An Initial Framework for NFT Auction Mechanism Design: Impossibility Results and Solutions

Andy Arditi* UNI: ava2123 ava2123@columbia.edu

Pranav Garimidi* UNI: pg2682 pg2682@columbia.edu

Dean Hirsch* UNI: dh3070 deanh@cs.columbia.edu

Iason Milionis* UNI: im2605 jm@cs.columbia.edu

1 Introduction

1.1 Background

Non-fungible tokens (or NFTs in shorthand notation) are blockchain assets (e.g., tokens) that are not interchangeable, i.e., that are "incapable of mutual substitution." (Merriam-Webster) One is able to precisely distinguish any one from any other and point to the owner of each one, whereas, for instance, units of traditional cryptocurrencies (such as Bitcoin or Ethereum) are indistinguishable. NFTs can thus serve as certificates of ownership and/or rights for a particular object, physical or digital, and may be auctioned, just like an actual painting would, for example. The underlying asset is often only cryptographically connected to the NFT, since it would be inefficient to include the whole portion of a digital file in the case of digital assets, or simply impossible in the case of physical assets. Building upon that, NFTs can be efficiently implemented on top of a blockchain, and then the natural question that arises is how to perform an auction of the kind described above in a decentralized manner, through a blockchain. One more significant factor to consider, though, is that, besides representing the ownership of objects, some NFTs may also have intrinsic value representing ownership of themselves instead of an external object. This, in turn, makes it reasonable to define the "rarity" of an NFT relative to another. The pricing of such NFTs has been known to be highly uncertain, and auctions of those to be susceptible to extreme volatility (Kugler, 2021).

The above two cases for minting (otherwise also called "initial drop" or "launch") NFTs correspond to the amount of information that the initial-drop auction bidders would have at their disposal. First, one could have an NFT act *solely* as a digital certificate of ownership/rights for a predefined object 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 (Locke, 2021) or a digital art file (Reyburn, 2021), whereby the NFT's minting would be verifiably made by the original creator. The bidder in this case knows exactly the underlying object whose ownership (or whose rights) is concerned and might have an underlying valuation in mind for the specific NFT (as they would, for instance, for the original Mona Lisa painting). A different format is to have NFTs that define their own rarity through on-chain verifiable attributes, and thus serve as a piece of "digital art" themselves (not merely as a certificate of ownership/rights). In this second regime, the rarity of an NFT to be minted is unknown and determined pseudo-randomly with the minting process. Therefore, the bidder chooses the amount to bid *without knowing* in advance the rarity of the NFT they will receive (ideally[1]). Hence, the bidder in this case ideally does not have an a priori valuation of the NFT they bid on, but may have a probability distribution on potential valuations. Thus, it is naturally motivating to ask the question of how to design truthful, efficient, strategy-proof auction mechanisms with the above core principles of NFTs in mind. For the first significant part of our analysis we focus on the case where bidders have a clear valuation of the NFT for which they are bidding, in which case the auction is labeled as a "single-item NFT auction."

Blockchains, in particular, are uniquely positioned to be able to offer brand new characteristics in 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

*Alphabetical Order

[1] There have been cases where particularly technically-adept participants were able to settle the pseudorandomness of the minting mechanism and "snipe" extremely rare tokens, violating the intended random distribution of rare-attribute NFTs.

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 optimally-calculated difference between their bid and the 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, to try and play "guessing games" for the bids that might be placed by others. This fact is well-known in the recent sibling literature of transaction fee mechanism design, where a similar issue is present (Section 1.1.2 follows with more details).

In our new type of NFT auctioning framework, there are in general three categories of auction "game players" to be recognized and accounted for. First, there are the traditional bidders, who may place bids in a decentralized auction. Second, there is the seller, who may now attempt to intermeddle in the auction according to their best interests, as mentioned in the previous paragraph. Third, there is what is traditionally called the "auctioneer"; in our case, this is assumed to be a "smart contract," i.e., a self-executing piece of code that is able to execute the full auction proceeds on its own in the blockchain that it runs atop. Note that the mechanisms we will describe are crucially based on the assumption that the underlying mechanism will precisely, completely and error-free be coded on a smart contract implementation; in many cases, there are well-known security issues, but we will not deal with these practical attack considerations throughout the rest of our paper.

1.1.1 Existing NFT Auction Mechanisms

An overwhelming majority of NFT auctions currently take place off-chain, and hence do not fall under the scope 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 definition 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, off-chain auctions critically assume an almost complete level of trust in a third-party platform to execute the auction honestly, hence resembling traditional settings of auctions where there is a completely trustworthy auctioneer. In short, 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 project, we specifically focus on the case when the auctioneer is not trusted, and so considering on-chain auctions is a natural idea. There are only a couple of current NFT marketplaces which offer on-chain auctions, including Foundation (Howard, 2021) and SuperRare (Perkins, 2020).

For both off-chain and on-chain auctions, the set of marketplaces which are popular today generally do not offer second-price (Vickrey) auctions. For example, Foundation offers an English auction with a reserve price. In this type of auction, bidders can see the current highest bid before submitting their own bid. From our survey of the current NFT auction landscape, this appears to be the most popular auction format for individual, non-random NFTs (aside from "buy-now" sales). These first-price auctions suffer from being non-truthful and may incentivize participants to lie about their values for an NFT to a significant extent. Bidders do not have a clear idea how much they should bid. Additionally, when bidding on-chain, this can result in gas races where bidders race to get their bids included (Buterin, 2021a). This is exacerbated by the fact that in a first-price auction, bidders are incentivized to repeatedly cancel old bids and make new bids as the currently winning price moves.

1.1.2 Related Academic Work

The first relevant setting (Roughgarden, 2020, 2021), similar to the category of single-item NFT auctions that we examine, 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. NFT auctions differ mostly on two aspects: first, on the auction-conducting smart contract's ability to completely specify the outcome (chosen distribution of NFTs), and second, on the random value initial NFT drops (or, NFT launches) which are a complete divergence from the traditional paradigm as seen by prior work. However, there is also some semblance in the single-item case: in fact, our initial definitions and properties can be seen through this lens as smart-contract-mediated adaptations of those given by Roughgarden (2020), 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. Finally, in concurrent and independent work to ours, Chung and Shi (2021) also examine the design space of transaction fee auctions, and in particular, show that there is no transaction fee auction mechanism that satisfies all of the desiderata and always provides the miner with revenue from transactions.

We now move on to review existing work which discusses practical implementations of Vickrey auctions.

One practical issue that comes up with Vickrey auctions is that of bidder collusion. To see how the bidders can collude, consider the following example: suppose that bidder $i \in \{1, 2, \ldots, n\}$ intends to bid b_i, where b_1 is the largest intended bid. The colluding bidders can establish these intended bid values ahead of time, and agree to place the following actual bids: bidder 1 will bid b_1, while all the other bidders will bid the minimum reserve price. Thus, bidder 1 will win the auction, paying the minimum reserve price, and can then pay the other cooperating bidders each a small reward. Micali and Rabin (2014) show how to prevent this bidder collusion by using cryptographic tools to make the cartel agreements unenforceable. This is achieved by having the auctioneer reveal as little information as possible: the auctioneer proves to the winner that he is in fact the winner, and that the second price was indeed the second price. Thus, with all other information unknown, it is impossible to tell which party broke a cartel agreement. Further, the auctioneer can incentivize bidders to do so by paying a kickback to the second-highest bidder. While this problem of bidder collusion in Vickrey auctions is interesting, we do not consider the problem of bidder collusion in this report, and focus instead on the case where the auctioneer is not trustworthy.

For standard mechanism design, the problem of untrustworthy auctioneers has been studied by Akbarpour and Li (2018). 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. Akbarpour and Li (2018) show 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 of their bid value $commit(bid||nonce)$. In the reveal phase, bidders reveal $(bid, nonce)$. One issue with this implementation is that not all bids may be revealed in the reveal phase. Ferreira and Weinberg (2020) propose a solution to this problem by fining bidders for unrevealed commits, and paying these fines to the winning bidder. We propose a similar solution to incentivize that all bids are revealed. A core difference between our proposal and that of Ferreira and Weinberg (2020) is that we operate over a blockchain, while Ferreira and Weinberg (2020) attempt to describe a protocol without assuming one, leading to a different set of assumptions.

1.2 Our Contributions

In the landscape delineated in the above section, we make contributions with the target of formalizing the various implemented, proposed and conceived decentralized NFT auction mechanisms, to shed light on the design considerations a mechanism designer should have in mind when selecting or proposing their own, as well as make our own proposal for viable auction mechanisms.

In particular, we begin our endeavors with the examination of single-item NFT auctions like the initial offering of a newly-minted single NFT from an original artist or distributor, or the re-sale of a currently existing NFT. In Section 2, we formally define the notion of a single-item NFT auction mechanism as a collection of three rules: an allocation, a payment and a removal rule, and give three key properties that we would ideally like an auction mechanism to have. The removal rule is the new ability that a blockchain can guarantee through the automated, verifiable nature of smart contracts to remove part of any bid from the process, both from the seller and the bidders. Typically, this had been referred to as a burning rule, but we consider that its use is indeed much broader and burning is only one very restrictive implementation of such a rule. We continue in Section 3 to provide the reader with a dire outcome: it is impossible to construct such a "good" single-item auction mechanism. This critically states that the desirable properties we chose, even though they would able 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 4 we discuss relaxed properties we may hope to achieve even when it is not possible to devise such a collusion-proof protocol. We show that we can still achieve other desirable properties, in the case of an untrusted seller but assuming no collusion among the bidders, using a smart contract. We believe that such a protocol is plausible to implement in current blockchains. We show an example of such a protocol in Section 5 and prove the main properties for it in Theorem 2, which is a culmination of an extensive discussion and analysis of the protocol done in Section 6.

