it receives, where x_j, v_j and p_j denotes b_j's confirmation probability, true value, and expected payment. A miner in the miner-user coalition now gets a utility equal to the miner revenue it gets and the off-chain payment it receives. Within the whole coalition, the off-chain payments sum up to 0. Note that during this

negotiation, the miners may learn information about the users' true values. Therefore, we need to take this into consideration when defining "miners have a better strategy", given that the miners learn some private information from the users.

Plain Model

Definition 6.1.1 (Plain model ex post c-MUCP)**.** We say that a TFM in the plain model with an honest inclusion rule I satisfies ex post c-MUCP iff for any coalition $\mathcal{C}$ consisting of the miner $\mathcal{M}$ and a set $\mathcal{U}$ of no more than c number of users, for any true value vector $\mathbf{v}_\mathcal{U}$ of users in $\mathcal{U}$, for any bid vector $\mathbf{b}_{-\mathcal{C}}$ posted by users not in $\mathcal{C}$, for any strategy $(I^*, \mathbf{b}^*_\mathcal{C})$ of the coalition,

- Either there exists a miner's strategy $(I', \mathbf{b}'_\mathcal{M})$ such that for any strategic bid $\mathbf{b}_\mathcal{U}$ from the users, the miner alone gets higher utility compared to colluding with the users:

$$\mathsf{util}^\mathcal{M}(I', (\mathbf{b}_\mathcal{U}, \mathbf{b}_{-\mathcal{C}}, \mathbf{b}'_\mathcal{M})) \geq \mathsf{util}^\mathcal{M}(I^*, (\mathbf{b}_{-\mathcal{C}}, \mathbf{b}^*_\mathcal{C})),$$

  where $\mathsf{util}^\mathcal{M}(I, \mathbf{b})$ represents the miner's utility when it adopts inclusion rule I on bid vector $\mathbf{b}$. Here,$(I', \mathbf{b}'_\mathcal{M})$ is called a *defecting strategy* of $\mathcal{M}$.
- Or there exists a strategic vector $\mathbf{b}'_\mathcal{U}$ for the set of users $\mathcal{U}$, with which the user coalition gets higher utility compared to colluding with the miner:

$$\mathsf{util}^\mathcal{U}(I, (\mathbf{b}_{-\mathcal{C}}, \mathbf{b}'_\mathcal{U})) \geq \mathsf{util}^\mathcal{U}(I^*, (\mathbf{b}_{-\mathcal{C}}, \mathbf{b}^*_\mathcal{C})),$$

  where $\mathsf{util}^\mathcal{U}(I, \mathbf{b})$ represents $\mathcal{U}$'s joint utility when the input bid vector is $\mathbf{b}$ and the miner adopts inclusion rule I. Here, $\mathbf{b}'_\mathcal{U}$ is called *defecting strategy* of $\mathcal{U}$.

MPC-Assisted Model

We will first define MUCP in the idealized MPC-assisted model, i.e., $\mathcal{F}_{\text{MPC}}$-hybrid model, where we assume that a trusted party implements $\mathcal{F}_{\text{MPC}}$ (see Figure 2.2). Later in [Ke: fill], we will explain how to extend the definitions in real-world instantiations to guarantee computationally sound reasoning.

Definition 6.1.2 (MPC-assisted model ex post (ρ, c)-MUCP)**.** We say that a TFM in the mpc-assisted model satisfies ex post (ρ, c)-MUCP iff for any coalition $\mathcal{C} = \mathcal{M} \cup \mathcal{U}$, where $\mathcal{M}$ consists of no more than ρ fraction of the miners and $\mathcal{C}$ consists no more than c number of users, for any true value vector $\mathbf{v}_\mathcal{U}$ of users in $\mathcal{U}$, for any bid vector $\mathbf{b}_{-\mathcal{C}}$ posted by users not in $\mathcal{C}$, for any strategic bid vector $\mathbf{b}^*_\mathcal{C}$ of the coalition,

- Either there exists a miner's strategy $\mathbf{b}'_\mathcal{M}$ such that for any strategic bid $\mathbf{b}_\mathcal{U}$ from the users, the miner alone gets higher utility compared to colluding with the users:

$$\mathsf{util}^\mathcal{M}(\mathbf{b}_\mathcal{U}, \mathbf{b}_{-\mathcal{C}}, \mathbf{b}'_\mathcal{M}) \geq \mathsf{util}^\mathcal{M}(\mathbf{b}_{-\mathcal{C}}, \mathbf{b}^*_\mathcal{C}),$$

  where $\mathsf{util}^\mathcal{M}(\mathbf{b})$ represents the miner coalition $\mathcal{M}$'s utility when the input bid vector is $\mathbf{b}$. Here, $\mathbf{b}'_\mathcal{M}$ is called *defecting strategy* of $\mathcal{M}$.

- Or there exists a strategic vector $\mathbf{b}'_{\mathcal{U}}$ for the set of users $\mathcal{U}$, with which the user coalition gets higher utility compared to colluding with the miner:

$$\mathsf{util}^{\mathcal{U}}(\mathbf{b}_{-\mathcal{C}}, \mathbf{b}'_{\mathcal{U}}) \geq \mathsf{util}^{\mathcal{U}}(\mathbf{b}_{-\mathcal{C}}, \mathbf{b}^*_{\mathcal{C}}),$$

  where $\mathsf{util}^{\mathcal{U}}(\mathbf{b})$ represents $\mathcal{U}$'s joint utility when the input bid vector is $\mathbf{b}$. Here, $\mathbf{b}'_{\mathcal{U}}$ is called *defecting strategy* of $\mathcal{U}$.

Bayesian MUCP

If, in addition, we assume that the honest users' true values are sampled from some distribution $\mathcal{D}$, then we can also define Bayesian MUCP. Again, note that in the plain model, the coalition can decide their strategy after observing honest users' bids, it does not make sense to consider the Bayesian notion of incentive compatibilities in the plain model. Therefore, we only define Bayesian MUCP in the MPC-assisted model

Definition 6.1.3 (MPC-assisted model Bayesian (ρ, c)-MUCP)**.** Let $\mathcal{D}$ denote the joint distribution of honest users' true values. We say that a TFM in the mpc-assisted model satisfies Bayesian (ρ, c)-MUCP iff for any coalition $\mathcal{C} = \mathcal{M} \cup \mathcal{U}$, where $\mathcal{M}$ consists of no more than ρ fraction of the miners and $\mathcal{C}$ consists no more than c number of users, for any true value vector $\mathbf{v}_{\mathcal{U}}$ of users in $\mathcal{U}$, for any strategic bid vector $\mathbf{b}^*_{\mathcal{C}}$ of the coalition,

- Either there exists a miner's strategy $\mathbf{b}'_{\mathcal{M}}$ such that for any strategic bid $\mathbf{b}_{\mathcal{U}}$ from the users, the miner alone gets higher utility compared to colluding with the users:

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

  where $\mathsf{util}^{\mathcal{M}}(\mathbf{b})$ represents the miner coalition $\mathcal{M}$'s utility when the input bid vector is $\mathbf{b}$. Here, $\mathbf{b}'_{\mathcal{M}}$ is called *defecting strategy* of $\mathcal{M}$.
- Or there exists a strategic vector $\mathbf{b}'_{\mathcal{U}}$ for the set of users $\mathcal{U}$, with which the user coalition gets higher utility compared to colluding with the miner:

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

  where $\mathsf{util}^{\mathcal{U}}(\mathbf{b})$ represents $\mathcal{U}$'s joint utility when the input bid vector is $\mathbf{b}$. Here, $\mathbf{b}'_{\mathcal{U}}$ is called *defecting strategy* of $\mathcal{U}$.

