A Framework for Single-Item NFT Auction Mechanism Design

Jason Milionis
jm@cs.columbia.edu
Columbia University
Computer Science
New York, NY, USA

Andy Arditi
ava2123@columbia.edu
Columbia University
Computer Science
New York, NY, USA

Dean Hirsch
deanh@cs.columbia.edu
Columbia University
Computer Science
New York, NY, USA

Pranav Garimidi
pg2682@columbia.edu
Columbia University
Computer Science
New York, NY, USA

ABSTRACT
Lately, Non-Fungible Tokens (NFTs), i.e., uniquely discernible assets on a blockchain, have skyrocketed in popularity by addressing a broad audience. However, the typical NFT auctioning procedures are conducted in various, ad hoc ways, while mostly ignoring the context that the blockchain provides, i.e., new possibilities, but at the same time new challenges in auction design. One of the main targets of this work is to shed light on the vastly unexplored design space of NFT Auction Mechanisms, especially in those characteristics that fundamentally differ from traditional and more contemporaneous forms of auctions. We focus on the case that bidders have a valuation for the auctioned NFT, i.e., what we term the single-item NFT auction case. In this setting, we formally define an NFT Auction Mechanism, give the properties that we would ideally like to have such a perfect mechanism to satisfy (broadly known as incentive compatibility and collusion resistance) and prove that it is impossible to have such a perfect mechanism. Even though we cannot have an all-powerful protocol like that, we move on to consider relaxed notions of those properties that we may desire the protocol to satisfy, as a trade-off between implementability and economic guarantees. Specifically, we define the notion of an equilibrium-truthful auction, where neither the seller nor the bidders can improve their utility by acting non-truthfully, so long as the counter-party acts truthfully. We also define asymptotically second-price auctions, in which the seller does not lose asymptotically any revenue in comparison to the theoretically-optimal (static) second-price sealed-bid auction, in the case that the bidders’ valuations are drawn independently from some distribution. We showcase why these two are very desirable properties for an auction mechanism to enjoy, and construct the first known NFT Auction Mechanism which provably possesses such formal guarantees.

CCS CONCEPTS
• Theory of computation → Algorithmic mechanism design; Computational pricing and auctions; Information systems → Online auctions; Applied computing → Online auctions.

KEYWORDS
blockchain; decentralized finance; on-chain auctions; NFTs

ACM Reference Format:
Jason Milionis, Dean Hirsch, Andy Arditi, and Pranav Garimidi. 2022. A Framework for Single-Item NFT Auction Mechanism Design. In Proceedings of the 2022 ACM CCS Workshop on Decentralized Finance and Security (DeFi ’22), November 11, 2022, Los Angeles, CA, USA. ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3560832.3563436

1 INTRODUCTION
Non-fungible tokens (or NFTs in shorthand notation) are blockchain assets (e.g., tokens) that are not interchangeable, i.e., that are “in-capable of mutual substitution” [9]. Chaotic circumstances have repeatedly arisen in the past during public auctions of NFTs, and became especially prominent recently [2], whereby the creation of a disorganized market was blamed on “the way (the organization) set up the sale and the inefficiency of its smart contract.” This showcases the dire and long-overdue need for the study of the NFT Auction Mechanism Design Space to identify the unique aspects of such mechanisms and address the impending issues.

In this work, we provide a formal model for reasoning about how to perform an auction of the kind described above in a decentralized manner, through a blockchain. We focus on the case where bidders have a valuation of the NFT for which they are bidding, in which case the auction is labeled as a “single-item NFT auction.” Note that this includes many common real-world auction formats involving predefined objects, such as something that is in the physical world, e.g., a painting or an album, or a digital object, like the first tweet of Jack Dorsey [8] or a digital art file [11].

Blockchains, in particular, are uniquely positioned to be able to offer brand new characteristics in such auctions. The single biggest advantage is self-evidently the decentralization: everyone is able to take part in any auction, and the auction’s result is verifiable using on-chain data without any explicit trust assumptions to any third-party “auctioneer.” Additionally, there is the ability to be more creative in the types of rules allowed for bids (such as the ability to
direct a portion of a bid someplace else other than the seller, as we will see below), hence broadening the design spectrum. However, these innovative features do not come without a cost; in particular, blockchain implementations of auctions, besides allowing for a greater design space, allow for a far richer set of attack vectors that can be used to compromise or corrupt an auctioning procedure. For instance, due to the free participation in any auction and the pseudonymous identities, there may be “fake bids,” i.e., the seller of the auction may be incentivized to submit their own set of bids to the mechanism, if that could possibly lead to an increased profit for them. Even more importantly, this leads to a great concern of directly applying the second-price auction mechanism to the decentralized setting we have been discussing: the seller can try to arbitrarily approximate the winning bid with their “fake bids” (that they will submit alongside the bids of actual bidders), thus forcing the winning bidder to pay their full bid instead of the optimal second highest actual bid. This would essentially transform the auction to a first-price auction, which would no longer incentivize individual bidders to report their actual valuations for the item to the mechanism, but rather, try and play “guessing games” for the bids that might be placed by others. This fact is well-known in the recent literature of transaction fee mechanism design, where a mechanism, but rather, to try and play “guessing games” for the bids that might be placed by others. This fact is well-known in the recent literature of transaction fee mechanism design, where a similar issue is present (see Section 1.2 for more details).