After the above outcomes, we shift gears to talking about NFT launches. Section 7 differs from the others in that we consider the case where there is a *collection* of NFTs to be distributed, rather than a single NFT. We discuss the desired properties of such an NFT launch, and then review existing NFT launch designs, pointing out which properties they each satisfy, and which they do not. We then generalize the important design challenges in the space. Finally, we briefly suggest some new designs which could overcome these challenges.

2 Definition and Properties

Consider that each bidder $i \in [n]$ 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 $(x, \boldsymbol{p}, \boldsymbol{r})$ where:

- $x : \mathbb{R}^n_{\geq 0} \to [n]$ 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.
- $\boldsymbol{p} : \mathbb{R}^n_{\geq 0} \to \mathbb{R}^n_{\geq 0}$ is the *payment rule* which determines the complete payment that each bidder shall pay, i.e., bidder $i \in [n]$ pays amount $p_i(b_1, \ldots, b_n) \leq b_i$ in total.
- $\boldsymbol{r} : \mathbb{R}^n_{\geq 0} \to \mathbb{R}^n_{\geq 0}$ 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 players of the game (seller and bidders alike). 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.

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

The mechanism in point has been defined under the consideration of the *environment* under which decentralized auctions may be run. In particular, we assume that this mechanism will be executed through a smart contract, which shall function as the (automated) auctioneer for the item (NFT). Definition 1 arises naturally through this consideration: a smart contract is able to enforce the "removal" of a portion of the bid, guarantee the transfer of funds to the seller, as well as allocate the NFT in a perfectly secure way, so long as the underlying programming code is assumed to exactly codify the above ingredients of the triplet that is the auction mechanism. The possibilities for removing a portion of a bid include but are not limited to burning the amount in question, or saving it for later auctions or organic growth of the auctioning ecosystem, for example; the crucial characteristic of the "removal" is that the funds need to be unavailable for the players of the game, i.e., the seller and the bidders in the auction. 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.

The utilities of the participants in the auction are, then, as below:

1. For the seller, the utility is the amount that they receive:

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

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 just negative or zero, depending on the amount that they were forced to pay: $u_j(b_1, \ldots, b_n) = -p_j(b_1, \ldots, b_n)$.

Before we define the properties, we will need to define 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 $\boldsymbol{b}'$ (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). 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 in any manner. More specifically, we define the joint utility of such an OCA as the aforementioned sum, where $\boldsymbol{b}_{-S}$ are the bids of the rest of the bidders that are not part of the coalition:

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

Some times we may write $u_{\text{joint}(i,s)}$ when 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 three properties that we believe are best able to capture the dynamic interests of a decentralized, single-item NFT auction, 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:

$$\begin{aligned} u_{\text{seller}}(b_1, \ldots, b_n) &= \sum_{i=1}^{n} \left(p_i(b_1, \ldots, b_n) - r_i(b_1, \ldots, b_n) \right) \\ &\geq \sum_{i=1}^{n} \left(p_i(b_1, \ldots, b_m) - r_i(b_1, \ldots, b_m) \right) - \sum_{j=n+1}^{m} r_j(b_1, \ldots, b_m) . \end{aligned}$$

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 $\boldsymbol{b}_{-i} \in \mathbb{R}_{\geq 0}^{n-1}$), 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

$$u_i(b_1, \ldots, b_{i-1}, v_i, b_{i+1} \ldots, b_n) \geq u_i(b_1, \ldots, b_{i-1}, b_i, b_{i+1} \ldots, b_n) .$$

3. *Off-Chain Agreement resistance* (OCA-proofness): For any set of bidders $S \subseteq [n]$, bids by bidders not belonging to S: $b_{|S|+1}, \dots, b_n$ (jointly denoted as $\boldsymbol{b}_{-S} \in \mathbb{R}^{n-|S|}_{\geq 0}$), if we consider the OCA of the seller s with a set S of bidders, then this coalition of game players is not able to obtain higher joint utility through any bids they might agree to submit, i.e., for any $\boldsymbol{b'} \in \mathbb{R}^k_{\geq 0}$ (for any $k \geq 0$),

$$u_{\text{joint}(S,s)}(\boldsymbol{b'}, \boldsymbol{b}_{-S}) \leq u_{\text{joint}(S,s)}(\boldsymbol{b}_S, \boldsymbol{b}_{-S}) \,,$$

where $\boldsymbol{b}_S \in \mathbb{R}^{|S|}_{\geq 0}$ are the bids that the bidders in the OCA would submit on their own without being part of that OCA (if we additionally have the bidder IC property above, these b_i's are equal to v_i's by the above property; if not, then it would be $b_i \in \operatorname{argmax}_{b \in \mathbb{R}_{\geq 0}} u_i(b, \boldsymbol{b}_{-i}) \ \forall i \in S$).

3 Impossibility of desirable mechanism

In what follows, we prove the following impossibility result.

Theorem 1 (No desirable 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. In particular, there is no such mechanism satisfying all three desirable properties.*

As noted in the Introduction, part of this result appears to be similar to alternative procedures concurrently and independently proposed and examined by Chung and Shi (2021).

Proof. The whole proof is based upon the interesting observation which we will prove below that any mechanism according to Definition 1 satisfying bidder IC and OCA-proofness (from Definition 2) must be such as to remove all funds, i.e., $\boldsymbol{r} = \boldsymbol{p}$. Then, it is evident that no desirable mechanism can exist, because the dependence of the removed amount on the current bids is detrimental to the OCA-proofness of the mechanism: in particular, suppose that such a mechanism exists. Now, consider the OCA of all bidders with the seller: all participants of the OCA agree that only the normal winner (say $i \in [n]$) of the auction will submit a single bid of $b = 0$ to the smart contract acting as the auctioneer, and that winner will pay the same amount that a typical auction with truthful valuations would want them to pay, i.e., $p_i(v_1, \dots, v_n)$ where $v_j \ \forall j \in [n]$ are considered to be the true private valuations of the bidders, but instead now this payment will stay within the coalition (and may be distributed to the members of the OCA in such a way that every member, who would previously obtain zero, has an incentive to participate). Thus, the joint utility of the OCA will be $(v_i - p_i(v_1, \dots, v_n)) + p_i(v_1, \dots, v_n) = v_i$ whereas their utility under the auction mechanism would have been $v_i - p_i(v_1, \dots, v_n)$. Thus, if even one bidder pays amount $p_i(b_1, \dots, b_n) > 0$ then the auction mechanism would not be OCA-proof. However, it must be the case that for some bids $b_1, \dots, b_n$, some bidder pays $p_i(b_1, \dots, b_n) > 0$, because if not, then the auction is trivially not bidder IC. Hence, this is a contradiction of the two desired properties, and there is no such mechanism.

We now move on to the second part of the proof, i.e., proving that if the mechanism is required to not match exactly the full removal rule given above ($\boldsymbol{r} = \boldsymbol{p}$), then no possible mechanism can simultaneously satisfy the properties of bidder IC and OCA-proofness. First, we make the critical observation that OCA-proofness implies the classical Individual Rationality property from the traditional auction theory (non-winning bidders pay zero), by contradiction: suppose that there was a bidder $i \in [n]$ that bid b_i such that $x(b_1, \dots, b_n) \neq i$ (they were not the auction winner), then the OCA of that bidder (along with any other non-winning bidder for whom the Individual Rationality property does not hold, i.e., they pay a strictly positive amount) with the winning bidder and the seller would obtain strictly greater joint utility by not submitting (or equivalently, submitting zero) bids for the users whose payments would not satisfy the Individual Rationality property, since the joint utility will be greater by that (positive) amount. Hence, a mechanism that is not Individually Rational would necessarily not be OCA-proof, which is a contradiction. Hence, the conclusion is that the mechanism in question has to make only the winning bidder pay a potentially non-zero amount (and potentially remove some portion of that amount). This further implies that the only source of revenue for the seller is the payment (minus the removed amount) of the winning bidder.

To proceed, we now remark that the payment rule $\boldsymbol{p}$ alongside the allocation rule x are subject to Myerson's Lemma (Myerson, 1981) since they have to satisfy the property of bidder IC, which is precisely the sense of DSIC (dominant strategy incentive compatibility) used in auction theory. In particular, this has two implications: (we will refer to them as property 1 and 2, respectively, in what follows)

1. The allocation rule x is monotone, which in our notation, means that if for some bids $b_1, \dots, b_n$ it holds that $x(b_1, \dots, b_n) = i$ for some $i \in [n]$, then $\forall b'_i \geq b_i : x(b'_i, \boldsymbol{b}_{-i}) = i$ and also if for some bids $b_1, \dots, b_n$ it holds that $x(b_1, \dots, b_n) \neq i$ for some $i \in [n]$, then $\forall b'_i \leq b_i : x(b'_i, \boldsymbol{b}_{-i}) \neq i$.

2. The payment rule $p_i(b_1, \dots, b_n)$ where $i = x(b_1, \dots, b_n)$ is the unique rule that imposes the payment which is the minimum bid b such that $x(b, \boldsymbol{b}_{-i}) = i$.

The rest of the proof will proceed on the basis of a contradiction: assume that there is a mechanism such that $u_{\text{seller}}(b_1, \dots, b_n) > 0$ for some *specific* bids $b_1, \dots, b_n$ (this is equivalent to saying that there are bids that make the supposed equality $\boldsymbol{r} = \boldsymbol{p}$ not hold) and prove that this is inconsistent with the properties of OCA-proofness and the above 2 properties arising from the bidder IC. In particular, the target is, of course, to construct an OCA that would violate the OCA-proofness of the mechanism. We shall construct the simplest possible OCA: that of one bidder with the seller. Target: without the OCA, the bidder would be hopeless (cannot get the item), but with the OCA, the bidder can now magically get the item, and the coalition will marginally benefit from the OCA (i.e., the seller's utility will outbalance any loss from the individual utility of the bidder, since they would not get the item in the first place without the OCA). We will construct such an OCA. According to that, we analyze the joint utilities and reverse engineer:

- With the OCA's false, agreed-upon bid b'_i (everybody else bids their true values, due to bidder IC): $u_{\text{joint}(i,s)}(b'_i, \boldsymbol{v}_{-i}) = u_{\text{seller}}(b'_i, \boldsymbol{v}_{-i}) + (v_i - p_i(b'_i, \boldsymbol{v}_{-i}))$.
- Without the OCA, due to bidder IC, the bidder would bid their true value, insufficient to get them the item (but at the same time, they pay zero as we proved above, thus obtaining zero utility): $u_{\text{joint}(i,s)}(v_i, \boldsymbol{v}_{-i}) = u_{\text{seller}}(v_i, \boldsymbol{v}_{-i}) + (0)$.