As discussed in Section 1.1.4, when $\mathcal{U}$ has a defecting strategy, we assume that the miners behave honestly because we always consider MUCP in conjunction with c-dominant-strategy MIC. This new notion assumes that even if the miner sees the bidding strategy of the users in $\mathcal{U}$, being honest is still the best strategy. Since c-dominant-strategy MIC is equivalent to MIC in the ex post setting, we only redefine it in the Bayesian setting in the MPC-assisted model. Formally,

Definition 6.1.4 (c-dominant-strategy ρ-MIC)**.** Given $\rho \in (0, 1]$ and an integer $c \geq 1$. We say that a TFM in MPC-assisted model satisfies c-dominant-strategy Bayesian ρ-MIC iff for any non-empty coalition $\mathcal{M}$ of at most ρ fraction of the miners, for any set $\mathcal{U}$ of c number of users, for any true value vector $\mathbf{v}_{\mathcal{U}}$ of users in $\mathcal{U}$, for any miner's strategy $\mathbf{b}_{\mathcal{M}}$, it must be that

$$\mathop{\mathbb{E}}_{\mathbf{b}_{-\mathcal{U}} \sim \mathcal{D}_{-\mathcal{U}}} \left[\mathsf{util}^{\mathcal{M}}(\mathbf{v}_{\mathcal{U}}, \mathbf{b}_{-\mathcal{U}}, \mathbf{b}_{\mathcal{M}})\right] \leq \mathop{\mathbb{E}}_{\mathbf{b}_{-\mathcal{U}} \sim \mathcal{D}_{-\mathcal{U}}} \left[\mathsf{util}^{\mathcal{M}}(\mathbf{v}_{\mathcal{U}}, \mathbf{b}_{-\mathcal{U}})\right].$$

[Ke: does mucp guarantee that the coalition is not stealing money from honest users? can we prove that when colluding users perform its best strategy, honest user's utility is at least their honest utility?]

6.2 Impossibility Results under MUCP

6.2.1 Relationship Between SCP and MUCP

Ex Post UIC + MIC + 1-MUCP Implies Ex Post 1-SCP

It turns out that when $c = 1$, i.e., when there is only one user in the coalition, ex post MUCP and ex post SCP are equivalent assuming the mechanism satisfies ex post UIC and ex post MIC. Formally,

Theorem 6.2.1. *If a TFM satisfies ex post UIC, ex post MIC, and ex post* 1*-MUCP, then it satisfies ex post* 1*-SCP. This holds for both the plain model and the MPC-assisted model.*

Proof. We prove the statement for TFMs in the plain model since the same proof holds for MPC-assisted model TFMs. Let I denote the honest inclusion rule of the TFM. For the sake of contradiction, suppose that the mechanism does not satisfy 1-SCP. Then there exists a coalition $\mathcal{C}$ consisting of the miner $\mathcal{M}$ and a user i with true value v_i, an honest uses' bid vector $\mathbf{b}_{-i}$, and a strategy $(I^*, \mathbf{b}^*_{\mathcal{C}})$ such that

$$\mathsf{util}^{\mathcal{C}}(I, (\mathbf{b}_{-i}, v_i)) < \mathsf{util}^{\mathcal{C}}(I^*, (\mathbf{b}_{-i}, \mathbf{b}^*_{\mathcal{C}})).$$

Since coalition $\mathcal{C}$'s strategy $(I^*, \mathbf{b}^*_{\mathcal{C}})$ improves their joint utility, there must exist a way of joint utility redistribution such that the following holds:

$$\mathsf{util}^{\mathcal{M}}(I, (\mathbf{b}_{-i}, v_i)) < \mathsf{util}^{\mathcal{M}}(I^*, (\mathbf{b}_{-i}, \mathbf{b}^*_{\mathcal{C}})) \text{ and } \mathsf{util}^{i}(I, (\mathbf{b}_{-i}, v_i)) < \mathsf{util}^{i}(I^*, (\mathbf{b}_{-i}, \mathbf{b}^*_{\mathcal{C}})).$$

Since the mechanism satisfies 1-MUCP, it must be that either there exists a miner's strategy $(I', \mathbf{b}'_{\mathcal{M}})$ such that

$$\mathsf{util}^{\mathcal{M}}(I, (\mathbf{b}_{-i}, v_i)) < \mathsf{util}^{\mathcal{M}}(I^*, (\mathbf{b}_{-i}, \mathbf{b}^*_{\mathcal{C}})) \leq \mathsf{util}^{\mathcal{M}}(I', (\mathbf{b}_{-i}, v_i, \mathbf{b}'_{\mathcal{M}})); \tag{6.1}$$

or there exists a strategic bid vector $\mathbf{b}'_i$ for user i such that

$$\mathsf{util}^{i}(I, (\mathbf{b}_{-i}, v_i)) < \mathsf{util}^{i}(I^*, (\mathbf{b}_{-i}, \mathbf{b}^*_{\mathcal{C}})) \leq \mathsf{util}^{i}(I, (\mathbf{b}_{-i}, \mathbf{b}'_i)). \tag{6.2}$$

However, (6.1) contradicts MIC and (6.2) contradicts UIC. Thus, it must be the case that the mechanism also satisfies 1-SCP. □

However, this result does not directly hold for the Bayesian incentive compatibility in the MPC-assisted model. To see this, note that had we consider a Bayesian notion of incentive compatibility, we will get the following "Bayesian" version of (6.1) [1]:

$$\mathop{\mathbb{E}}_{\mathbf{b}_{-i} \sim \mathcal{D}_{-i}} \mathsf{util}^{\mathcal{M}}(\mathbf{b}_{-i}, v_i) < \mathop{\mathbb{E}}_{\mathbf{b}_{-i} \sim \mathcal{D}_{-i}} \mathsf{util}^{\mathcal{M}}(\mathbf{b}_{-i}, \mathbf{b}^*_{\mathcal{C}}) \leq \mathop{\mathbb{E}}_{\mathbf{b}_{-i} \sim \mathcal{D}_{-i}} \mathsf{util}^{\mathcal{M}}(\mathbf{b}_{-i}, v_i, \mathbf{b}'_{\mathcal{M}}), \tag{6.3}$$

[1] we ignore the inclusion rule since Bayesian notion of incentive compatibility only makes sense in the MPC-assisted model, which guarantees the honest implementation of inclusion rule

where $\mathcal{D}_{-i}$ denotes the joint distribution of honest users' true values. This does NOT directly contradict Bayesian MIC. Instead, we can get a similar conclusion either with a stronger notion of MIC or with a weaker notion of SCP.