In our new type of NFT auctioning framework, it is important to emphasize that the seller may now attempt to actively intermeddle in the auction according to their best interests; hence, the auction participants should include not only the bidders but also the seller. Our goal in this framework is to answer the following questions:

(1) How can we formally define a decentralized auctioning mechanism? What are the properties that would constitute perfect theoretical security in such an auction?
(2) Is it possible to create such a perfect auction mechanism?
(3) If this is not possible, then can the properties be relaxed such that they still provide meaningful economic security guarantees?
(4) With these relaxed definitions in place, is it possible now to construct a viable auction mechanism, and if yes, what are the necessary core ideas of such a protocol?

We address all these questions, and introduce our framework: consider that each bidder \( i \in [n] \) (where by \([n]\) we denote the set \( \{1, 2, \ldots, n\}) \) has a (private) valuation \( v_i \geq 0 \) for the auctioned NFT, and reports (bids) \( b_i \geq 0 \) to the mechanism.

**Definition 1** (Single-item NFT Auction Mechanism). A single-item NFT auction mechanism is described by a triplet \((x, p, r)\) where:

1. \( x : \mathbb{R}_{\geq 0}^n \rightarrow [n] \) is the allocation rule that determines which bidder receives an item, i.e., bidder \( x(b_1, \ldots, b_n) \in [n] \) is determined to be the single-item NFT auction winner, who gets the NFT.
2. \( p : \mathbb{R}_{\geq 0}^n \rightarrow \mathbb{R}_{\geq 0} \) is the payment rule which determines the complete payment that each bidder shall pay, i.e., bidder \( i \in [n] \) pays amount \( p_i(b_1, \ldots, b_n) \leq b_i \) to the mechanism.
3. \( r : \mathbb{R}_{\geq 0}^n \rightarrow \mathbb{R}_{\geq 0} \) is the removal rule (or else, burning rule, as it appears in the prior literature; as we note below, there is no need for the funds to be “burnt” in the standard way) specifying how much of the bidders’ bids shall be removed from the current auctioning proceeds, i.e., an amount \( r_i(b_1, \ldots, b_n) \leq p_i(b_1, \ldots, b_n) \) of the \( i \)-th bidder’s payment \( p_i(b_1, \ldots, b_n) \) is “removed” from the possession of all auction participants (seller and bidders alike).

It is implied from the above definition that at the end of the auction, the seller receives an amount of \( \sum_{i=1}^{n} (p_i(b_1, \ldots, b_n) - r_i(b_1, \ldots, b_n)) \).

To the best of our knowledge, Definition 1 is able to capture all of the existing decentralized, single-item NFT auction designs that have been implemented or put forward to date.

Note that the removed amount is considered to be included in the payment rule, i.e., the payment rule subsumes all funds either removed from the bidder or given to the seller. We find this definition to be more natural, because the bidder does not directly care whether a spent amount from their payment is directed to the seller or not.

The crucial characteristic of the “removal” is that the funds need to be unavailable for the auction participants (the seller and the bidders); there is no need to burn them. The utilities of the auction participants are

(1) For the seller, the utility is the amount that they receive:
\[
u_{\text{seller}}(b_1, \ldots, b_n) = \sum_{i=1}^{n} (p_i(b_1, \ldots, b_n) - r_i(b_1, \ldots, b_n)).
\]

(2) For the winning bidder \( i = x(b_1, \ldots, b_n) \), the utility is the difference between their true private valuation and the amount they paid: \( u_i(b_1, \ldots, b_n) = v_i - p_i(b_1, \ldots, b_n) \).

(3) For the non-winning bidders \( j \neq i \in [n] \), the utility is non-positive: \( u_j(b_1, \ldots, b_n) = -p_j(b_1, \ldots, b_n) \).

We continue to intuitively review the three key properties that we would ideally like a perfect auction mechanism to have. These are: the seller incentive compatibility (the seller should not be able to submit fake bids to the auction mechanism in order to improve his/her utility), the bidder incentive compatibility (traditionally known as strategy-proofness, it requires that the bidders have as a dominant strategy to bid truthfully, i.e., to set \( b_i = v_i \), and the off-chain-agreement resistance (or OCA-proofness): that the bidders and the seller cannot improve their utility by colluding off-chain and agreeing to some arcane strategy of coordinating to bid strategically to the mechanism). For the formal treatment of those, please refer to Section 2.

However, we move on with a dire result: our Theorem 1 states that it is impossible to construct such a perfect single-item auction mechanism, while still obtaining revenue from the auction proceeds. In other words, the only auction mechanisms that satisfy the above three properties are the trivial ones, where no-one pays anything, and the seller receives no revenue. Clearly, this is an undesirable auction format, so Theorem 1 critically states that those commonly-thought desirable properties, even though they would allow to defend against collusions of bidders and the seller and at the same time safeguard the interest of all involved parties, cannot be satisfied altogether; some kind of concession has to made in order to construct viable protocols.
In Section 3 we discuss relaxed properties we can hope to achieve even when it is not possible to devise a perfectly collusion-proof protocol. We give the definition of an equilibrium-truthful auction (Definition 3) and an asymptotically second-price auction (Definition 4), for which we argue that, when combined, are good enough relaxations of desirable properties for an incentivizing, reasonable NFT Auction Mechanism to possess. Equilibrium-truthfulness provides a formal truthfulness guarantee for the case of the auction having an untrustworthy seller without collusion among the bidders, and at the same time, if the seller is trustworthy, guarantees that the bidders will bid truthfully according to their true valuations. In other words, so long as the counter-party (bidders/seller) is legitimate, the other side of the auction participants (seller/bidders) will have as a dominant strategy to behave truthfully. Asymptotic second-price auction mechanisms give an approximation guarantee for the revenue obtained through the auction: the seller will asymptotically enjoy the same revenue as the optimal static strategy-proof auction, which is the second-price sealed-bid auction. In this way, any auction format that both satisfies equilibrium-truthfulness and is asymptotic second-price is attractive both for the seller (who will be able to obtain almost the best revenue he/she could hope for) and for the bidders (who will know that they do not need to collude with the seller, or fear that the seller will attempt to “scan” them into paying more than they should).