We want to construct the OCA, i.e., make it so that $u_{\text{seller}}(b'_i, \boldsymbol{v}_{-i}) + (v_i - p_i(b'_i, \boldsymbol{v}_{-i})) > u_{\text{seller}}(v_i, \boldsymbol{v}_{-i})$. If $b'_i < v_i$ (i agreed to underbid), then i would not get the object (due to the x's monotonicity as above property 1), so i must overbid, and thus on its own, would obtain negative utility. Thus, we set its valuation to be $v_i = p_i(b'_i, v_{-i}) - \alpha$ for some $\alpha > 0$ to be discovered later, and we now only want $\alpha < u_{\text{seller}}(b'_i, \boldsymbol{v}_{-i}) - u_{\text{seller}}(v_i, v_{-i})$ for the OCA to game the mechanism.

By the assumption of the *specific* bids that make the removed amount not be equal to the whole payment, we have that $\exists b_1, \dots, b_n : u_{\text{seller}}(b_1, \dots, b_n) > 0$. Then, we claim that there exist *specific* bids $b'_1, b'_2, \dots, b'_n$ and bidder $i \in [n]$ such that $u_{\text{seller}}(b'_i, \boldsymbol{b}'_{-i}) > u_{\text{seller}}(0, b'_{i-})$. Proof by contradiction: if there were not, then for all bids b'_i and $i \in [n]$, it would be true that $u_{\text{seller}}(b'_i, \boldsymbol{0}) \leq 0 \Rightarrow u_{\text{seller}}(b'_i, \boldsymbol{0}) = 0$ and since this is for all i, $u_{\text{seller}}(b_1, \dots, b_n) = 0$ for all bids $b_1, \dots, b_n$ which is a contradiction since we assumed there are some bids that do not satisfy this.

We will now use these specific values $b'_1, b'_2, \dots, b'_n$ that we proved they exist: define the valuations of other users $\boldsymbol{v}_{-i} = \boldsymbol{b}'_{-i}$, and i's as has already been discussed (α is still free to be chosen later). The final observation needed is that the specific bid of a bidder should not matter for the determination of the removal rule $\boldsymbol{r}$ as long as they are "in the same range of item-getting" (i.e., if both a current bid and an alternative bid would either make them win or lose). Formally, we will prove through this observation that $u_{\text{seller}}(v_i, \boldsymbol{v}_{-i}) = u_{\text{seller}}(0, \boldsymbol{v}_{-i})$ (because this also holds for the payment rule, by properties 1 and 2 above), i.e., it does not matter to the seller whether a non-winning bidder bid some $v_i > 0$ or exactly zero. Then, set any α such that $0 < \alpha < u_{\text{seller}}(b'_i, \boldsymbol{b}'_{-i}) - u_{\text{seller}}(0, \boldsymbol{b}'_{-i})$ and the OCA is complete, thus our proof by contradiction is completed.

We now proceed to prove the above observation: we have that if $x(b_1, \dots, b_n) \neq i$ and $x(b'_i, \boldsymbol{b}_{-i}) \neq i$, then $r_i(b_1, \dots, b_n) = r_i(b'_i, \boldsymbol{b}_{-i}) = 0$ by Definition 1 and OCA-proofness as above. In the other case that $x(b_1, \dots, b_n) = x(b'_i, \boldsymbol{b}_{-i}) = i$, if we assume that $r(b_i, \boldsymbol{b}_{-i}) > r(b'_i, \boldsymbol{b}_{-i})$ then the mechanism would not be OCA-proof: consider the situation where all bidders have valuations $v_j = b_j$ for all $j \in [n]$ and the OCA of i with the seller: $u_{\text{joint}(i,s)}(b_i, \boldsymbol{b}_{-i}) = u_{\text{seller}}(b_i, \boldsymbol{b}_{-i}) + (b_i - p_i(b_i, \boldsymbol{b}_{-i})) = b_i - r(b_i, \boldsymbol{b}_{-i})$ but could be manipulated with the OCA because the payments remain the same: $u_{\text{joint}(i,s)}(b'_i, \boldsymbol{b}_{-i}) = u_{\text{seller}}(b'_i, \boldsymbol{b}_{-i}) + (b_i - p_i(b'_i, \boldsymbol{b}_{-i})) = b_i - r(b'_i, \boldsymbol{b}_{-i}) > b_i - r(b_i, \boldsymbol{b}_{-i}) = u_{\text{joint}(i,s)}(b_i, \boldsymbol{b}_{-i})$. This is a contradiction to the OCA-proofness of the mechanism. (the other case that $r(b_i, \boldsymbol{b}_{-i}) < r(b'_i, \boldsymbol{b}_{-i})$ also leads to the same contradiction by taking $v_i = b'_i$, but $v_j = b_j$ for $j \neq i$) □

4 Discussion of OCA-proofness and Relaxed Properties

We note that, while we have shown that is not possible to devise a good protocol that will be OCA-proof and at the same time bidder incentive compatible, the notion of OCA-proofness is a stronger one than what we need; having an OCA-proof protocol would be very reassuring, but a non-OCA-proof protocol could also maybe be safe for practical purposes. For example, the type of OCA used for our impossibility proof above requires some bidders to collude with the seller. However, the problem we might imagine we would like to deal with is the one where the bidders do not trust the seller, and hence they should be reluctant to jointly take any collusive action with the seller. Moreover, in the case that the seller and the bidders are all trusted, there will be no need to the proposed protocol below, and a much more straightforward second-price auction could be implemented. Therefore, it makes sense to discuss other possible guarantees.

As discussed, we cannot promise a second-price auction that is seller IC, bidder IC, and OCA-proof without any additional assumptions. Instead, we give a relaxed definition:

Definition 3 (Equilibrium-truthful auction)**.** An *equilibrium-truthful auction* is an auction in which:

1. The seller's dominant strategy, assuming bidders bid truthfully (i.e. $b_i = v_i$), is to not post any fake bid

2. Each bidder's dominant strategy is to bid truthfully, assuming the seller is not posting any fake bids.

Further, we require that running such an auction will make economic sense for the seller. For this, we define the following auction property:

Definition 4 (Asymptotically second-price auction)**.** Given an auction in which at least n participant are known to participate, and have their private value independently drawn from the same distribution D, we say that an auction is *asymptotically second-price* when the expected reward of the seller is $(1 - o(1)) \cdot \mathbb{E}[B_2]$, assuming bidders bid truthfully and that there are no fake bids. Here B_2 is the second highest price, and $o(1)$ goes to 0 as $n \to \infty$.

We will show that under some assumptions on the distribution from which the valuations v_i are drawn, there is an equilibrium-truthful asymptotically second-price auction that can be executed on a smart contract (see Theorem 2).

5 Proposed Protocol

In this section we propose a protocol that will satisfy equilibrium-truthfulness. We start with basic assumptions, and then provide an initial protocol, which will be revised a few times to overcome the difficulties that emerge. This shows the motivation and necessity of the components involved in the final version of the protocol, presented at the end of this section.

5.1 Relaxed Assumptions and Notations

We relax the assumptions that lead to the optimality of Vickrey auctions. We aim to devise a contract for an auction with the following assumptions:

1. The seller has a minimal value v at which he's willing to sell his NFT.

2. Bidder i places a single bid b_i, and has a private value v_i for the NFT. It should be noted that this is not a strong assumption, as it can be seen that there is no reason for any bidder to place more than one bid in our protocol. This assumption is mainly in place to simplify notations.

3. The seller may place multiple *fake bids* in the auction, if they wish to do so.

4. The utility of bidder i is $u_i = v_i - p_i$ if i is the winning bidder, otherwise $u_i = 0$.

5. There is some continuous distribution D from which all v_i are randomly drawn independently. We denote by f the probability density function of D, and by F the cumulative distribution function.

6. Bidders are not colluding. Thus, each b_i is a function of only v_i and D.

Our main result is the following:

Theorem 2. *Under the assumptions given in Section 5.1, if* $\sup_{s \in support(f)} \frac{1-F(s)}{f(s)} < \infty$*, then there is a protocol (described further in this section) that is equilibrium-truthful and asymptotically second-price.*

5.2 Basic idea

We would like all parties to commit to their bid ahead of time on the blockchain, without revealing it. This way, the auctioneer won't be able to inspect the maximum after the auction is done and insert a new bid that is only slightly lower than the maximum. We would also like to have all the parties aware of all the commits. So the protocol is as follows:

1. Commit phase: The auctioneer sets up a contract, to which bidders should should send commits (i.e. $hash(bid||nonce)$), and the auctioneer sends the NFT to the contract. A minimum bidding value m is also advertised.
2. With predefined conditions, e.g. enough time and blocks have passed, the commit phase ends.
3. Verification phase: Now all parties send to the contract their actual bids (and nonce for verification).
4. At the end of the verification phase, the contract knows who the highest bidder is, and who the second highest bidder is.
5. The highest bidder is the only party allowed to send to the contract the money, which must be at least the second highest bid, in which case it triggers a transaction of the NFT from the contract to the winner, and allows the auctioneer to withdraw the money received in the contract.