Bayesian UIC + 1-dominant-strategy MIC + Bayesian 1-MUCP Implies Bayesian 1-SCP

Theorem 6.2.2. *For any $\rho \in (0, 1]$, if a TFM in the MPC-assisted model satisfies Bayesian UIC, 1-dominant-strategy ρ-MIC, and Bayesian $(\rho, 1)$-MUCP, then it satisfies Bayesian $(\rho, 1)$-SCP.*

Proof. By the same proof as with Theorem 6.2.1. If a TFM is Bayesian 1-MUCP but not 1-SCP, then it either contradicts 1-dominant MIC (see (6.3)), or it contradicts UIC (see Bayesian version of (6.2)). □

6.2.2 Bounds on Miner Revenue

Combining the previous bounds on miner revenues under SCP and Theorem 6.2.1, we get the following results.

Theorem 6.2.3. *We have the following impossibility results for TFMs under MUCP:*

1. *If a TFM in the plain model satisfies ex post UIC, ex post MIC, and ex post 1-MUCP, then the miner revenue must be 0. In addition, if the block size is finite, there is no non-trivial TFM that satisfies three properties simultaneously.*
2. *For any $\rho \in (0, 1]$, if a TFM in the MPC-assisted model satisfies Bayesian UIC, 1-dominant strategy ρ-MIC, and $(\rho, 1)$-MUCP, then the miner revenue must be 0.*
3. *Suppose honest user i's true values is sampled independently from $\mathcal{D}_i$ for $i = 1, \ldots, n$. For any $\rho \in (0, 1]$, $h \geq 1$ and any $d \geq c \geq 1$, if an MPC-assisted TFM in the (h, ρ, c, d)-environment satisfies Bayesian UIC, 1-dominant MIC, and MUCP, then the miner revenue is at most $\sum_{i=1}^{h} \mathbb{E}_{X_i \sim \mathcal{D}_i}[X_i]$.*

Proof. Results 1 can be obtained by combining Theorem 6.2.1 and results in [CS23]. Results 2 can be obtained by combining Theorem 6.2.1 and Theorem 3.2.4. Results 3 can be obtained by combining Theorem 6.2.1 and Theorem 5.5.1. □

[Ke: how about approx? add approx if time allows]

6.3 Feasibility under MUCP

With the relaxation of MUCP, we now revisit the mechanisms for finite block sizes and prove that these mechanisms satisfy MUCP.

6.3.1 Burning Posted Price with Random Selection

For easier reading, we restate the burning posted price auction with random selection below.

MPC-assisted, burning 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.

Theorem 6.3.1. *The MPC-assisted posted price with random selection satisfies ex post UIC, ex post ρ-MIC, and ex post (ρ, c)-MUCP for arbitrary $\rho \in (0, 1]$, arbitrary $c \geq 1$.*

Proof. UIC and MIC of the mechanism have been proved in Theorem 3.3.1. In this proof, we focus on proving (ρ, c)-MUCP. For any coalition $\mathcal{C}$ consisting of a non-empty set $\mathcal{U}$ of no more than c users and a non-empty set $\mathcal{M}$ of no more than ρ fraction of miners $\mathcal{M}$, let $\mathbf{v}$ denote the vector of true values of $\mathcal{U}$.

For any strategy $S_{\mathcal{C}}(\mathbf{v})$ of $\mathcal{C}$, consider the following user-only strategy $S_{\mathcal{U}}(\mathbf{v})$ for $\mathcal{U}$: Upon observing honest users' bid vector $\mathbf{b}_{-\mathcal{C}}$, $S_{\mathcal{U}}(\mathbf{v})$ runs $S_{\mathcal{C}}(\mathbf{v})$ on $\mathbf{b}_H$ and gets $\mathbf{b}^*$. The user coalition $\mathcal{U}$ then bids $\mathbf{b}^*$.

Now we show that for any $\mathbf{b}_{-\mathcal{C}}$, either $S_{\mathcal{U}}(\mathbf{v})$ is a defecting strategy of $\mathcal{U}$ w.r.t. strategy $S_{\mathcal{C}}$, or the honest strategy $HS_{\mathcal{M}}$ is a sensible defecting strategy of $\mathcal{M}$ w.r.t. strategy $S_{\mathcal{C}}$. To see why this is the case, note that the miner revenue is always 0. Therefore, by the construction of $S_{\mathcal{U}}(\mathbf{v})$

$$\mathsf{util}^{\mathcal{C}}(S_{\mathcal{C}}(\mathbf{v}), \mathbf{b}_{-\mathcal{C}}) = \mathsf{util}^{\mathcal{U}}(S_{\mathcal{U}}(\mathbf{v}), \mathbf{b}_{-\mathcal{C}}). \tag{6.4}$$

There are two possible cases:

- If $\mathsf{util}^{\mathcal{U}}(S_{\mathcal{C}}(\mathbf{v}), \mathbf{b}_{-\mathcal{C}}) \leq \mathsf{util}^{\mathcal{U}}(S_{\mathcal{U}}(\mathbf{v}), \mathbf{b}_{-\mathcal{C}})$, then $S_{\mathcal{U}}(\mathbf{v})$ is a defecting strategy of $\mathcal{U}$ by definition;
- If $\mathsf{util}^{\mathcal{U}}(S_{\mathcal{C}}(\mathbf{v}), \mathbf{b}_{-\mathcal{C}}) > \mathsf{util}^{\mathcal{U}}(S_{\mathcal{U}}(\mathbf{v}), \mathbf{b}_{-\mathcal{C}})$, then by Equation (6.4), it must be that

  $$\mathsf{util}^{\mathcal{M}}(S_{\mathcal{C}}(\mathbf{v}), \mathbf{b}_{-\mathcal{C}}) = \mathsf{util}^{\mathcal{C}}(S_{\mathcal{C}}(\mathbf{v}), \mathbf{b}_{-\mathcal{C}}) - \mathsf{util}^{\mathcal{U}}(S_{\mathcal{C}}(\mathbf{v}), \mathbf{b}_{-\mathcal{C}}) < 0.$$

  Note that when $\mathcal{M}$ acts honestly on its own, its utility is 0. Therefore, the honest strategy HS is a defecting strategy of $\mathcal{M}$.

Therefore, the mechanism is (ρ, c)-MUCP by Definition 6.1.2. □

6.3.2 LP-Based Mechanism with Random Selection

For easier reading, we restate the LP-based mechanism with random selection below. Here, we make one modification of the mechanism, which is emphasized with a green background. Still, here we assume that each honest user i's true value is sampled independently from some distribution $\mathcal{D}_i$, and that each $\mathcal{D}_i$ has the same median m such that $\Pr_{x \sim \mathcal{D}_i}[x \geq m] = \frac{1}{2}$ for any i. Meanwhile, strategic user's true value can be any arbitrary value.

MPC-assisted, LP-based mechanism with random selection

Parameters: the block size k, the environment $(h, *, 1, d)$, the distribution median m.