Perhaps most importantly, we show that we can construct a protocol that is both equilibrium-truthful as well as asymptotically second-price. Our proposed protocol is described in detail in Section 3 and its properties are provided in Theorem 2 and formally proven in Section 4. We believe that our proposed protocol is plausible to implement in current blockchains.

The core ideas behind our protocol are to:

1. create a sealed-bid-like environment where bidders commit to their bids in advance without revealing their bid (by using some assumed-secure hiding commitment scheme),
2. incentivize the revelation of all commits to discourage the seller from trying to insert fake bids so as to attempt to convert the auction to a first-price one to maximize their revenue (by requiring an amount \( L \) to be locked up in the contract along with any commit/bid, which will be returned to the bidder after revealing their bid) and properly punish the unrevealed bids (by completely removing the locked amount \( L \) from the possession of all auction participants: both the bidders and the seller), and
3. discourage the seller from selling the NFT to themselves, through possible very high fake bids to gain information from this futile auction and attempt to perform another auction where he/she will have the information edge on the bidders (we economically prevent this by burning some of the money transferred from the winning bidder to the seller, according to a predefined fee function \( g(\text{second highest price}) \)).

The definition of an asymptotic second-price auction (Definition 4) presupposes that the bidders’ valuations are drawn independently from some distribution \( D \). Notice that, since the protocol we design will only depend on a reasonable lower bound \( L \) of a quantity computable from the distribution \( D \), the assumption that the bidders’ private value stems from the same distribution \( D \) is not particularly strong, because any reasonable heavy-tailed distribution with appropriate entropy and similar-to-expected characteristics can be utilized to compute the asked lower bound.

For the analysis of this protocol, we ignore potential gas fees; for example imagine the auction running on a highly-efficient L2 network. The impact of gas fees would have been a reduction of the utilities of the bidders, but if we further assume “normality” on the underlying chain’s transaction fee mechanism, then the effect of those gas fees would be exactly the same across all bidders, resulting in a simple scaling of all the utilities.

### 1.1 Existing NFT Auction Mechanisms

An overwhelming majority of (single-item) NFT auctions currently take place off-chain, and hence fall short of the formal structure of our framework, as adumbrated above. These off-chain auctions generally take place on centralized web-based platforms, such as OpenSea. The entire auction takes place on this centralized platform, including the design choice of an auction format, as well as the bidding process. Once the auction is complete, the NFT is transferred to the winner, and the winner is charged accordingly. These final settlements are usually the only transactions to take place on-chain. A comparative advantage of off-chain auctions is that they minimize the number of on-chain transactions, and therefore minimize transaction fees. However, current off-chain auctions critically assume an almost complete level of trust in a third-party platform to execute the auction honestly (without almost any verifiability present), hence resembling traditional settings of auctions where there is a completely trustworthy auctioneer. In short, currently-employed off-chain auctions crucially sacrifice security for convenience.

On the other hand, with an on-chain auction, the entire protocol is absolutely transparent and verifiable by the participants. Most importantly, bidders are not required to trust a third-party to run the auction. In this work, we specifically focus on the case when the seller might not be trusted, and so considering on-chain auctions is a natural idea. Currently, there are a few NFT marketplaces which offer on-chain auctions, including Foundation [7] and SuperRare [10].

For both off-chain and on-chain auctions, the set of marketplaces which are popular today generally do not enjoy the truthfulness guarantee of actual second-price sealed-bid (Vickrey) auctions, something which would be very desirable to have. The most popular auction format for single-item NFTs appears to be an English auction format, where the current price is dynamically ascending according to the revealed bids of the users and the highest bid wins. In traditional environments this is paralleled to second-price auctions. However, in an environment of complete anonymity, such as the blockchain and the aforementioned marketplaces, where crucially the seller may (and most likely will) be incentivized to maliciously affect the auction in order to enjoy maximum revenue, the English auction devolves into a first-price auction, because the seller may continually submit “fake bids” in order to artificially increase the price. The outcome will then approximate a first-price, non-truthful auction. Bidders do not have a clear idea how much they should bid. Further, when bidding on-chain, this can result in gas races where bidders rush to get their bids included [2, 3].
1.2 Related Academic Work

From the auction mechanism design perspective with the unique characteristics of blockchains in mind, the first relevant setting [13, 12] is that of transaction fee mechanism design; there, we are interested in designing an auction mechanism for transaction fees (sometimes also called gas fees) on the blockchain. There is, undoubtedly, some semblance in the single-item NFT Auction Mechanism Design case: in fact, our initial definitions and properties can be seen through this lens as smart-contract-mediated adaptations of those given by Roughgarden [13], but with the crucial difference that the original definitions are on a setting where a single entity (the miner) has (mostly) dictatorial control over the allocation of winning bids. NFT auctions critically differ on two aspects: first, on the auction-conducting smart contract’s ability to completely specify the outcome (the smart contract is akin to a “third party” that has the ability to regulate the distribution of NFTs; this introduces major mechanism feasibility differences with transaction fee mechanisms), and second, on the availability and permanent recording of all bids or some appropriate (potentially private) transformation of them on the blockchain. The latter is also in stark contrast to the transaction fee mechanism design space, where crucially not all gas fee bids can be recorded on the blockchain, but only the winning ones. These differences make the respective design spaces significantly different from the perspective of realizable/feasible mechanisms, as we will move on to see from Section 3 onwards.