5.3 Immediate problems

While the proposed protocol indeed prohibits the seller from inserting fake bids adaptively by inspecting other bids, we can spot several problems with the protocol:

- Funds are not guaranteed. That is, when bidders bid, they are not required to lock their bid amount to ensure funding in case they win. This problem seems to be a fundamental issue with sealed bids, since by locking any amount the bidder reveals information (in the form of an upper bound) on their bid. Furthermore, any funds locking for this purpose will be difficult to verify on the blockchain, requiring additional zero-knowledge proofs for the bid not to be revealed. We revisit this issue in Section 5.9.

  We also note that it might be a feature of the protocol and not a shortcoming; a bidder might want to bid high on some NFT, but not want to liquidate the funds for it unless he wins. It is still necessary, however, to economically discourage the winner from not buying the NFT.

- Bids are revealed after the auction. That is, in the verification phase, all bidders must reveal their bid to everyone. This might be undesirable, as the only information really required to be revealed is the identity of the highest bidder, and the second highest price.

- As we stated the protocol, it is not clear what happens if some bidder does not reveal their bid in the verification phase. We will address this issue in the next proposed protocol. Allowing revelations to be withheld by the bidders is a problem which introduces a different attack on the protocol by the seller. The attack is as follows: the seller can submit multiple bids in a range of values, and at the end of the verification phase, but before it is finished, reveal only the bid that is closest to the highest bid from below. This then allows the seller to receive a potentially higher price for the NFT than they would normally get in the second price auction. We note that this is a problem which cannot be solved by simply blocking the seller from submitting multiple bids, since the seller could generate arbitrarily-many private/public key pairs (sybils) to achieve the same attack.

5.4 Incentivizing Reveals

To avoid an attack by the seller of committing to many bids and only revealing the one that benefits them, we demand that all commits are revealed. However, we do not want to have the auction get stuck if one party does not reveal their commit, as that introduces possible DoS attacks (from malicious parties as well as a cheating seller who is not happy with the result after seeing the revealed bids). To that end, we require a relatively high amount (to be decided) of money, L, to be locked up in the contract by any commit. This money will be returned to the bidder after revealing their bid, and will be burned otherwise.

The locked up amount L will also not be returned to the winner if they do not transfer the funds. To discourage the winning bidder from not following up with their offer, we can also decide on the following. First, denote by b_2 the second-highest bid revealed. Then $\min\{L, b_2\}$ will be transferred to the seller in any case. Then, if this value was equal to b_2, the NFT is automatically freed for the winner, and otherwise the winner should still send an additional $b_2 - L$ to the contract before the NFT is freed.

5.5 Further considerations

While the revised protocol solves part of the problems previously discussed, some new problems arise.

- This protocol is prone to a different attack by the seller: have a high initial bid, that is higher than the expected highest bid. Then, after the verification phase when all bids are revealed, the bidder learns the distribution of the bids and and can make one more auction to use it. It is unclear whether this is an attack that is usually concerning, but this might be a problem in light of the revelations of all the bids in the verification phase.
- It is unclear how the value of L should be chosen. The optimal value will probably depend on further assumptions. This issue is analyzed in depth in the analysis section.
- Funds are not guaranteed if L is too low. On the other hand, parties will be discouraged from bidding if L is too large.
- Bids are still revealed by the end of the auction.

5.6 Adding a fee function

We want to discourage the seller from selling the NFT to themselves, through possible sybils (though the attack has nothing to do with sybils, we emphasize by that that it's not enough to block the seller from bidding). We therefore also burn some of the money transferred from the winner to the auctioneer, with some predefined *fee* function $g(\text{second highest price})$.

Remark: we take the money from the seller and not from the bidder, to make the protocol still behave like a second-price auction in the perspective of the bidders, in order to enjoy the optimality guarantees of Vickrey auctions.

For completeness, we detail the protocol:

1. Commit phase: The seller sets up a contract, to which bidders should should send commits (i.e. $hash(bid||nonce)$), and the seller sends the NFT to the contract. A minimum bidding value m is also advertised.
2. Bidders send their commits to the contract, together with a sum of money L (in tokens) to be locked up. Without both of these ingredients, the bid is not taken into account.
3. With predefined conditions, e.g. enough time and blocks have passed, the commit phase ends.
4. Verification phase: Now all parties send to the contract their actual bid amounts (and nonce for verification). 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 burned. The contract can now look for the highest revealed bid b_1 and second-highest revealed bid b_2.

6. $\min\{b_2, L\}$ 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 (as usual, the parties will need to initiate a transaction to the contract to receive their share). Otherwise, if this amount was equal to L, then an additional $b_2 - L$ amount should be sent from the winner to the contract before the NFT is unlocked. The seller will only be able to 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 pay the $L < b_2$ they have locked up, but do not receive the NFT.

5.7 New problems

We are now burning potentially a lot of money that would ideally be sent to the seller. This discourages the seller from using this type of auction. In fact, as we will see in the analysis, we need to require only a small portion of the winning bid to be burnt, assuming nice enough properties of the distribution of bids and that there is a large enough amount of bidders.

Additionally, it is unclear how we should decide on the fee function g. Some suggestions are presented in the analysis section.

5.8 Using Secure Multi-Party Computation

Arguably a better remedy for the last proposed attack by the seller is to use a zero-knowledge proof of the two highest bidders. For example, we could use a protocol such as that proposed in Sheikh and Mishra (2010) for secure multi-party computation (MPC) of the maximum, to find the highest bidder, then with another round find the second highest bidder (in which the highest bidder does not participate). This way, the only new information revealed to all other parties is the highest and second highest bids. The problem is, however, that too many messages would need to be sent back and forth, rendering this solution too impractical.

However, the protocol described by Sheikh and Mishra (2010) is not exactly good enough for our purposes, as it assumes a trusted third party, which would be undesirable for our setting. It may still be possible to use ideas for a general secure multi-party computation, such as those referenced by Goldreich (2002).

Using these ideas, for example by making a secure computation of the identity of the two highest bids before revealing any bid amounts, will potentially alleviate the need for winner payments burning, because the described attack by the seller to gain information about the bids will only reveal minimal information, and this will also solve the problem of the sealed bids ending up being revealed by all parties.

The information that would still leak in the most optimistic scenario, however, is the second highest bid, as the winner has to pay that amount. So the seller might still post a particularly large bid to learn the highest bid among the honest participants, then take advantage of it in a second auction, but this would be all the information he has.

5.9 Validating Bid Amounts

One problem with the described protocol is that committed bid amounts are not validated (i.e. the protocol does not check that a bidder actually has the funds to pay her bid). A malicious (and extremely wealthy) adversary A could perform the following attack on an auction:

1. In the commit phase, A deposits L, and commits to an extremely large bid value, X.

2. In the reveal phase, A reveals X, and wins the auction.

3. A *does not* pay the remaining $X - L$ to the smart contract.

The adversary A above would lose her deposit L every time she carried out such an attack. One could very reasonably argue that this type of attack is not a concern, as the locked up amount L sufficiently disincentivizes it. However, it is natural to ask whether it is possible to achieve a commit scheme in which committed values can be validated by the smart contract, still while revealing no information to the other bidders or the auctioneer.

5.9.1 Submarine Bids

Breidenbach et al. (2017) introduced "Submarine sends" in 2017. A Submarine send is a way for an Ethereum user to lock up funds in a fresh account (with a fresh address), and give a smart contract access to these funds. We can use the Submarine send as a subroutine for our protocol to enforce that all bids are backed by funds.

Let c denote the address of the smart contract orchestrating the auction. Let `Forwarder` denote the Ethereum smart contract which simply consists of sending all one's funds to address c.

1. To commit to a bid value b_i, bidder i sends b_i ETH to address $\hat{a} = hash(c||\text{nonce}||\text{byteCode}(\texttt{Forwarder}))$.

2. To reveal a bid value b_i, bidder i sends her b_i value and nonce to c. c then instantiates the `Forwarder` contract on address $\hat{a}$ using Ethereum's `CREATE2` operation (Buterin, 2018), causing the funds from $\hat{a}$ to be transferred to c.

3. c then verifies that the value from $\hat{a}$ is indeed equal to b_i. If not, the funds are not eligible for refund. Otherwise, bidder i is eligible to claim a refund of $b_i - p_i$ (where p_i is the amount i must pay given the results of the auction).

With the above commitment scheme, bids must be deposited ahead of time to the submarine addresses. Upon revealing, if any revealed values don't match the commit values, the deposited funds will be lost forever, as only the auction contract c has access to them (the bidder cannot feasibly compute the private key of $\hat{a}$, and it is enforced that $\hat{a}$ has start-up code to send all of its funds to c).

The secrecy of the commit operation using Submarine sends relies on the fact that transactions to Submarine address are indistinguishable from any other transactions on the blockchain. However, at the time of an auction, it may be possible for a savvy bidder to identify transactions which are likely to be relevant to the auction if, for instance, he has a good prior estimate of the bid distribution.

6 Protocol Analysis

The main purpose of this section is to give a proof of Theorem 2. We prove this in steps, by showing that the described attacks do no longer work, in a suitable sense.

6.1 Multiple Bids by the Seller

We analyze the attack of multiple bids in the model without fees. Suppose that B_1 and B_2 are the highest and second-highest bids, respectively. Suppose the seller tries the attack of submitting multiple bids. We will analyze under which conditions we can be sure this attack is not worthwhile for the seller.

To this end, we model the attack in a way that is favorable for the seller: the seller chooses a set S of bids (potentially by sybils), and a moment before the verification phase ends, the seller inspects all revealed bids and chooses which of the bids in S to reveal. It is favorable for the seller in the sense that the seller is assumed to inspect the complete information of the verification phase before taking action.

Observe that the seller does not lose anything by revealing all commits that are lower than B_1, and in fact will gain by getting to keep the locked up funds used for those bids. Assuming the seller does not aim for a second auction and wants to sell the NFT in the current auction, 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).