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

Mechanism:

- *Allocation Rule.* Let $d' = c + d$. Let $\widetilde{\mathbf{b}} = (\widetilde{b}_1, \ldots, \widetilde{b}_s)$ denote the bids that are at least m. If $s \leq k$, confirm all bids in $\widetilde{\mathbf{b}}$. Otherwise, randomly select k bids from $\widetilde{\mathbf{b}}$ to confirm.
- *Payment rule.* Each confirmed bid pays m.
- *Miner revenue rule.*

  Let $\mathbf{y} := (y_0, y_1, \ldots, y_n)$ be any feasible solution to the following linear program:

$$\forall i \in [n] : \ 0 \leq y_i \leq min(i, k) \cdot m \tag{6.5}$$

$$\forall 0 \leq j \leq d : \quad \sum_{i=0}^{n-d'} q_i \cdot y_{i+j} = \frac{m \cdot min(h, k)}{4} \tag{6.6}$$

  where $q_i = \frac{1}{2^{n-d'}} \binom{n-d'}{i}$ is the probability of observing i heads if we flip $n-d'$ independent fair coins. The total miner revenue is y_t where t is the number of confirmed bids in the block.

Theorem 6.3.2. *Suppose the block size is k. Fix any $h \geq 2$, and any $d \leq \frac{1}{16}\sqrt{\frac{h}{2 \log h}}$. The LP-based mechanism with random selection satisfies ex post UIC, c-dominant strategy Bayesian MIC, and Bayesian MUCP in an $(h, *, c, d)$-environment.*

Proof. First, UIC has been proved in Theorem 5.4.1. For c-dominant strategy MIC, note that we replace d in the original LP-based mechanism with $d' = c + d$. This guarantees that as long as there are $n - d'$ honest users' bids that are kept private from the miners, the expected miner revenue does not change (Theorem 5.4.1). Therefore, following a similar reasoning as in Theorem 5.4.1, even if the miner sees c bids from c users and controls at most d bids, its expected miner revenue does not change. c-dominant strategy MIC thus follows.

In the rest of this proof, we focus on proving that the mechanism satisfies MUCP. Fix an arbitrary coalition $\mathcal{C}$ consisting of a non-empty set $\mathcal{U}$ of no more than c users and a non-empty set $\mathcal{M}$ of no more than ρ fraction of miners $\mathcal{M}$. Let $\mathbf{v}$ denote the vector of true values of $\mathcal{U}$. For any strategy $S_{\mathcal{C}}(\mathbf{v})$ of $\mathcal{C}$, define the following user-only strategy $S_{\mathcal{U}}(\mathbf{v})$ for $\mathcal{U}$: It simply runs $S_{\mathcal{C}}(\mathbf{v})$ and gets $\mathbf{b}^*$. The user coalition then bids $\mathbf{b}^*$.

Next, we show that either $S_{\mathcal{U}}(\mathbf{v})$ is a defecting strategy of $\mathcal{U}$ w.r.t. coalition strategy $S_{\mathcal{C}}$, or the honest strategy $HS_{\mathcal{M}}$ is a defecting strategy of $\mathcal{M}$ w.r.t. $S_{\mathcal{C}}$. To see why this is the case, note that the linear program guarantees that no matter how the coalition bids, as long as it meets the requirement of the environment, the expected miner revenue is always the same: $\mathop{\mathbb{E}}_{\mathbf{b}_{-\mathcal{C}} \sim \mathcal{D}_{-\mathcal{C}}} [\mu(\mathbf{b}_{-\mathcal{C}}, \mathbf{b}^*)] = \frac{m \cdot \min(h,k)}{4}$. Therefore, by the construction of $S_{\mathcal{U}}(\mathbf{v})$,

$$\mathop{\mathbb{E}}_{\mathbf{b}_{-\mathcal{C}} \sim \mathcal{D}_{-\mathcal{C}}} \mathsf{util}^{\mathcal{C}}(S_{\mathcal{C}}(\mathbf{v}), \mathbf{b}_{-\mathcal{C}}) = \mathop{\mathbb{E}}_{\mathbf{b}_{-\mathcal{C}} \sim \mathcal{D}_{-\mathcal{C}}} \mathsf{util}^{\mathcal{U}}(S_{\mathcal{U}}(\mathbf{v}), \mathbf{b}_{-\mathcal{C}}) + \frac{m \cdot \min(h,k)}{4}. \tag{6.7}$$

There are two possible cases.

- If $\mathop{\mathbb{E}}_{\mathbf{b}_{-\mathcal{C}} \sim \mathcal{D}_{-\mathcal{C}}} \mathsf{util}^{\mathcal{U}}(S_{\mathcal{C}}(\mathbf{v}), \mathbf{b}_{-\mathcal{C}}) \leq \mathop{\mathbb{E}}_{\mathbf{b}_{-\mathcal{C}} \sim \mathcal{D}_{-\mathcal{C}}} \mathsf{util}^{\mathcal{U}}(S_{\mathcal{U}}(\mathbf{v}), \mathbf{b}_{-\mathcal{C}})$, then $S_{\mathcal{U}}(\mathbf{v})$ is a defecting strategy of $\mathcal{U}$ by definition;
- If $\mathop{\mathbb{E}}_{\mathbf{b}_{-\mathcal{C}} \sim \mathcal{D}_{-\mathcal{C}}} \mathsf{util}^{\mathcal{U}}(S_{\mathcal{C}}(\mathbf{v}), \mathbf{b}_{-\mathcal{C}}) > \mathop{\mathbb{E}}_{\mathbf{b}_{-\mathcal{C}} \sim \mathcal{D}_{-\mathcal{C}}} \mathsf{util}^{\mathcal{U}}(S_{\mathcal{U}}(\mathbf{v}), \mathbf{b}_{-\mathcal{C}})$, then by Equation (6.4), it must be that $\mathop{\mathbb{E}}_{\mathbf{b}_{-\mathcal{C}} \sim \mathcal{D}_{-\mathcal{C}}} \mathsf{util}^{\mathcal{M}}(S_{\mathcal{C}}(\mathbf{v}), \mathbf{b}_{-\mathcal{C}}) < \frac{m \cdot \min(h,k)}{4}$. Note that when $\mathcal{M}$ acts honestly on its own, its utility is $\frac{m \cdot \min(h,k)}{4}$. Therefore, the honest strategy HS is a defecting strategy of $\mathcal{M}$.

Therefore, the mechanism is (ρ, c)-MUCP by Definition 6.1.2. □

Chapter 7

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

So far in the paper, we have assumed that the MPC-assisted transaction fee mechanisms are 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, since this is a good starting point as an initial exploration. All the impossibility results in our paper hold even with generic MPC. However, as we will discuss in Section 7.5, some MPC-assisted mechanisms described in this thesis actually need not employ generic MPC to be instantiated in practice — we describe efficient instantiations for our specific protocols in Section 7.5. In Section 7.4, we will discuss how to modify the instantiation such that it can tolerate a majority of corrupted miners. In Section 7.6, we will formally define computational incentive compatibility and show that all our MPC-assisted TFMs, when implemented in a real-world MPC-assisted model, satisfy computational incentive compatibility.