The design space of transaction fee auctions is also examined in concurrent and independent work by Chung and Shi [5], where, in particular, the authors show that there is no transaction fee auction mechanism that satisfies all of the desiderata and always provides the miner with revenue from transactions.

For standard mechanism design, the problem of trustworthy auctioneers has been studied by Akhbarpour and Li [1]. They define the notion of a credible mechanism where the auctioneer has a dominant strategy to follow the mechanism. For example, a Vickrey auction is not credible because the auctioneer has an incentive to fabricate a bid right below the highest bid to drive up their revenue. Akhbarpour and Li [1] shows that, while a mechanism is desired to be sealed-bid, credible, and strategy-proof, only two out of the three properties can be satisfied if only winners of the auction make payments. A natural question to ask is whether this impossibility result can be bypassed using cryptographic commitments to make it harder for the auctioneer to act untrustworthy. For the case of Vickrey auctions this would make it so that the auctioneer does not have knowledge of the bids’ values until they are publicly known. One implementation of such a sealed-bid auction is to have two phases: a commit phase, and a reveal phase. In the commit phase, all bidders output some cryptographic commit to their bid value: commit(bid|nonce). In the reveal phase, bidders reveal the pair (bid, nonce). One issue with this implementation is that not all bids may be revealed in the reveal phase. Ferreira and Weinberg [6] propose a solution to this problem by fining bidders for unrevealed commits, and paying these fines to the winning bidder. Our solution similarly intends to incentivize the revelation of all bids. A core difference between our proposal and that of Ferreira and Weinberg [6] is that we operate over a blockchain, inheriting its corresponding capabilities and assumptions; in particular, the anonymity of the seller as an entity (that allows them to provide fake bids to the auction) means that providing the penalties to him/her would no longer guarantee their Incentive Compatibility. Even if one fixes this issue, the additional attack of the seller bidding to win the NFT themselves (just to gain information about the auction participants’ bids) and repeatedly running auctions arises. We show a protocol that deals with these attacks, and formally prove that it is properly incentivizing all auction participants to behave truthfully.

Lastly, Chen et al. [4] treats NFTs as asset-backed securities, allowing for fractional ownership of NFTs through what they call “repurchase protocol” which then enables performing repeated auctions of the same NFT. However, the proposed method is still vulnerable to common shortcomings of traditional auction mechanisms when applied to blockchain settings, as mentioned before.

2 PROPERTIES AND IMPOSSIBILITY RESULT

Before we define the properties, we will need to formally describe an Off-Chain Agreement (OCA) for our purposes, along with the joint utilities of its participants. In particular, an OCA is a coalition of bidders $S \subseteq [n]$ with the seller $s$ who collude and submit a specific agreed-upon set of bids $b' \in \mathbb{R}_{\geq 0}^k$ for some $k \geq 0$ (usually different from the ones that they would normally submit) to the auction mechanism such that in total, they would be better off in their joint utility (which is the sum of their individual utilities)$^2$. In this way, they will be able to also obtain better individual utilities, for instance by splitting the joint utility differential that arises among all of them. More specifically, we define the joint utility of such an OCA as the aforementioned sum, where $b_{-S}$ are the bids of the rest of the bidders that are not part of the coalition:

$$u_{\text{joint}}(S, s)(b', b_{-S}) = u_{\text{seller}}(b', b_{-S}) + \sum_{i \in S} u_i(b', b_{-S}).$$

We shall write $u_{\text{joint}}(s)$ whenever we want to signify the particular coalition of the bidder $i$ with the seller and the meaning is clear from the context. We now move forward to provide the properties that one could possibly desire such an NFT auction mechanism to satisfy.

Definition 2 (Desirable properties of a single-item NFT Auction Mechanism). We give the three properties as follows:

1. **Seller incentive-compatibility (seller IC):** For any set of bids by bidders $b_1, \ldots, b_n$, if the seller submitted any set of “fake bids” $b_{n+1}, \ldots, b_m$, then they would not be able to obtain any more utility than they already get without submitting any fake bids:

$$u_{\text{seller}}(b_1, \ldots, b_n) = \sum_{i=1}^{n} (p_i(b_1, \ldots, b_n) - r_i(b_1, \ldots, b_n))$$

$$\geq \sum_{i=1}^{n} (p_i(b_1, \ldots, b_m) - r_i(b_1, \ldots, b_m))$$

$$- \sum_{j=m+1}^{m} r_j(b_1, \ldots, b_m).$$

$^2$We assume that the OCA cannot consist of fully bypassing the auction mechanism, i.e., the NFT still has to be distributed through the mechanism, even, e.g., when all the parties agree to collude and only the agreed winner submits a fake bid of 0.
(2) Bidder incentive-compatibility (bidder IC): For any bidder \( i \in [n] \), for any set of bids by (all but the \( i \)-th) bidders \( b_1, \ldots, b_{i-1}, b_{i+1}, \ldots, b_n \) (jointly denoted as \( b_{-i} \in \mathbb{R}_{\geq 0}^{n-1} \)), the \( i \)-th bidder’s utility is maximized exactly when they bid their true valuation, i.e., for any potential bid \( b_i \), it holds that 
\[
u_i(b_1, \ldots, b_{i-1}, v_i, b_{i+1}, \ldots, b_n) \geq \nu_i(b_1, \ldots, b_{i-1}, b_i, b_{i+1}, \ldots, b_n).
\]