Assuming the utility function of the seller is linear in their amount of money, the seller therefore attempts to maximize:

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

where the expectation is over the values of B_1 and B_2. We note that by linearity of expectations, we can rewrite $\mathbb{E}|\{s \in S : s > B_1\}|$ as $\sum_{s \in S} \mathbf{Pr}(B_1 < s)$.

Theorem 3. *If there exists a set S for the seller with a positive expected reward, then there also exists a set of size 1 with a positive expected reward.*

Proof. Suppose there is a set for which there's a positive expected reward. Then there is a set with minimal size among all such sets. This set is nonempty, since for the empty set the expected reward 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 assumption, if s_1 is taken out, the reward becomes non-positive, and in particular decreases. On the other hand, the expected reward increases by $L\,\mathbf{Pr}(s_1 > B_1)$ and decreases only when $B_2 < s_1 < B_1$ by: $\mathbb{E}[s_1 - B_2 | s_2 < B_2 < s_1 < B_1]\,\mathbf{Pr}(s_2 < B_2 < s_1 < B_1) + (s_1 - s_2)\,\mathbf{Pr}(B_2 < s_2 < s_1 < B_1)$ (where we used the law of total expectation).

The claim that the set $\{s_1\}$ has a positive reward is equivalent to the claim that $L\,\mathbf{Pr}(s_1 > B_1) < \mathbb{E}[s_1 - B_2 | B_2 < s_1 < B_1]\,\mathbf{Pr}(B_2 < s_1 < B_1)$, so the claim will be follow if we show that

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

is at least

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

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

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

which is larger because $s_1 - B_2 > s_1 - s_2$. □

Therefore, in order to ensure there is no favorable attack for the seller, we need only ensure that there is no favorable singleton set of commits it can do. That is:

Corollary 1. *There exists a favorable attack for the seller if and only if there exists a value s such that*

$$\mathbb{E}[s - B_2 | B_2 < s < B_1]\,\mathbf{Pr}(B_2 < s < B_1) > L\,\mathbf{Pr}(s > B_1)$$

Thus, to make the attack unfavorable, the following condition is enough:

Corollary 2. *There is no favorable attack for the seller if and only if*

$$L \geq \max_s \frac{\mathbb{E}[s - B_2 | B_2 < s < B_1]\,\mathbf{Pr}(B_2 < s < B_1)}{\mathbf{Pr}(s > B_1)}$$

As an example, we estimate this value in a simple model of the private values of the n bidders. First, we model their values as being drawn independently at random from a common distribution D with a probability density function $f(x)$ and cumulative distribution function $F(x) = \int_{-\infty}^{x} f(t)dt$. Then

$$\mathbf{Pr}(s > B_1) = \mathbf{Pr}(\forall i : b_i < s) = F(s)^n$$

Further, by symmetry,

$$\mathbf{Pr}(B_2 < s < B_1) = n \cdot \mathbf{Pr}(B_2 < s < B_1 \wedge \text{bidder 1 is the winner}) = nF(s)^{n-1}(1 - F(s))$$

And also, assuming the bids are nonnegative and using the fact that $\mathbb{E}[X] = \int_0^\infty \mathbf{Pr}(X > x)dx$ for any nonnegative random variable X

$$\mathbb{E}(B_2 | B_2 < s < B_1) = \int_0^s \mathbf{Pr}(B_2 > x | B_2 < s < B_1)dx = \frac{1}{\mathbf{Pr}(B_2 < s < B_1)} \int_0^s \mathbf{Pr}(x < B_2 < s < B_1)dx$$

$$= \frac{1}{\mathbf{Pr}(B_2 < s < B_1)} \int_0^s (\mathbf{Pr}(B_2 < s < B_1) - \mathbf{Pr}(B_2 < x < B_1))dx$$

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

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

Plugging all together to the bound on L, noting that in the case of no favorable attack by the seller there will be no deviation from the optimality and truthfulness of second-price auctions, we obtain the simplified:

Corollary 3. *When all bidders draw their private value independently at random from a common distribution with a cumulative probability function F, the auction is equilibrium-truthful exactly when*

$$L \geq \max_s \frac{1}{F(s)^n} \int_0^s F(x)^{n-1}(1-F(x))dx \,,$$

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

Example 1. Suppose the common distribution is the uniform distribution [0,1]. For any $s > 1$ the lower bound obtained will be the same as for $s = 1$, so we will restrict the calculations to $s \in [0,1]$. Here, $F(x) = x$, and so the integral evaluates to $\int_0^s (x^{n-1} - x^n)dx = \frac{s^n}{n} - \frac{s^{n+1}}{n+1}$, so we need $L \geq \max_{s\in[0,1]} \left(\frac{1}{n} - \frac{s}{n+1}\right) = \frac{1}{n}$.

For a general uniform distribution $[a,b]$, it can be seen by linearity that the condition becomes $L \geq \frac{b-a}{n}$

Example 2. It can be argued that bids follow a more heavy-tailed distribution. So as an additional example, we take the exponential distribution $f_X(x) = \lambda e^{-\lambda x}$ for $x > 0$ (and 0 otherwise). Here $F(x) = 1 - e^{-\lambda x}$, and so

$$\int_0^s F(x)^{n-1}(1-F(x))dx = \int_0^s (1-e^{-\lambda x})^{n-1} e^{-\lambda x} dx$$

Making a change of variables, $e^{-\lambda x} = t$ we get

$$= \frac{1}{\lambda} \int_{e^{-\lambda s}}^1 (1-t)^{n-1}dt = \frac{1}{\lambda} \int_{e^{-\lambda s}}^1 (1-t)^{n-1}dt = \frac{1}{n\lambda}(1-t)^n \Big|_1^{e^{-\lambda s}} = \frac{1}{n\lambda}(1-e^{-\lambda s})^n = \frac{1}{n\lambda} F(s)^n$$

Dividing by $F(s)^n$ we finally get the lower bound $L \geq \frac{1}{n\lambda}$, (this turned out to be independent of s, and is therefore also the maximum lower bound needed).

In fact, the examples we've seen are special cases of the following theorem:

Theorem 4. *Suppose all bidders draw their private value independently at random from a common distribution on values in $[a,\infty)$ for some $a > 0$, with a cumulative probability function F and density function f. Also suppose 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 $L(n) \sim \frac{1}{n} \cdot \sup_s \frac{1-F(s)}{f(s)}$ as $n \to \infty$.*

We note that this is exactly what we obtained in the two examples: in the uniform distribution, $\sup_s \frac{1-F(s)}{f(s)} = \sup_s \frac{1-s/(b-a)}{1/(b-a)} = b-a$, and in the exponential distribution, $\sup_s \frac{1-F(s)}{f(s)} = \sup_s \frac{e^{-\lambda s}}{\lambda e^{-\lambda s}} = \frac{1}{\lambda}$.

Proof. We sketch the proof of this theorem, in order to not get too technical. It should be stated that we might need additional "niceness" properties on the functions f and F, but we are not trying to be as precise as possible with this theorem.

We first fix a value of s. Assuming some "niceness" properties as alluded to above, we will assume we can switch between the limit of $n \to \infty$ and the maximum over s. Hence, we first find the limit for any s, then take the maximum over the values of s.

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

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

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

$$\int_0^s F(x)^{n-1}(1-F(x))dx \approx \frac{1-F(s)}{f(s)} \int_0^{F(s)} t^{n-1}dt = \frac{1-F(s)}{nf(s)}F(s)^n$$

Thus, dividing by $F(s)^n$ we finally have the limit per s be $L(n,s) = \frac{1-F(s)}{nf(s)}$. Taking the maximum over s now proves the result. □

Thus, we see that the value of L will always dependent on the number of anticipated participants (assuming the "niceness" properties of the distribution, which are arguably expected) in a way the reduces the amount required to be locked up in order to avoid the several commits attack.

6.2 Repeated Auction

Suppose the highest bid is B_1 and second-highest bid is B_2, unknown to the seller. We will analyze the conditions that make the two-auction attack by the seller worthwhile.

We make the assumption that the bids B_1 and B_2 will stay exactly the same in the event of a second auction, analyzing the amount the seller loses.

We note that model is possibly in favor of the seller, because in the second phase B_2 might not bother submitting, but it might also be no in his favor, because other bidders might want to bid higher than B_1 in this situation.

First, we recall the attack model: the seller posts a very high bid ∞ from a sybil, learns the bids in the verification phase while selling the item to himself, then posts a second auction with the same parameters.

In the first auction, it pays B_1 but receives $B_1 - g(B_1)$ where f is the fee function, so in total he loses $g(B_1)$. Had the seller not done this attack, the total gain would be $B_2 - g(B_2)$.

In the second auction, the seller posts $B_1 - \varepsilon$ as a bid, thus selling the item for a price of essentially B_1, receiving $B_1 - g(B_1)$ 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 better off with this attack if and only if

$$B_2 - g(B_2) < B_1 - 2g(B_1)$$

Or equivalently

$$B_1 - B_2 > 2g(B_1) - g(B_2)$$

Thus, if we find an f such that $B_1 - B_2 \leq 2g(B_1) - g(B_2)$ (at least in expectation), the attacker would be discouraged from proceeding with such an attack. Assuming a nondecreasing fee function, we have $g(B_1) \geq g(B_2)$, and 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.

There are models in which this does not necessarily leads to high fees: in the example model of n uniformly random private values in $[0,1]$, we have $\mathbb{E}B_1 - \mathbb{E}B_2 = \frac{n}{n+1} - \frac{n-1}{n+1} = \frac{1}{n+1}$ (we used here the well known expected values of the order statistics of the uniform distribution. They can also be easily computed with the CDF functions computed in the analysis of the previous attack, of multiple bids). Thus, it is enough to take a constant $g(x) = \frac{1}{n+1}$ where n is the number of anticipated participants. For a more general uniform distribution on $[a,b]$ it follows by linearity that the bound in expectation is $\frac{b-a}{n+1} \leq g(B_1)$. If the seller assumes an interval of $[m, M]$ where m is the minimal accepted bid and M is unknown, it can estimate M from the highest bid as $B_1 \approx M$, making it reasonable to choose

$$g(x) \approx \frac{x-m}{n+1}$$

Thus, for a large number of expected participants (which will need to be posted in the auction contract, but participants have the choice of whether or not to trust the auction with the posted parameters), the seller will lose only a small fraction of the highest bid, while maintaining the truthfulness of the auction, potentially inviting more participants and increasing the reward.