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. Each player is modeled as probabilistic poltnomial-time (p.p.t.) algorithms. Specifically, the execution of the mechanism is parametrized with some security parameter λ. A p.p.t. algorithm (or *efficient* algorithms henceforth) runs in time upper bounded by some polynomial function in λ. 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.

A function $f(\cdot)$ is *negligible* iff for any polynomial function $\mathsf{poly}(\cdot)$, $f(x) \leq 1/\mathsf{poly}(x)$. We say that two distribution ensembles $\{X_\lambda\}_\lambda$ and $\{Y_\lambda\}_\lambda$ are *computationally indistinguishable*, denoted as $\{X_\lambda\}_\lambda \equiv_c \{Y_\lambda\}_\lambda$, iff for all non-uniform p.p.t. algorithm $\mathcal{A}$, there exists a negligible function $\mathsf{negl}(\cdot)$ such that for any $\lambda \in \mathbb{N}$, $|\Pr[x \leftarrow X_\lambda, \mathcal{A}(x) = 1] - \Pr[y \leftarrow Y_\lambda, \mathcal{A}(y) = 1]| \leq \mathsf{negl}(\lambda)$.

7.1 Building Blocks

We first introduce some building blocks used in the protocol.

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.

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 $|I| < t$,

$$\{\{s_i\}_{i \in I} : (s_1, \ldots, s_m) \leftarrow \mathsf{share}(s)\} \equiv \{\{u_i\}_{i \in I} : u_i \text{ uniformly randomly chosen from } \mathbb{F}\}.$$

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 $\mathcal{A}$from 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

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

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:

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

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

1. Upon receiving crs generation query $\mathtt{gen}$ 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 $(\mathtt{prove}, \mathsf{stmt})$ 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 7.1.1 (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.*

7.2 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 Section 7.1. Moreover, we will describe our protocol Π_{MPC} assuming that players have access to an ideal functionality $\mathcal{F}_{\mathrm{TFM}}$ which computes the rules of the auction — the formal description of $\mathcal{F}_{\mathrm{TFM}}$ will be provided at the end of Π_{MPC}. The ideal functionality $\mathcal{F}_{\mathrm{TFM}}$ 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 $\mathcal{F}_{\mathrm{MPC}}$

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 $\mathcal{F}_{\mathrm{TFM}}$.
2. For each $j \in [m]$, the functionality $\mathcal{F}_{\mathrm{TFM}}$ 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 7.2.1. *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 $\mathcal{F}_{\mathrm{MPC}}$ (See Figure 2.2) in the $\mathcal{F}_{\mathrm{TFM}}$-hybrid model as long as the number of colluding miners is less than $\frac{m}{2}$.*

7.3 Proof of Theorem 7.2.1

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 $\mathcal{F}_{\mathrm{MPC}}$ 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 $\mathcal{H}_{\mathrm{miner}}$ and $\mathcal{K}_{\mathrm{miner}}$ to denote the set of honest miners and corrupted miners, respectively. Formally, the simulator $\mathcal{S}$ interacting with $\mathcal{F}_{\mathrm{MPC}}$ behaves as follows.

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

Sharing Phase

1. Let $\mathcal{C}$ be an empty set.
2. Emulate honest miner $h \in \mathcal{H}_{\mathrm{miner}}$ 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 $\{\tau_h\}_{h\in\mathcal{H}_{\mathrm{miner}}}$, and $\overrightarrow{\mathsf{ek}}$ be the vector of $\{\mathsf{ek}_h\}_{h\in\mathcal{H}_{\mathrm{miner}}}$.
3. For each corrupted miner $k \in \mathcal{K}_{\mathrm{miner}}$, 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 $k \in \mathcal{K}_{\mathrm{miner}}$, let the share $X_{i,k}$ be a uniformly random element in the finite field $\mathbb{F}$. For every honest miner $h \in \mathcal{H}_{\mathrm{miner}}$, 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 $\pi_i \leftarrow \mathsf{NIZK}.\widetilde{\mathsf{P}}(\mathsf{CRS}, \overrightarrow{\tau}, \mathsf{stmt}_i)$, 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 $\mathsf{NIZK}.\mathsf{V}(\mathsf{CRS}, \mathsf{stmt}_\ell, \pi_\ell) = 0$ 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.

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 7.3.1. *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 Section 7.1), 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 7.3.2. *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 7.3.3. *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 7.3.4. *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 $\mathcal{F}_{\mathrm{TFM}}$. 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 $\mathcal{F}_{\mathrm{MPC}}$ in the $\mathcal{F}_{\mathrm{TFM}}$-hybrid model as long as the adversary controls only a minority number of miners.

7.4 MPC Protocol in the Presence of Majority-Miner Coalitions

So far, we have focused on instantiating the MPC protocol when the coalition controls only a minority of the miners. But 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, $\mathcal{F}_{\mathrm{TFM}}$ should also be instantiated with a corrupt majority MPC protocol.

7.5 Efficient Instantiations of MPC-Assisted Mechanisms

Certain MPC-assisted mechanisms proposed in this paper, including [Ke: list them] achieve incentive compatibility in the ex-post setting. This allows us to instantiate these mechanisms without using the above generic MPC, which can be expensive to implement in practice. Instead, 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.

7.6 Computational Incentive Compatibility

For computational sound reasoning, we additionally make the following assumption for the true value distributions: we assume that $\mathcal{D}(1^\lambda)$ outputs an n that is upper bounded by some fixed polynomial function in λ; and moreover, every player's type can be encoded as some string whose length is upper bounded by some fixed polynomial function in λ.

Definition 7.6.1 (Utility-equivalent emulation)**.** We say that a mechanism Π is a utility-equivalent emulation of another mechanism Π' w.r.t. a coalition $\mathcal{C}$ and the distribution $\mathcal{D}$, iff for any vector $\mathbf{v}_\mathcal{C}$ denoting the true values of the users in $\mathcal{C}$, for any non-uniform p.p.t. strategy $S_\mathcal{C}(\mathbf{v}_\mathcal{C})$ in the mechanism Π, there exists a non-uniform p.p.t. strategy $S'_\mathcal{C}(\mathbf{v}_\mathcal{C})$, such that

1. The following probability ensembles are computationally indistinguishable:
   - *Real-world experiment.* Call $\mathbf{v}_{-\mathcal{C}} \sim \mathcal{D}(1^\lambda)$ to sample the number of honest players and their true values, execute the real-world mechanism Π with a set of honest players with true values $\mathbf{v}_{-\mathcal{C}}$, and a coalition $\mathcal{C}$ with true values $\mathbf{v}_\mathcal{C}$ and executing the strategy $S_\mathcal{C}(\mathbf{v}_\mathcal{C})$; and output the utilities of every member of $\mathcal{C}$ in the execution.
   - *Ideal-world experiment.* Call $\mathbf{v}_{-\mathcal{C}} \sim \mathcal{D}(1^\lambda)$ to sample the number of honest users and their true values, execute the mechanism Π' with honest users whose true values are denoted $\mathbf{v}_{-\mathcal{C}}$, and a coalition $\mathcal{C}$ which has the types $\mathbf{v}_\mathcal{C}$ executing the strategy $S'_\mathcal{C}(\mathbf{v}_\mathcal{C})$; finally, output the utilities of every member of $\mathcal{C}$ in the mechanism Π'.
2. Furthermore, the above also holds if $S_\mathcal{C}(\mathbf{v}_\mathcal{C})$ is the honest strategy of $\mathcal{C}$ in mechanism Π and $S'_\mathcal{C}(\mathbf{v}_\mathcal{C})$ is the honest strategy of $\mathcal{C}$ in Π'.