(3) Off-Chain Agreement resistance (OCA-proofness): For any set of bidders \( S \subseteq [n] \), bids by bidders not belonging to \( S \): \( b_{|S|+1}, \ldots, b_n \) (jointly denoted as \( b_{-S} \in \mathbb{R}_{\geq 0}^{n-|S|} \)), if we consider the OCA of the seller \( s \) with a set \( S \) of bidders, then this coalition of auction participants is not able to obtain higher joint utility through any bids they might agree to submit, i.e., for any \( k \geq 0 \) and for any \( b' \in \mathbb{R}^k \), 
\[
u_{\text{joint}}(s, S)(b', b_{-S}) \leq \nu_{\text{joint}}(s, S)(b_S, b_{-S}),
\]
where \( b_S \in \mathbb{R}_{\geq 0}^{|S|} \) are the bids that the colluding bidders would submit on their own without being part of that OCA.

We are now ready to state our impossibility result.

**Theorem 1.** (No perfect NFT auction mechanism). There is no single-item NFT auction mechanism as considered in Definition 1, satisfying the properties of bidder IC and OCA-proofness of Definition 2, where some bidder pays the seller some non-zero (positive) amount.

Theorem 1 implies that the only feasible perfect mechanisms necessitate paying nothing to the seller, which is clearly undesirable as an auction format. In the interest of brevity, we omit the analytical, case-based proof of this Theorem.

3 RELAXATION OF PROPERTIES AND OUR PROPOSED PROTOCOL

Theorem 1 implies that we cannot hope for the design of a perfect auction mechanism. However, we may be interested to relax our provided guarantees such that the design of a mechanism with sufficiently secure “relaxed” properties is feasible. In particular, we give a first relaxed definition that defends the individual incentives to bid truthfully given that there exists at least some side that is not rogue:

**Definition 3.** An auction is equilibrium-truthful when:

1. The seller’s dominant strategy, assuming bidders bid truthfully (i.e. \( b_i = v_i \)), is to not post any fake bids, and
2. Each bidder’s dominant strategy is to bid truthfully, assuming the seller is not posting any fake bids.

Further, in order for such an auction mechanism to be additionally considered reasonable, it makes sense that we require that running such an auction will produce a favorable economic result for the seller. For this, we define the following auction property:

**Definition 4.** Given an auction in which at least \( n \) participants are known to participate, and have their valuations independently drawn from the same distribution \( D \), we say that an auction is asymptotically second-price when the expected utility of the seller is 
\[(1 - o(1)) \cdot \mathbb{E}[b_2],
\]
assuming bidders bid truthfully and that there are no fake bids, where \( b_2 \) is the second-highest bid, and \( o(1) \) goes to 0 as \( n \to \infty \).

**Theorem 2.** Assume that there is a continuous distribution \( D \) from which the valuations \( v_i \) are independently drawn. We denote by \( f \) the probability density function of \( D \), and by \( F \) the cumulative distribution function. Then, if 
\[
\sup_{s \in \text{support}(f)} \frac{1 - F(s)}{f(s)} < \infty,
\]
then there exists a protocol that is both equilibrium-truthful and asymptotically second-price, as per Definitions 3 and 4.

Notice that the requirement of Equation (1) is satisfied for many natural distributions, including the uniform, exponential, and log-normal distributions. For the main ideas of the protocol, please refer to Section 1. We now detail the protocol, and prove Theorem 2 in Section 4.

(1) The seller sets up a contract, to which bidders should send commits (i.e., \( \text{hash}(\text{bid}||\text{nonce}) \)), and the seller sends the NFT to the contract.

(2) Commit phase: Bidders send their commits to the contract, together with an amount \( L \) (in tokens) to be locked up. Without both of these ingredients, the bid is not taken into account.

(3) After adequate time has passed, e.g., enough blocks confirmed, the commit phase ends.

(4) Verification phase: Now bidders send to the contract their actual bid amounts (and nonce for verifying). We emphasize that the funds themselves are not sent to the contract at this point, but only the bid amounts the bidders committed to.

(5) At the end of the verification phase, for any unrevealed bid, their locked up amount \( L \) is removed. The contract can now look for the highest revealed bid \( b_1 \) and second-highest revealed bid \( b_2 \).

(6) An amount of \( \min\{b_2, L\} \) from the winner remains at the contract for the seller. If \( b_2 \leq L \), then the NFT is unlocked for transfer to the winner (the other bidders that revealed their commits will need to initiate a transaction to the contract to receive back their locked amounts). Otherwise, if \( b_2 > L \), then an additional \( b_2 - L \) amount should be sent from the winner to the contract before the NFT is unlocked. The seller can withdraw \( x = g(x) \), where \( x \) is the funds received in the contract by the winner (including the initial \( \min\{b_2, L\} \)).

(7) After some predefined time, the auction ends. If the highest bidder did not transfer the funds, they still lose the \( L < b_2 \) they have locked up, and additionally do not receive the NFT.