As an additional example, when the distribution is exponential with parameter λ, it can also be computed that $\mathbb{E}[B_1 - B_2] = \frac{1}{\lambda}$, and here we see no diminishing fees for large n if we choose a constant fee function. However, the expected value of B_2 is $\sim \frac{\log n}{\lambda}$, and hence we can take $g(x) = \frac{x}{\log n}$.

Given that using these fees potentially makes the auction more trusted, it might indeed be beneficial for the seller to set up an auction with our proposed protocol, and if a large number of participants is anticipated, the expected loss for the seller is small.

For completeness we state and prove the relevant theorem:

Theorem 5. *Suppose the distribution D satisfies*

$$\sup_{x \in support(f)} \frac{1 - F(x)}{f(x)} < \infty \,,$$

then the fee function $g(x)$ can be chosen to be linear, $g(x) = \alpha x$, where α depends only on n and on the distribution D, such that the repeated auction attack will have a negative value for the seller and $\lim_{n\to\infty} \alpha(n) = 0$.

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

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

Let $A = \sup_x \frac{1-F(x)}{f(x)}$, which is finite by assumption. Then we finally get:

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

We will now show that we can always choose

$$\alpha = 2\frac{\mathbb{E}[B_1 - B_2]}{\mathbb{E}[B_2]}$$

We evidently have $\mathbb{E}[B_1 - B_2] < \mathbb{E}[g(B_2)]$. To show that $\lim_{n\to\infty} \alpha = 0$, we divide to 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've shown that the difference $\mathbb{E}[B_1] - \mathbb{E}[B_2]$ is bounded. Therefore $\alpha \leq 2\frac{A}{\mathbb{E}[B_2]} \to 0$.
- $\lim_{n\to\infty} \mathbb{E}[B_1] < \infty$, and is equal to 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 we'll show below), hence the numerator approaches 0 and the denominator is nondecreasing, giving $\lim_{n\to\infty} \alpha = 0$.

  It remains to prove that $\lim_{n\to\infty} \mathbb{E}[B_2] = M$. Since $B_2 \leq B_1$, it's enough to prove that $\liminf \mathbb{E}[B_2] \geq M$. 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 $\forall\varepsilon > 0$ we have $\lim_{n\to\infty} \mathbf{Pr}(B_1 < M - \varepsilon) = 0$, it would follow that $\liminf \mathbb{E}[B_2(n)] \geq M - \varepsilon$, for any $\varepsilon > 0$, so $\liminf \mathbb{E}[B_2(n)] \geq M$ and we will be done.

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

This completes the proof. □

6.3 Proof of Theorem 2

Proof. We prove that the protocol described with choosing the funds requirements L as in Theorem 4 and the fee function g as in Theorem 5, satisfies the theorem.

We first show that the auction described is equillibrium-truthful.

Indeed, suppose the seller does not place any fake bids. Then, since the amount each bidder pays in case they win is the second highest price, we have the usual assumptions for the optimality of second-price auction, and it follows that the best strategy of any bidder is to place $b_i = v_i$. On the other hand, assuming all bids are $b_i = v_i$, the seller has no better strategy than to place no fake bids, as is shown in Theorem 4 and Theorem 5.

Theorem 5 now also shows that the auction is asymptotically second-price, since the reward for the seller is $(1-\alpha)\mathbb{E}[B_2]$ and $\alpha = o(1)$ (with the notations of Theorem 5). □

As already shown, examples of distributions for which the assumption holds is any uniform distribution on a bounded interval, and the exponential distribution.

7 NFT Launches

While designing single-item auctions is one part of the design space, another key application of blockchains' auction mechanisms is designing protocols for NFT launches. These differ in that a distributor has multiple items they are selling. The design space for NFT launches is rich with many different considerations. (While we are discussing NFT drops in particular, much of this applies generally to selling any scarce goods on a blockchain.) Some properties we would like are as follows:

1. Fairness: NFT launches often incorporate a form of randomness where a buyer isn't sure of the exact NFT they have received until finalizing a purchase. This randomness should not be able to be manipulated by buyers with advanced technical knowledge or large amounts of capital; every bidder should have an equal chance of getting NFTs of different rarity. We also include sybil-resistance as contributing to a launch's fairness where bidders shouldn't be able to make multiple false identities to manipulate the mechanism.

2. No-Race Condition: Often times NFT launch can devolve into who can get their transaction on the blockchain first. This can cause the launch to effectively become a first price gas auction. We do not want our launch to reduce to a secondary auction apart from the primary mechanism.

3. Ease of Use: To make participation accessible to everyone we want it to be easy to participate in a NFT launch. This includes keeping the bidding language simple, not gaining advantages from advanced technical knowledge, and having dominant strategies for the bidders.

4. Trustlessness: Bidders should not have to trust that the distributor running the launch will not manipulate the mechanism. The mechanism should be designed so bidders can verify the launch is being executed correctly, and the distributor should find it in their best interest to follow the mechanism.

5. Robustness: The launch should perform well for the distributor even if the distributor does not have good information on the bidder demand curves. For instance the launch should be able to adapt if lots of inventory is going unsold and likewise should adapt if demand is far outstripping supply. The launch should end with the distributor netting revenue relatively close to what they could have netted if they knew the bidder's demand exactly. This is more relevant in a NFT launch because this is the first time these items are sold, and so there is no explicit market data to help price the items.

7.1 Model

For the sake of analyzing different NFT launch formats, we define a concrete model to work with. We assume the launch is looking to distribute some set of m items M where each item in M has an associated rarity score. Then there are a set of n bidders N where each bidder has some value for each of the items in M. In order to sell these items, a distributor runs a mechanism that solicits bids from bidders and uses those bids to determine an allocation of the items. To keep the bidding language simple, we assume that bids can at most include a single price and quantity. The distributor also has some payment rule that takes all the bids and determines how much each bidder should pay in the end. In our model, we allow for it to be possible for a bidder to not receive any NFTs and still have to pay. We do not consider how NFTs are distributed after the launch and assume there is some mechanism in place for this transfer.

We restrict ourselves to mechanisms that happen entirely on-chain. Regarding trustlessness, this gives the benefit of reducing the amount of ways a distributor can manipulate an auction as the allocation rules/payment rules are specified in the smart contract. This also makes it so distributors can't secretly censor certain bids. Having everything happen on-chain does introduce the difficulty of users having to pay gas costs on top of their bids and limits the calculations the smart contract itself can do. For the purposes of this analysis, we assume these gas costs are non-negligible and seek to minimize them when possible.

We also make the distinction between continuous and sequential bidding. In continuous bidding, bids are submitted and cleared as they arrive. Whereas in sequential bidding, there are distinct bidding and clearing phases. For the purposes of running the launch, we assume there is a shared global clock between the distributor and all the bidders. Examples uses of this clock are to signal the end of a bidding phase and the start of a clearing phase, or a mechanism making the current price for minting as a function of the time.

Finally we note, assuming a launch is universally fair (not a strong assumption as discussed below), bidders are effectively competing for copies of a token that has some distribution of possible values it can take on. Thus we assume the bidder has a single value for winning the launch equal to their expected value of this token. We use this modelling assumption to reason about bidders having dominant strategies.

With this general model in place, some of the desirable properties we want come as a byproduct. For the remaining properties, we formalize the following definitions. In this report we do not formally prove any of these definitions hold but rather use them as ways to informally reason about different launch mechanisms and differentiate them.

Definition 5 (Universal Fairness)**.** *L is universally fair if for any two bidders i, j who bid for the same quantity of goods, $\mathbf{Pr}[i \text{ receives item of rarity } r | i \text{ wins}] = \mathbf{Pr}[j \text{ receives item of rarity } r | j \text{ wins}]$, regardless the prices i or j bid at.*

Definition 6 (Sybil Resistant)**.** *L is sybil resistant if a bidder i can not use sybils to increase their chance of winning the launch or receiving a higher rarity good.*

We consider sybils as two or more distinct bids made by the same bidder.

Definition 7 (Δ-Race Free)**.** *L is Δ-race free if for times $t, t' \in [0, \Delta]$, entering a bid at time $t < t'$ is not a strongly dominant strategy for every possible bidder.*

For this analysis, we assume Δ is sufficiently large (e.g. 24 hrs), and so we drop the Δ and simply talk about "race freeness". Note that depending on a bidder's valuation, it might be a dominant strategy for them to enter the auction as early as possible (e.g. in a descending price auction where a bidder's value is higher than the starting price). But for the race free condition to hold, it suffices that there exists at least one valuation where it would benefit a bidder to wait. Trivially any sequential bidding mechanism is race free, since the order of bids made during the bidding phase is inconsequential on the final allocations. We will often say a mechanism that avoids 'race conditions' to imply it is race free.

Definition 8 (Trustless)**.** *L is trustless if the distributor is strongly worse off by submitting fake bids.*

Definition 9 (Incentive Compatible)**.** *L is incentive compatible if bidders have a unique dominant strategy for participating in the launch.*

Definition 10 (α-Robust Revenue)**.** Let R be the expected maximum revenue a distributor could receive if they had complete knowledge of the distributions bidders were drawing their bids from. Then L is *α-revenue robust* if the expected revenue L generates is at least αR.

Definition 11 (α-Robust Welfare)**.** Let W be the expected maximum social welfare a launch could generate. Then L is *α-welfare robust* if the expected welfare L generates is at least αW. Welfare here is defined by the sum of the valuations of winning bidders.