Henceforth, we also refer to $S'_\mathcal{C}(\mathbf{v}_\mathcal{C})$ a *utility-equivalent strategy* of $S_\mathcal{C}(\mathbf{v}_\mathcal{C})$ in the mechanism Π.

Definition 7.6.2 (Computational incentive compatibility)**.** We say that an MPC-assisted transaction fee mechanism Π is *computationally* UIC (resp. ρ-MIC, (ρ, c)-SCP) in an environment $\mathcal{E}$ iff the following hold: there exists an ideal mechanism Π' such that Π is a utility-equivalent emulation of Π' w.r.t. $\mathcal{C}$ and $\mathcal{D}$; and moreover, Π' is UIC (resp. ρ-MIC, (ρ, c)-SCP) in an environment $\mathcal{E}$.

Definition 7.6.3 (Computational MUCP)**.** We say that an MPC-assisted transaction fee mechanism Π satisfies computational MUCP iff for any miner-user coalition $\mathcal{C}$ consisting of a set of miners $\mathcal{M}$ and a set of users $\mathcal{U}$ with any true value vector $\mathbf{v}_\mathcal{U}$, any coalition's strategy $S_\mathcal{C}(\mathbf{v}_\mathcal{U})$, there exists an ideal mechanism $\widetilde{\Pi}$ such that

- Π is a utility-equivalent emulation of $\widetilde{\Pi}$ w.r.t. $\mathcal{C}$ and $\mathcal{D}$;

- Either there exists a p.p.t. strategy $S^*_{\mathcal{M}}$ for the miners $\mathcal{M}$, and a utility equivalent strategy $\widetilde{S}^*_{\mathcal{M}}$ in $\widetilde{\Pi}$ such that $\widetilde{S}^*_{\mathcal{M}}$ is a defecting strategy for $\mathcal{M}$ w.r.t. $\widetilde{S}_{\mathcal{C}}(\mathbf{v}_{\mathcal{U}})$, where $\widetilde{S}_{\mathcal{C}}(\mathbf{v}_{\mathcal{U}})$ is a utility equivalent strategy of $S_{\mathcal{C}}(\mathbf{v}_{\mathcal{U}})$ in $\widetilde{\Pi}$;
  Or there exists a p.p.t. strategy $S^*_{\mathcal{U}}(\mathbf{v}_{\mathcal{U}})$ for the users $\mathcal{U}$, and a utility equivalent strategy $\widetilde{S}^*_{\mathcal{U}}(\mathbf{v}_{\mathcal{U}})$ in $\widetilde{\Pi}$ such that $S^*_{\mathcal{U}}(\mathbf{v}_{\mathcal{U}})$ is a defecting strategy for $\mathcal{U}$ w.r.t. $\widetilde{S}_{\mathcal{C}}(\mathbf{v}_{\mathcal{U}})$, where $\widetilde{S}_{\mathcal{C}}(\mathbf{v}_{\mathcal{U}})$ is a utility equivalent strategy of $S_{\mathcal{C}}(\mathbf{v}_{\mathcal{U}})$ in $\widetilde{\Pi}$;

If a real-world mechanism Π securely realizes a ideal-world mechanism $\widetilde{\Pi}$, then Π is a utility-equivalent emulation of $\widetilde{\Pi}$ w.r.t. any coalition $\mathcal{C}$ that involves less than $\frac{1}{2}M$ fraction of miners by Theorem 7.2.1. Thus, we have the following results:

Theorem 7.6.4. *Given some $\rho \in (0, \frac{1}{2})$ and an integer $c \geq 1$. Let $\widetilde{\Pi}$ be a mechanism that satisfies UIC (resp. ρ-MIC or (ρ, c)-SCP) in the idealized MPC-assisted model. Let Π be a mechanism in the real-world MPC-assisted model that securely instantiates an ideal-world mechanism $\widetilde{\Pi}$. Then Π satisfies computational UIC (resp. computational ρ-MIC or computational (ρ, c)-SCP).*

Proof. By Theorem 7.2.1 and Definition 7.6.1, Π is a utility-equivalent emulation of $\widetilde{\Pi}$ w.r.t. any coalition $\mathcal{C}$ that involves less than half miners. Therefore, the theorem follows by Definition 7.6.2. □

Therefore, all the mechanisms described in this thesis in the idealized MPC-assisted model still satisfy computational incentive compatibility in real-world instantiations. Next, we prove that the two mechanisms, posted price auction with random selection and LP-based mechanism with random selection, also satisfy computational MUCP when instantiated with real-world cryptography.

Theorem 7.6.5 (Computational MUCP of MPC-assisted posted price auction with random selection)**.** *Given some $\rho \in (0, \frac{1}{2})$ and an integer $c \geq 1$. Let Π be a mechanism in the real-world MPC-assisted model that securely instantiates posted-price auction with random selection $\widetilde{\Pi}$ (Figure 3.2). Then Π satisfies computational (ρ, c)-MUCP for arbitrary $c \geq 1$.*

Proof. By Theorem 7.2.1 and Definition 7.6.1, Π is a utility-equivalent emulation of $\widetilde{\Pi}$ w.r.t. any coalition $\mathcal{C}$ that involves less than half miners. Consider an arbitrary miner-user coalition $\mathcal{C}$ that consists of no more than half fraction of the miners $\mathcal{M}$ and a set of no more than c users $\mathcal{U}$ with true value $\mathbf{v}_{\mathcal{U}}$. For any strategy $S_{\mathcal{C}}(\mathbf{v}_{\mathcal{U}})$ of coalition $\mathcal{C}$ in Π, there exists a utility-equivalent $\widetilde{S}_{\mathcal{C}}(\mathbf{v}_{\mathcal{U}})$ in $\widetilde{\Pi}$.