4 PROTOCOL ANALYSIS

The main purpose of this section is to give an overview of the proof of Theorem 2 (see Section 4.3).

4.1 Fake Bids by the Seller

We analyze the attack of multiple bids in the model without fees, to reason about the equilibrium truthfulness for the seller, i.e., Definition 3. Suppose that \( B_1 \) and \( B_2 \) are the highest and second-highest bids, respectively. Suppose the seller tries to submit multiple fake bids. We will analyze under which conditions this attack does not benefit the seller’s utility. In particular, the seller chooses a set \( S \) of

3More generally, it is satisfied for most unimodal distributions, such that \( f(s) = O(f'(s)) \).
bids, and just before the verification phase ends, the seller gets to inspect all revealed bids and choose which of his/her fake bids in \( S \) to reveal.

Observe that the seller does not lose anything by revealing all \( S \) commits that are lower than \( B_1 \), and in fact will gain by getting to keep the locked up funds used for those bids. Now, assume that the seller wants to sell the NFT in the current auction (the other case of repeated auction would be examined in Section 4.2) they also must not reveal any commit higher than \( B_1 \). Therefore, the seller loses \( L \cdot \{s \in S: s > B_1\} \), but gains \( \max\{s\in S \cup \{B_2\}: s \leq B_1\} - B_2 \) (that is, the amount by which it raised the second-highest bid).

The seller therefore maximizes their utility as:

\[
\mathbb{E}\max\{0, \max\{s \in S: s \leq B_1\} - B_2\} - L \cdot \mathbb{E}\{s \in S: s > B_1\},
\]

where the expectation is over the values of \( B \) (that is, the amount by which it raised the second-highest bid).

Therefore, in order to ensure that the dominant strategy of the seller is not to place any fake bids, we need only ensure that there is no favorable single fake bid, which happens when:

\[
\text{Corollary 1. If the distribution of the } n \text{ bidders are drawn independently at random from a distribution } D \text{ with a probability density function } f(x) \text{ and cumulative distribution function } F(x) = \int_{-\infty}^{x} f(t)dt, \text{ we have that (we remind the reader that the probabilities are over the randomness of the } b_i \text{'s)}
\]

\[
\text{Pr}(B_1 < s) = \text{Pr}(\forall i : b_i < s) = F(s)^n.
\]

Further, by symmetry,

\[
\text{Pr}(B_2 < s < B_1) = n \cdot \text{Pr}(B_2 < s < B_1 \land \text{bidder 1 is the winner}) = nF(s)^{n-1}(1 - F(s)).
\]

Also, using the fact that \( \mathbb{E}[X] = \int_{-\infty}^{\infty} \text{Pr}(X > x)dx \) for any non-negative random variable \( X \), we get that

\[
\mathbb{E}(B_2|B_2 < s < B_1) = \int_{s}^{\infty} \text{Pr}(B_2 > x|B_2 < s < B_1)dx
\]

\[
= s - \int_{0}^{s} \frac{1}{F(s)^{n-1}(1 - F(s))} \int_{0}^{s} F(x)^{n-1}(1 - F(x))dx
\]

Substituting the formulae to the bound on \( L \), we obtain the simplified corollary:

\[
\text{Corollary 2. When all bidders draw their private value independently at random from a distribution with a cumulative probability function } F, \text{ the auction is equilibrium-truthful for the seller so long as there is no repeated auction exactly when}
\]

\[
L \geq n \cdot \frac{1}{F(s)^n} \int_{0}^{s} F(x)^{n-1}(1 - F(x))dx,
\]

where the maximum is taken over all values of \( s \) in the support of \( f \).

Under a distribution with specified lower bound on its support, the following Theorem can be shown (see Appendix B):

**Theorem 3.** Suppose that all bidders draw their private value independently at random from a distribution supported on \([a, \infty)\) for some \( a > 0 \), with a cumulative probability function \( F \) and density function \( f \), such that \( \frac{1 - F(s)}{f(s)} \) is bounded from above on the support of \( f \). Let \( L(n) \) be the lower bound on \( L \) for \( n \) bidders. Then, the auction is equilibrium-truthful for the seller so long as there is no repeated auction, if \( L(n) \sim \frac{1}{n} \cdot \sup_{x \in \text{support}(f)} \frac{1 - F(s)}{f(s)} \) as \( n \to \infty \).

### 4.2 Repeated Auction

Let, again, the highest bid be \( B_1 \) and the second-highest bid be \( B_2 \), unknown to the seller. We will analyze the conditions that make the seller’s utility be maximized without any repeated auction.

In the first auction, the seller pays \( B_1 \) and receives \( B_1 - g(B_1) \) (where \( g \) is the fee function), so in total they lose \( g(B_1) \). Had the seller not done this attack, the total gain would have been \( B_2 - g(B_2) \).

In the second auction, the seller posts \( B_1 - \varepsilon \) as a bid, thus receiving essentially \( B_1 - g(B_1) \) for the item. Thus, in total, instead of receiving \( B_2 - g(B_2) \), the seller receives \( B_1 - 2g(B_1) \). So the seller is acting truthfully if and only if

\[
B_2 - g(B_2) \geq B_1 - 2g(B_1) \iff B_1 - B_2 \leq 2g(B_1) - g(B_2).
\]

Thus, if we find a fee function \( g \) such that \( B_1 - B_2 \leq 2g(B_1) - g(B_2) \) (in expectation), the seller would be acting truthfully, and not try to insert fake bids so as to perform repeated auctions. Assuming a nondecreasing fee function, where \( g(B_1) \geq g(B_2) \), it is enough to require that \( B_1 - B_2 \leq g(B_1) \), i.e., that the fee provides an estimate for an upper bound on the gap between the two highest bids.