The exact bounds on α that different mechanisms can guarantee is beyond the scope of this report. Instead we informally differentiate between mechanisms where α can be arbitrarily small and mechanisms where α should be relatively large for both revenue and welfare. Thus, for the analysis, we drop the α and simply speak towards whether or not mechanisms are revenue robust and welfare robust.

For designing mechanisms that are universally fair and sybil resistant, there are solutions that can generally be plugged into any launch mechanism that were proposed by Hasu and Agnihotri (2021) and Buterin (2021a) respectively. We outline their proposals below.

7.1.1 Satisfying Fairness through Post Launch Randomness

One of the desirable properties an NFT auction should satisfy is that winners of the launch should not be able to increase the chance they get a particularly rare NFT. Hasu and Agnihotri (2021) outline many cases where NFT launches did not follow this and contained exploits that let sophisticated users have an advantage over normal users by sniping rare NFTs. However, that same article also suggests a method for remedying this problem in any launch. The solution is to simply distribute tokens representing that a bidder has won an NFT, but to have those tokens contain no information about the NFT which it represents. Then after the launch has ended, some sort of verifiable randomness (such as Chainlink VRF) can be used to assign different token IDs to different NFTs according to the preset distribution. Since it is impossible to know the random assignment before it is computed, by implementing this step after the launch ends, there is no way for bidders to manipulate the rarity of which NFT they are given during the launch. This strategy can effectively be added to any launch mechanism to satisfy universal fairness.

7.1.2 Satisfying Sybil Resistance through Proof of Personhood

There are various ways in which an NFT drop can satisfy sybil-resistance. The most straightforward way to achieve this property (and also advocated for in Buterin (2021a)) is to enforce that each bidder is a human-being, rather than a sybil/bot. While it seems that this is an unreasonable task to perform on a blockchain, there have been developments over the past year of “proof of personhood” protocols which solve this problem. For example, Proof of Humanity (James, 2021) is a protocol which creates an on-chain registry of identities, each of which is linked to a unique human identity. Each identity is required to link to a web-profile, and is also required to be vouched for by a number of other verified human identities.

An NFT launch which seeks to achieve sybil-resistance can simply enforce that every bidder address is registered in the registry of human identities. An obvious concern with such a mechanism is that it violates privacy. To get around this, one could prove their human identity using a zero knowledge proof (Buterin, 2021b).

7.2 Current NFT Launches

With this framework, we analyze the following popular on-chain designs for NFT launches.

7.2.1 First Come First Serve

First come first serve (FCFS) is the simplest and most common way for a distributor to run a NFT launch. The distributor simply has to pick a price to mint the NFTs. Then the first m bidders who meet that price win an NFT. To satisfy universal fairness and sybil resistance, the distributor can use post launch randomness and proof of personhood.

While this design is simple, it has clear flaws. Firstly, FCFS is not race free. Generally, the price in FCFS mechanisms is set below the true market price to increase participation in the launch. Thus there is a race for all the bidders to be the first ones to guarantee they will receive an NFT. This demand ends up spilling into the gas auctions where bidders want to guarantee their transaction to mint an NFT will be included in the next block. This effectively reduces the primary FCFS mechanism to a first price gas auction at the consensus layer. This race condition also increases the difficulty of pariticipating in the auction. Technologically advanced bidders gain an advantage in how quickly they can post their bids.

On the other hand, FCFS mechanisms are easy to use in that the strategy for how to bid as a bidder is clear. If you think the price the distributor set is fair, submit a bid at that price. Additionally FCFS is trustless in that there is no way for the distributor to manipulate the mechanism. There is simply a posted price and if a transaction clears that price, it is given an NFT. There is no way for the distributor to affect the payments since they are already set.

Finally, FCFS is not revenue robust. If the distributor has no priors on the values of bidders, they have no way to determine what fixed price they should set. If the fixed price is above the bidder's valuations, then the distributor will be left with unpurchased supply. If the fixed price is too low, the distributor will have missed out on potential revenue. FCFS fairs a bit better in welfare robustness as the bidders with the highest values are the ones that will put the most effort into getting their bids posted first.

7.2.2 Batched Raffle

Batched raffles use a lottery format where bidders buy raffle tickets and random raffle tickets are chosen to be winners. This is the mechanism used by FairDrop (0xEssential, 2021). The mechanism has a long bidding period (24-48 hrs) where bidders can purchase a raffle ticket. The price for this ticket is determined by the distributor, and ETH corresponding to the price must be sent along with the bid as a security deposit. After the bidding period is over, winners (whose number is equal to the number of available NFTs) are chosen at random out of the outstanding tickets. Tickets that aren't winners have their ETH refunded.

This format satisfies universal fairness, as the only thing bidders control is how many raffle tickets they have. They are only given specific NFTs after the launch has ended. The long bidding period also eliminates any race conditions since bidders have time to purchase tickets. This also makes the process easy for bidders. They simply have to decide whether the ticket price is worth it. Assuming the existence of some sort of verifiable randomness (i.e. Chainlink VRF), the distributor has no way to manipulate the mechanism so trustlessness is also satisfied.

The only properties which aren't satisfied are revenue and welfare robustness. Without any knowledge of the bidder's expected bid, the distributor has no information to choose the price of the raffle tickets. Thus similar to FCFS, if the price is too high, not enough tickets will be sold. And if the price is too low, potential revenue will be lost. On the welfare side, bidders have no way to signal they have higher value for winning than others. To combat revenue robustness, a potential variation of this format would be to keep the cost of the raffle ticket for losers as well. Then even if the ticket is under priced, there will be more demand for the tickets driving up revenue. However, an increase in ticket holders lowers the value of each ticket, potentially canceling this effect out.

7.2.3 Batch Auction

Hasu and Agnihotri (2021) mentions another mechanism called "batch auctions," which were used by Jay Pegs Auto Mart. In batch auctions, each bidder i deposits some amount of ETH d_i to a smart contract during the bidding phase. Once the bidding phase is over and the clearing phase begins, the smart contract allocates to bidder i a fractional token of value $d_i \cdot \frac{m}{\sum_j d_j}$. Users can then trade these fractional tokens, and redeem 1 token for 1 NFT item.

This mechanism is fair in the sense that everyone is rewarded proportionally to their bid, and nobody is left with nothing. It avoids race conditions by defining separate bidding and clearing phases. The mechanism achieves sybil resistance naturally - the amount of fractional-tokens that an individual receives is exactly proportional to his total bid. This mechanism is also both revenue robust and welfare robust: the individuals who value the tokens the most will bid the highest and pass that revenue over to the auctioneer.

The main flaw of this mechanism is that it is difficult to use. It's not immediately obvious how much you are getting by bidding a certain amount, making the bidding strategy unclear. There is also potential incentive for the distributor to make bids to drive the prices of the tokens up. Then after the clearing phase, if multiple bidders don't have enough token to purchase an NFT, the distributor can sell their tokens again to bidders. Since bidders have already invested, they are likely to take unfavorable trades at this point to recoup some gain. Additionally, this entire process means users must make at-least 3 on-chain transactions: one for the initial bid, another to trade fractional tokens to reach an integer value, and yet another to exchange the tokens for the NFT. These additional transactions not only make it difficult for users, but also expensive in terms of gas costs.

7.3 Design Challenges

The entire design space of NFT launches is very large, with trade-offs made for every choice. One of the primary tensions is the balance between robustness and making the launch accessible to a large community. Making sure everyone has opportunity to win means that prices have to be relatively low. However, a mechanism that allows bidders to win with low bids means that bidders with high valuations and large bids might not win in the launch. Additionally, these same low prices have to be charged to everyone to keep incentive compatibility. These last two points show how it is hard to keep fairness and robustness at the same time. One option to help with revenue robustness is to keep prices low and keep revenue from losers as well. However, this violates ex-post individual rationality.

Another important choice a distributor has to make is how to take care of the race condition. One choice is to simply use sequential bidding. However, it can be a negative user experience to have to wait a long time on a large bid before knowing if they won. Another way to remove the negative externalities with race conditions is to take the mechanism off-chain and post winners to the chain afterwards. Using a third-party platform means there are no gas fees, but relies on much stronger trust assumptions on the third party acting like they say they will. The guaranteed trust from seeing the public contract code is lost. Another option is to run the mechanism on an L2 with low gas fees and settling the final transactions on L1. The problem with this today is that many potential bidders might not be on the L2. This might drive away potential bidders who do not want to face the cost/inconvenience of moving some of their assets to the L2 to participate in the auction.

Finally, on a more practical point, mechanism designers have to be careful when implementing their launches. Small vulnerabilities in smart contracts can allow for sophisticated bidders to cheat the mechanism and fixing a contract once its deployed can be very difficult. For this reason, it can be argued that simple is better, even at the cost of certain desirable properties.

7.4 New Proposed Mechanisms

None of the above three on-chain mechanisms achieve all three of the following: 1) avoiding race-conditions in the case of a scarce auction, i.e. when $m < n$, 2) resources are allocated to the bidders with the highest valuations, 3) easy and inexpensive to use. In the following subsections, we give initial ideas for mechanisms which may satisfy these three properties.

7.4.1 Multi-Unit Vickrey Auctions

We propose an alternative sequential bidding mechanism that delivers NFTs to the bidders that have the highest value for them

1. The auctioneer posts a reserve price r
2. Over 24 hours, bidders are free to place a bid including a price and a quantity they are willing to purchase (the auctioneer can set a cap on the maximum number of NFTs per bid and use proof of personhood to prevent multiple bids from the same person)
3. The bidders are served in decreasing order from highest to lowest bids with the bidders being allocated the quantity they bid for until the supply runs out
4. If a winner wins k NFTs, they pay the sum of the k highest bids of losers. If one of those bids was for q NFTs, that counts as q of the k highest bids. Also if any of the k bids fall below the reserve price, the reserve price is the value of all the remaining bids for the payment calculation. All losers have their bids refunded.