Now consider the following p.p.t. strategies: The miner's strategy $S^*_{\mathcal{M}}$, which is the honest strategy of $\mathcal{M}$ in Π; and $S^*_{\mathcal{U}}(\mathbf{v}_{\mathcal{U}})$ simply runs $S_{\mathcal{C}}(\mathbf{v}_{\mathcal{U}})$. By definition, these two strategies are p.p.t. algorithms. Since by Theorem 6.3.1, either the honest strategy $\widetilde{S}^*_{\mathcal{M}}$ of $\mathcal{M}$ in $\widetilde{\Pi}$ is a defecting strategy w.r.t. $\widetilde{S}_{\mathcal{C}}(\mathbf{v}_{\mathcal{U}})$, or the users' strategy $\widetilde{S}^*_{\mathcal{U}}(\mathbf{v}_{\mathcal{U}})$ which simply runs $\widetilde{S}_{\mathcal{C}}(\mathbf{v}_{\mathcal{U}})$ is a defecting strategy of $\mathcal{U}$. The theorem then follows by observing that $\widetilde{S}^*_{\mathcal{M}}$ is a utility equivalent strategy of $S^*_{\mathcal{M}}$ and $\widetilde{S}^*_{\mathcal{U}}(\mathbf{v}_{\mathcal{U}})$ is a utility equivalent strategy of $S^*_{\mathcal{U}}(\mathbf{v}_{\mathcal{U}})$. □

Theorem 7.6.6. *Suppose the block size is k. Fix any $\rho \in (0, \frac{1}{2})$, $h \geq 2$, and any $d \leq \frac{1}{16}\sqrt{\frac{h}{2\log h}}$. Let Π be a mechanism in the real-world MPC-assisted model that securely instantiates the LP-based mechanism with random selection (Figure 5.4). Then Π satisfies computational MUCP in an (h, ρ, c, d)-environment for arbitrary $c \geq 1$.*

Proof. By the same proof as Theorem 7.6.5. □

References

[ACH11] Gilad Asharov, Ran Canetti, and Carmit Hazay. Towards a game theoretic view of secure computation. In *Eurocrypt*, 2011. 1.1, 1.1.3, 1.1.3, 1.1.5

[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. 1.1, 1.1.3, 1.1.3, 1.1.5

[AEC21] Guillermo Angeris, Alex Evans, and Tarun Chitra. A note on bundle profit maximization, 2021. 1.1.5

[AHV14] Colleen Alkalay-Houlihan and Adrian Vetta. False-name bidding and economic efficiency in combinatorial auctions. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 28, 2014. 1.1.5

[AL11] Gilad Asharov and Yehuda Lindell. Utility dependence in correct and fair rational secret sharing. *Journal of Cryptology*, 24(1), 2011. 1.1, 1.1.3, 1.1.3, 1.1.5

[AL20] Mohammad Akbarpour and Shengwu Li. Credible auctions: A trilemma. *Econometrica, Econometric Society*, 2020. 1.1.5

[AS16] Noga Alon and Joel H Spencer. *The probabilistic method*. John Wiley & Sons, 2016. 5.3.3

[BCD+] Vitalik Buterin, Eric Conner, Rick Dudley, Matthew Slipper, and Ian Norden. Ethereum improvement proposal 1559: Fee market change for eth 1.0 chain. https://github.com/ethereum/EIPs/blob/master/EIPS/eip-1559.md. 1, 1.1.5, 2.2

[BCLL22] Massimo Bartoletti, James Hsin-yu Chiang, and Alberto Lluch Lafuente. Maximizing extractable value from automated market makers. In *Financial Cryptography and Data Security: 26th International Conference, FC 2022, Grenada, May 2–6, 2022, Revised Selected Papers*, 2022. 1.1.5

[BDKJ23] Kushal Babel, Philip Daian, Mahimna Kelkar, and Ari Juels. Clockwork finance: Automated analysis of economic security in smart contracts. In *IEEE Symposium on Security and Privacy*, 2023. 1.1.5

[BEOS19] Soumya Basu, David A. Easley, Maureen O'Hara, and Emin Gün Sirer. Towards a functional fee market for cryptocurrencies. *CoRR*, abs/1901.06830, 2019. 1, 1.1.5, 2.2

[Can00] Ran Canetti. Security and composition of multiparty cryptographic protocols. *Jour-*

nal of CRYPTOLOGY, 13:143–202, 2000. 2.3

[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. 1.1, 1.1.3, 1.1.3, 1.1.5

[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. 1.1, 1.1.3, 1.1.3, 1.1.5

[CM12] Jing Chen and Silvio Micali. Collusive dominant-strategy truthfulness. *J. Econ. Theory*, 147(3):1300–1312, 2012. 1.1.5

[CRS24] Hao Chung, Tim Roughgarden, and Elaine Shi. Collusion-resilience in transaction fee mechanism design. *arXiv preprint arXiv:2402.09321*, 2024. 1, 1.1.4, 1.1.5

[CS23] Hao Chung and Elaine Shi. Foundations of transaction fee mechanism design. In *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 3856–3899. SIAM, 2023. 1, 1.1.1, **??**, **??**, 1.1.1, 1.1.5, 1.1.5, 1.1.5, 3.1, 3.2, 3.2, 3.2.2, 6.2.3

[DM17] Alan Deckelbaum and Silvio Micali. Collusion, efficiency, and dominant strategies. *Games Econ. Behav.*, 103:83–93, 2017. 1.1.5

[DR07] Yevgeniy Dodis and Tal Rabin. Cryptography and game theory. In *AGT*, 2007. 1.1, 1.1.3, 1.1.3, 1.1.5

[EFW22] Meryem Essaidi, Matheus V. X. Ferreira, and S. Matthew Weinberg. Credible, strategyproof, optimal, and bounded expected-round single-item auctions for all distributions. In Mark Braverman, editor, *13th Innovations in Theoretical Computer Science Conference, ITCS 2022, January 31 - February 3, 2022, Berkeley, CA, USA*, volume 215 of *LIPIcs*, pages 66:1–66:19, 2022. 1.1, 1.1.3, 1.1.3, 1.1.5

[EW09] Joseph Engelberg and Jared Williams. ebay's proxy bidding: A license to shill. *Journal of Economic Behavior and Organization*, 72:509–526, 10 2009. 1.1.5

[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. 1, 1.1.5, 2.2

[FW20] Matheus V. X. Ferreira and S. Matthew Weinberg. Credible, truthful, and two-round (optimal) auctions via cryptographic commitments. In Péter Biró, Jason D. Hartline, Michael Ostrovsky, and Ariel D. Procaccia, editors, *EC '20: The 21st ACM Conference on Economics and Computation, Virtual Event, Hungary, July 13-17, 2020*, pages 683–712. ACM, 2020. 1.1, 1.1.3, 1.1.3, 1.1.5

[GH05] Andrew V. Goldberg and Jason D. Hartline. Collusion-resistant mechanisms for single-parameter agents. In *SODA 2005*, pages 620–629, 2005. 1.1.5, 3.4, 3.4, 1

[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. 1.1, 1.1.3, 1.1.3, 1.1.5

[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. 1.1, 1.1.3, 1.1.3, 1.1.5

[GL79] Jerry Green and Jean-Jacques Laffont. On coalition incentive compatibility. *The Review of Economic Studies*, 46(2):243–254, 04 1979. 1.1.5

[GLR10] Ronen Gradwohl, Noam Livne, and Alon Rosen. Sequential rationality in cryptographic protocols. In *FOCS*, 2010. 1.1, 1.1.3, 1.1.3, 1.1.5

[GMR90] Daniel A. Graham, Robert C. Marshall, and Jean-Francois Richard. Phantom bidding against heterogeneous bidders. *Economics Letters*, 32(1), 1990. 1.1.5

[GMW87] O. Goldreich, S. Micali, and A. Wigderson. How to play any mental game. In *ACM symposium on Theory of computing (STOC)*, 1987. 7.2