**Theorem 4 (Appendix C).** If the distribution \( D \) satisfies

\[
\sup_{x \in \text{support}(f)} \frac{1 - F(s)}{f(s)} < \infty,
\]

then choosing the fee function \( g(x) = ax \) for an appropriate \( a = o(1) \), i.e., such that \( \lim_{n \to \infty} a(n) = 0 \), guarantees that the seller’s dominant strategy is to not run repeated auctions by inserting fake bids.

### 4.3 Final Proof of Theorem 2

Proof. We prove that the protocol described with choosing the locked-amount \( L \) as in Theorem 3 and the fee function \( g \) as in Theorem 4, satisfies the theorem.

We first show that the auction described is equilibrium-truthful. Indeed, for Item 2 of Definition 3, suppose that the seller does not place any fake bids. Then, since the amount each bidder pays in case they win is the second highest price, the auction satisfies the conditions for the bidder IC of second-price auction, and it follows that the best strategy of any bidder is to bid truthfully (i.e., \( b_i = v_i \)).

On the other hand, for Item 1 of Definition 3, assuming all bids are truthful, the seller’s dominant strategy is to not place any fake bids, as shown in Theorems 3 and 4.

Finally, Theorem 4 additionally shows that the auction is asymptotically second-price, since the utility for the seller is \( (1 - \alpha)\mathbb{E}[B_2] \) with \( \alpha = o(1) \).
REFERENCES

[1] Mohammad Akhbarpour and Shengwu Li. 2018. Credible mechanisms. In Proceedings of the 2018 ACM Conference on Economics and Computation, Ithaca, NY, USA, June 18-22, 2018. Eva Tardos, Edith Elkind, and Rakesh Vohra, (Eds.) ACM, 371–372. doi: 10.1145/3219666.3219669.

[2] Daniel Van Boom. [n. d.] How Bored Ape Yacht Club Broke Ethereum. en. (2022). Retrieved May 5, 2022 from https://www.cnet.com/personal-finance/crypto/how-bored-ape-yacht-club-broke-ethereum/.

[3] Vitalik Buterin. 2021. Alternatives to selling at below-market-clearing prices for achieving fairness (or community sentiment, or fun). https://vitalik.ca/general/2021/08/22/prices.html.

[4] Hongyan Chen, Yukun Cheng, Xiaotie Deng, Wenhuan Huang, and Luizbon Rong. 2022. Absent: securitization and repurchase scheme for non-fungible tokens based on game theoretical analysis. arXiv preprint arXiv:2202.02199.

[5] Hao Chung and Elaine Shi. 2021. Foundations of transaction fee mechanism design. (2021). arXiv:2111.01351 [cs.GT].

[6] Matheus V. X. Ferreira and S. Matthew Weinberg. 2020. Credible, truthful, and two-round (optimal) auctions via cryptographic commitments. In Proceedings of the 21st ACM Conference on Economics and Computation (EC ’20). Association for Computing Machinery, Virtual Event, Hungary, 683–712. isbn: 9781450385541.

[7] Tim Roughgarden. 2021. Transaction fee mechanism design. In Proceedings of the 22nd ACM Conference on Economics and Computation (EC ’21). Association for Computing Machinery, Budapest, Hungary, 792. isbn: 9781450385541. doi: 10.1145/3465456.3467591.

A PROOF OF LEMMA 1

Suppose there is a set with a positive expected utility. Then there is a set with minimal size among all such sets. This set is nonempty, since for the empty set the expected utility is 0. Let this set be $S$. Assume for the sake of contradiction that $|S| > 1$, and take $s_1 = \max S$ and let $S_2 = \max \{s \in S : s < s_1\}$.

By the assumption that $S$ has minimal size among all sets with positive expected utility, if $s_1$ is taken out, the utility becomes non-positive, and in particular decreases. On the other hand, the expected utility decreases only when $B_2 < s_1 < B_1$, by: $E[S_1 - B_2 | B_2 < s_1 < B_1] Pr(s_2 < s_1 < B_1) + (s_1 - s_2) Pr(B_2 < s_1 < B_1)$ (where we used the law of total expectation).

Therefore, the claim that the set $\{s_1\}$ has a positive utility is equivalent to the claim that $LPr(s_1 > B_1) < E[S_1 - B_2 | B_2 < s_1 < B_1] Pr(B_2 < s_1 < B_1)$, so the Lemma will follow if we show that

$$E[S_1 - B_2 | B_2 < s_1 < B_1] Pr(B_2 < s_1 < B_1)$$

is at least

$$E[S_1 - B_2 | B_2 < s_1 < B_1] Pr(s_2 < B_2 < s_1 < B_1) + (s_1 - s_2) Pr(B_2 < s_1 < B_1)$$

Indeed, by further conditioning on $B_2$ and using the law of total expectation, we can see that the first value equals

$$\mathbb{E}[S_1 - B_2 | B_2 < s_1 < B_1] Pr(s_2 < B_2 < s_1 < B_1) + (s_1 - s_2) Pr(B_2 < s_1 < B_1),$$

which is larger than the second value because $s_1 - B_2 > s_1 - s_2$.