As usual this auction can be made universally fair by using post launch randomness. If the distributor wishes the NFTs to be distributed to as many unique individuals as possible, they can implement a cap on the quantity in a bid and use proof of personhood to thwart sybils. This auction is also race free since it uses sequential bidding. Furthermore, from well known results on Vickrey auctions, we know the auction is bidder incentive compatible. This makes the bidding strategy easy: simply bid your value for a token by the number of tokens you want. (This can be more complicated depending on how bidder valuations are modeled; i.e. a bidder might want a specific number of tokens to have a higher chance of getting a specific NFT). Since the auction awards NFTs to the highest bidders, it is also welfare robust. We do not comment on the revenue robustness of this mechanism.

One downside is that this auction is not trustless for the same reason Vickrey auctions in single item settings are not trustless. The distributor has the ability to submit false bids driving up the price of the highest losing bids. However, we note that this has a smaller effect in the multi-unit case. When the number of NFTs being sold is large, and there are many participants, bids tend to be close together near the middle of the pack. Thus the gap between the highest losing bid and the lowest winning bid might not be very large. If the distributor decides to raise prices by having themselves beat the lower bidders, then they will incur a cost in unsold inventory.

7.4.2 Multi-item Dutch Sale

We propose a multi-item Dutch Sale:

1. The auctioneer defines a very large initial price of p_0, a step value Δ, and a minimum reserve price r
2. The sale begins at t_0: any bidder can purchase an item for price p_0
3. For any subsequent timestep t_i: a bidder can purchase an item for price $p_i = p_0 - (i \cdot \Delta)$
4. The sale ends when all m items have been sold, or when p_i dips below the auctioneer's minimum reserve price r

This mechanism enables the bidders with the highest valuations to secure items: the m bidders with the highest valuations can simply buy the item as soon as the price drops below their valuation. Universal fairness can be implemented using the standard post launch randomness. If the distributor wishes the NFTs to be distributed to as many unique individuals as possible, they can implement a cap on purchases per address and use proof of personhood to thwart sybils. This launch is also trustless since the distributor can't affect the price curve or demand with any bids.

However, there are a couple of major issues with the mechanism. Firstly, it is not easy to use in that it is not clear what the best bidding strategy is. For example, bidder i with the highest valuation v_i may not want to buy the item at the first timestep t_j for which $p_j \leq v_i$. Bidder i could reason that he will still be able to get an item at a future timestep $j + k$, at a price $p_{j+k} < p_j$. Secondly, whether or not there are race conditions depends on the initial values of p_0 and Δ. If p_0 is set below the market-clearing price, then the sale will degenerate into a FCFS model. In practice, this can be avoided by having the auctioneer set p_0 to a price which is unreasonably high. If Δ is too large, then the difference between demand for adjacent timesteps will be large, causing a demand spike and therefore a race condition. However, one could argue that this race condition is not so relevant: as long as there are less bidders than remaining items at the timestep, and the duration of the timestep is sufficiently long enough for all bidders to have an opportunity to get their bid in, then the ordering of bids does not matter for that timestep. The mechanism is welfare robust as bidders with the highest valuations should be the first to make transactions. Revenue robustness is more ambiguous because there is a general equilibrium where bidders wait to buy even after the price drops below their value to get a better deal. If multiple bidders do this, the mechanism's revenue can fall quite far from optimal.

One possible improvement to this is mechanism would be to have a self-adjusting price function (rather than a simple step function). Similar to ideas from Roughgarden (2020), one could imagine a price function which responds dynamically to the demand from previous blocks. The auctioneer could set a "target rate" of sales, and then measure the historical rate of sales in previous blocks to adjust the current price accordingly. We suggest such a line of inquiry for future work.

8 Future directions

In this section, we provide some future directions that we find would be interesting for consideration as future research.

Relaxation of properties We believe that the properties of seller IC as well as OCA-proofness that we defined above for the single-item NFT Auction framework are overly stringent in their requirements towards the implemented auction mechanism. We discussed to a certain extent how the rigidity of OCA-proofness might be relaxed, but it is clear that more research is needed on that matter, to develop relaxed notions of collusion-proofness that nevertheless succeed to provide enough stability alongside flexibility on the underpinned auction mechanism.

Secure computation One interesting direction to explore would be to investigate how to embed ideas for revealing as little as possible about the bids, in order to avoid attacks by the seller and to keep non-winning sealed bids private.

Non-negligible gas fees Analysis of the proposed protocol in the presence of non-negligible transaction fees might show under which conditions the protocol can be made practical given that transactions to the blockchain are not free.

Further analysis of NFT launches This report provided an introductory glance into the design space of NFT launches. Work on formal analysis for equilibrium bidder/seller behavior with different assumptions on the distributions that bidders draw their bids from remains open. There exists some classical literature, which we posit that it is effectively connected to that, on *common-value* or *interdependent-value* auctions (see Roughgarden and Talgam-Cohen (2016), for instance), and we suggest the idea of connecting those to the aforementioned NFT launches. Additionally, we constrained the defined framework to purely on-chain mechanisms and simple bidding languages. Expanding these assumptions could potentially lead to more interesting design choices. Another choice which could imply fruitful results is to attempt to violate certain constraints like universal fairness and letting bidders make separate bids for different NFTs in the same collection. All the above demonstrate that the richness of the design space to be explored is quite vast and many potential designs remain open for analysis.

References

0xEssential. Fairdrop: A proof-of-concept for fair nft drops, 2021. URL `https://fairdrop.0xessential.com/`.

Mohammad Akbarpour and Shengwu Li. Credible mechanisms. In Éva Tardos, Edith Elkind, and Rakesh Vohra, editors, *Proceedings of the 2018 ACM Conference on Economics and Computation, Ithaca, NY, USA, June 18-22, 2018*, page 371. ACM, 2018. doi: 10.1145/3219166.3219169. URL `https://doi.org/10.1145/3219166.3219169`.

Lorenz Breidenbach, Philip Daian, Florian Tramèr, and Ari Juels. To sink frontrunners, send in the submarines, 2017. URL `https://hackingdistributed.com/2017/08/28/submarine-sends/`.

Vitalik Buterin. Eip-1014: Skinny create2, 2018. URL `https://eips.ethereum.org/EIPS/eip-1014`.

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

Vitalik Buterin. An approximate introduction to how zk-snarks are possible, 2021b. URL `https://vitalik.ca/general/2021/01/26/snarks.html`.

Hao Chung and Elaine Shi. Foundations of transaction fee mechanism design, 2021.

Matheus V. X. Ferreira and S. Matthew Weinberg. Credible, truthful, and two-round (optimal) auctions via cryptographic commitments. In *Proceedings of the 21st ACM Conference on Economics and Computation*, EC '20, page 683–712, New York, NY, USA, 2020. Association for Computing Machinery. ISBN 9781450379755. doi: 10.1145/3391403.3399495. URL `https://doi.org/10.1145/3391403.3399495`.

Oded Goldreich. Secure multi-party computation, 2002. URL https://www.wisdom.weizmann.ac.il/~oded/PSX/prot.pdf.

Hasu and Anish Agnihotri. A guide to designing effective nft launches, 2021. URL https://www.paradigm.xyz/2021/10/a-guide-to-designing-effective-nft-launches/.

Lindsay Howard. Foundation: A complete guide to collecting nfts and how auctions work, 2021. URL https://help.foundation.app/en/articles/4742997-a-complete-guide-to-collecting-nfts-and-how-auctions-work.

Stewart James. Proof of humanity - an explainer, 2021. URL https://blog.kleros.io/proof-of-humanity-an-explainer/.

Logan Kugler. Non-fungible tokens and the future of art. *Commun. ACM*, 64(9):19–20, aug 2021. ISSN 0001-0782. doi: 10.1145/3474355. URL https://doi.org/10.1145/3474355.

Taylor Locke. Jack Dorsey sells his first tweet ever as an NFT for over $2.9 million, March 2021. URL https://www.cnbc.com/2021/03/22/jack-dorsey-sells-his-first-tweet-ever-as-an-nft-for-over-2point9-million.html.

Merriam-Webster. Fungibility. URL https://www.merriam-webster.com/dictionary/fungibility.

Silvio Micali and Michael O. Rabin. Cryptography miracles, secure auctions, matching problem verification. *Commun. ACM*, 57(2):85–93, feb 2014. ISSN 0001-0782. doi: 10.1145/2574871. URL https://doi.org/10.1145/2574871.

Roger B. Myerson. Optimal auction design. *Mathematics of Operations Research*, 6(1):58–73, 1981. ISSN 0364765X, 15265471. URL http://www.jstor.org/stable/3689266.

Jon Perkins. How superrare timed auctions work, 2020. URL https://medium.com/superrare/how-superrare-timed-auctions-work-a351058a6120.

Scott Reyburn. JPG File Sells for $69 Million, as 'NFT Mania' Gathers Pace. *The New York Times*, March 2021. ISSN 0362-4331. URL https://www.nytimes.com/2021/03/11/arts/design/nft-auction-christies-beeple.html.

Tim Roughgarden. Transaction fee mechanism design for the ethereum blockchain: An economic analysis of eip-1559, 2020.

Tim Roughgarden. Transaction fee mechanism design. In *Proceedings of the 22nd ACM Conference on Economics and Computation*, EC '21, page 792, New York, NY, USA, 2021. Association for Computing Machinery. ISBN 9781450385541. doi: 10.1145/3465456.3467591. URL https://doi.org/10.1145/3465456.3467591.

Tim Roughgarden and Inbal Talgam-Cohen. Optimal and robust mechanism design with interdependent values. *ACM Trans. Econ. Comput.*, 4(3), jun 2016. ISSN 2167-8375. doi: 10.1145/2910577. URL https://doi.org/10.1145/2910577.

Rashid Sheikh and Durgesh Kumar Mishra. Protocols for getting maximum value for multi-party computations, 2010. URL https://ieeexplore.ieee.org/document/5489259.