[GO14] Jens Groth and Rafail Ostrovsky. Cryptography in the multi-string model. *Journal of cryptology*, 27(3):506–543, 2014. 7.1, 7.1, 7.1.1

[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. 7.1

[GTZ15] Juan A. Garay, Björn Tackmann, and Vassilis Zikas. Fair distributed computation of reactive functions. In *DISC*, volume 9363, pages 497–512, 2015. 1.1, 1.1.3, 1.1.3, 1.1.5

[GY22] Yotam Gafni and Aviv Yaish. Greedy transaction fee mechanisms for (non-)myopic miners, 2022. 1, 1.1.5

[Har] Jason Hartline. Lectures on optimal mechanism design. `http://users.eecs.northwestern.edu/~hartline/omd.pdf`. 3.2

[HT04] Joseph Halpern and Vanessa Teague. Rational secret sharing and multiparty computation. In *STOC*, 2004. 1.1, 1.1.3, 1.1.3, 1.1.5

[IML05] Sergei Izmalkov, Silvio Micali, and Matt Lepinski. Rational secure computation and ideal mechanism design. In *FOCS*, 2005. 1.1.5

[Kat08] Jonathan Katz. Bridging game theory and cryptography: Recent results and future directions. In *TCC*, 2008. 1.1, 1.1.3, 1.1.3, 1.1.5

[kCK09] Yeon koo Che and Jinwoo Kim. Optimal collusion-proof auctions. *Journal of Economic Theory*, pages 565–603, 2009. 1.1.5

[KDC22] Kshitij Kulkarni, Theo Diamandis, and Tarun Chitra. Towards a theory of maximal extractable value I: constant function market makers. *CoRR*, abs/2207.11835, 2022. 1.1.5

[KMSW22] Ilan Komargodski, Shin'ichiro Matsuo, Elaine Shi, and Ke Wu. log*-round game-theoretically-fair leader election. In *CRYPTO*, 2022. 1.1, 1.1.3, 1.1.3, 1.1.5

[KN08] Gillat Kol and Moni Naor. Cryptography and game theory: Designing protocols for exchanging information. In *TCC*, 2008. 1.1, 1.1.3, 1.1.3, 1.1.5

[LSZ19] Ron Lavi, Or Sattath, and Aviv Zohar. Redesigning bitcoin's fee market. In *The World Wide Web Conference, WWW 2019*, pages 2950–2956, 2019. 1, 1.1.5, 2.2

[MM12] Robert C. Marshall and Leslie M. Marx. *The Economics of Collusion: Cartels and Bidding Rings*. The MIT Press, 2012. 1.1.5

[Mye81] Roger B. Myerson. Optimal auction design. *Math. Oper. Res.*, 6(1), 1981. 3.1, 3.2.1, 3.2

[NB15] Alexey Nikitkov and Darlene Bay. Shill bidding: Empirical evidence of its effectiveness and likelihood of detection in online auction systems. *International Journal of Accounting Information Systems*, 16:42–54, 2015. 1.1.5

[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. 1.1, 1.1.3, 1.1.3, 1.1.5

[PS17] Rafael Pass and Elaine Shi. Fruitchains: A fair blockchain. In *PODC*, 2017. 1.1, 1.1.3, 1.1.3, 1.1.5

[QZG22] Kaihua Qin, Liyi Zhou, and Arthur Gervais. Quantifying blockchain extractable value: How dark is the forest? In *43rd IEEE Symposium on Security and Privacy, SP*, 2022. 1.1.5

[QZLG21] Kaihua Qin, Liyi Zhou, Benjamin Livshits, and Arthur Gervais. Attacking the defi ecosystem with flash loans for fun and profit. In *Financial Cryptography and Data Security: 25th International Conference, FC 2021, Virtual Event, March 1–5, 2021, Revised Selected Papers, Part I*, 2021. 1.1.5

[RBO89] T. Rabin and M. Ben-Or. Verifiable secret sharing and multiparty protocols with honest majority. In *STOC*, 1989. 7.2

[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. 1, **??**, 1.1.4, 1.1.5, 1.1.5, 2.2

[Rou21] Tim Roughgarden. Transaction fee mechanism design. In *EC*, 2021. 1, 1.1.1, 1.1.4, 1.1.5, 1.1.5, 1.1.5, 2.2

[SCW23] Elaine Shi, Hao Chung, and Ke Wu. What Can Cryptography Do for Decentralized Mechanism Design? In Yael Tauman Kalai, editor, *14th Innovations in Theoretical Computer Science Conference (ITCS 2023)*, volume 251 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 97:1–97:22, Dagstuhl, Germany, 2023. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. 1, 1, 2, 1.1.5

[TMIY12] Taiki Todo, Takayuki Mouri, Atsushi Iwasaki, and Makoto Yokoo. False-name-proofness in online mechanisms. In *AAMAS*, pages 753–762, 2012. 1.1.5

[WAS22] Ke Wu, Gilad Asharov, and Elaine Shi. A complete characterization of game-theoretically fair, multi-party coin toss. In *Eurocrypt*, 2022. 1.1, 1.1.3, 1.1.3, 1.1.5

[WSC24] Ke Wu, Elaine Shi, and Hao Chung. Maximizing Miner Revenue in Transaction Fee Mechanism Design. In Venkatesan Guruswami, editor, *15th Innovations in Theoretical Computer Science Conference (ITCS 2024)*, volume 287 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 98:1–98:23, Dagstuhl, Germany, 2024. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. 1, 3

[Yao] Andrew Chi-Chih Yao. An Incentive Analysis of Some Bitcoin Fee Designs (Invited

Talk). In *ICALP 2020*. 1, 1.1.5, 2.2

[YG97] Yi-Sheng Yu and Dun-He Gu. A note on a lower bound for the smallest singular value. *Linear algebra and its Applications*, 253(1-3):25–38, 1997. 5.2.1, 5.3.5

[Yok07] Makoto Yokoo. False-name bids in combinatorial auctions. *ACM SIGecom Exchanges*, 7(1):48–51, 2007. 1.1.5

[YSM01] Makoto Yokoo, Yuko Sakurai, and Shigeo Matsubara. Robust combinatorial auction protocol against false-name bids. *Artificial Intelligence*, 130(2):167–181, 2001. 1.1.5

[YSM04] Makoto Yokoo, Yuko Sakurai, and Shigeo Matsubara. The effect of false-name bids in combinatorial auctions: New fraud in internet auctions. *Games and Economic Behavior*, 46(1):174–188, 2004. 1.1.5

[ZCZ22] Zishuo Zhao, Xi Chen, and Yuan Zhou. Bayesian-nash-incentive-compatible mechanism for blockchain transaction fee allocation. https://arxiv.org/abs/2209.13099, 2022. 1, 1.1.5

[ZQC+21] Liyi Zhou, Kaihua Qin, Antoine Cully, Benjamin Livshits, and Arthur Gervais. On the just-in-time discovery of profit-generating transactions in defi protocols. In *IEEE Symposium on Security and Privacy, SP*, 2021. 1.1.5

[Zus] Patrick Zust. Analyzing and preventing sandwich attacks in ethereum. Bachelor's thesis. 1.1.5