B PROOF OF THEOREM 3

We first fix a value of $s$. By Fubini, we can switch between the limit of $n \to \infty$ and the supremum over $s$, if by first finding the limit for fixed $s$ and then taking the maximum over the support, we can obtain a finite result.

For a fixed value of $s$, we make a change of variables $t = F(s)$ in the integral, so $dx = \frac{dt}{f(F^{-1}(t))}$. Thus the integral is:

$$\int_0^\infty F(x)^n - 1 - F(x) dx = \int_0^\infty t^n - 1 - t \frac{dt}{f(F^{-1}(t))}.$$ 

We now note that the part of the integral that is bounded above by $F(s) - \epsilon$ is thus also upper-bounded by $O((F(s) - \epsilon)^n)$ and would therefore not contribute asymptotically even after dividing by $F(s)^n$ (note that we have $t < 1$ in the integral). Thus we can approximate this integral asymptotically by replacing $(1 - t)/f(F^{-1}(t))$ with its value on the upper limit of the integral; that is, replace $(1 - t)/f(F^{-1}(t))$ with $(1 - F(s))/f(F(s)) = (1 - F(s))/f(F(s))$. Now the integral becomes

$$\int_0^\infty F(x)^n - 1 - F(x) dx \sim 1 - F(s) \int_0^\infty \frac{F(s)}{nF(s)} dt = 1 - F(s)$$

Thus, dividing by $F(s)^n$ we finally obtain the limit as $L(n, s) = \frac{1 - F(s)}{nF(s)}$. Taking the supremum over $s$ yields the stated result.

C PROOF OF THEOREM 4

We find $\mathbb{E}[B_1] - \mathbb{E}[B_2]$. For that, we use the CDF of $B_1$ $F_B(x) = F(x)^n$, and the CDF of $B_2$ $F_B(x) = Pr(B_1 < x) + Pr(B_1 > x) \sim F(x)^n + nF(x)^n - 1 - F(x))$. Thus, since for a nonnegative random variable $X$ we have $\mathbb{E}[X] = \int_0^\infty Pr(X \geq x) dx = \int_0^\infty (1 - F(x)) dx$, we get that

$$\mathbb{E}[B_1] - \mathbb{E}[B_2] = \int_0^\infty \left(1 - F(x)^n - (1 - F(x)^n) - nF(x)^n - 1 - (1 - F(x))\right) dx$$

$$= n \int_0^\infty (1 - F(x)) dx$$

$$= n \int_0^\infty F(x)^{n - 1} \frac{1 - F(x)}{F(x)} dx.$$ 

Let $A = \sup_{x} \frac{1 - F(x)}{F(x)}$, which is finite by the Theorem’s premise. Then we finally obtain

$$\mathbb{E}[B_1] - \mathbb{E}[B_2] \leq nA \int_0^\infty F(x)^{n - 1} f(x) dx = AF(x)^n |_0^\infty = A$$

We will now show that we can always choose

$$\alpha = 2 \frac{\mathbb{E}[B_1] - \mathbb{E}[B_2]}{\mathbb{E}[B_2]}.$$
We evidently have $\mathbb{E}[B_1 - B_2] < \mathbb{E}[g(B_2)]$.

To show that $\lim_{n \to \infty} \alpha(n) = 0$, we distinguish between two cases:

- $\lim_{n \to \infty} \mathbb{E}[B_1] = \infty$. In this case, we also have $\lim_{n \to \infty} \mathbb{E}[B_2] = \infty$, since we have shown that the difference $\mathbb{E}[B_1] - \mathbb{E}[B_2]$ is finite. Therefore $\alpha(n) \leq 2 \frac{\mathbb{E}[B_2]}{\mathbb{E}[B_1]} \to 0$.

- $\lim_{n \to \infty} \mathbb{E}[B_1] = M < \infty$ for some finite $M$. We note that the limit necessarily exists, since this expectation is nondecreasing with $n$. In this case, $\mathbb{E}[B_2]$ also has the same limit $M$ (as shown below), hence the numerator approaches 0 and the denominator is nondecreasing, giving $\lim_{n \to \infty} \alpha(n) = 0$.

It remains to prove that $\lim_{n \to \infty} \mathbb{E}[B_2] = M$. Since $B_2 \leq B_1$, it suffices to prove that $\lim \inf \mathbb{E}[B_2] \geq M$. Indeed, we have that $B_2(n)$ is at least the minimum between the best bids between any two arbitrary disjoint sets of the $n$ bids. In particular, we obtain that $B_2(n)$ is at least the minimum between the maximum of the first $n/2$ bids and the maximum of the last $n/2$ bids. Thus if we show that for all $\varepsilon > 0$ we have that $\lim_{n \to \infty} \mathbb{P}(B_1 < M - \varepsilon) = 0$, it would follow that $\lim \inf \mathbb{E}[B_2(n)] \geq M - \varepsilon$, for any $\varepsilon > 0$, so $\lim \inf \mathbb{E}[B_2(n)] \geq M$ and we will be done.

For that, we notice that $\mathbb{P}(B_1 < M - \varepsilon) = F(M - \varepsilon)^n$, so it remains to show that $F(M - \varepsilon) < 1$. Indeed, if $F(M - \varepsilon) = 1$, we would also have $\mathbb{P}(B_1 < M - \varepsilon) = 1$ exactly (i.e. not in the limit), hence $\mathbb{E}[B_1] \leq M - \varepsilon$, a contradiction.

This concludes the proof.