Otter: A Provably MEV-Resilient Automated Market Maker via Surplus Redistribution*

Elaine Shi† Mengqian Zhang‡ Hao Chung§ Yuhao Li¶

Abstract

Miner extractable value (MEV) in automated market makers allows block builders to profit from transaction ordering and injected trades, imposing costs on users and contributing to builder centralization. We introduce Otter (Optimal Truthful Trading with Excess Redistribution), a two-asset batch AMM that achieves provable MEV resilience when the consensus layer provides censorship resilience and block space is uncongested. In particular, Otter makes truthful behavior a dominant strategy for both users and builders. Consequently, a builder cannot profit from strategic deviations, including reordering bids or injecting sybil bids. These guarantees continue to hold even when the builder is itself a trader with intrinsic value. Moreover, we show that our mechanism maximizes social welfare, in a strong sense, within a natural class of mechanisms satisfying the desired game-theoretic properties.

To achieve these guarantees, we introduce a new paradigm called *surplus redistribution*, which provably prevents residual surplus from being captured as MEV by redirecting it to the broader community. Specifically, the pool's output tokens need not be distributed entirely among the users in the current batch. Instead, any residual surplus may be transferred, for example, to a smart contract governed by the decentralized community. The accumulated surplus can subsequently be used to benefit community members in ways that preserve the mechanism's game-theoretic guarantees — for example, by subsidizing traders' transaction fees, rewarding liquidity providers, or returning assets to the pool to reduce price impact and slippage for future traders.

Our approach relies on the underlying consensus layer to provide censorship resilience. We motivate the necessity of this assumption through an impossibility result showing that the desired game-theoretic guarantees become unattainable when the builder is additionally allowed to censor transactions. Thus, our results also provide a mathematically formal demonstration of how consensus-level security guarantees can fundamentally expand what is achievable at the application (i.e., smart-contract) layer. Finally, we establish additional results characterizing the builder fee structures compatible with our desired incentive guarantees.

*The order of authorship is randomized. The random order was generated using the American Economic Association's Author Randomization Tool, and can be verified with confirmation code 6SqRwY10BsKd at https://www.aeaweb.org/journals/policies/random-author-order/search?RandomAuthorsSearch%5Bsearch%5D=6SqRwY10BsKd.

†Carnegie Mellon University and Oblivious Labs. elaineshi@cmu.edu. Supported through a grant from LayerZero via Oblivious Labs.

‡Carnegie Mellon University. mengqianzhang@cmu.edu. Supported through a grant from LayerZero via CMU, and a CyLab Secure Blockchain Initiative grant.

§LayerZero Labs. hao.chung@layerzerolabs.org

¶Columbia University. yuhaoli@cs.columbia.edu

Contents

1 Introduction

MEV, short for Miner (or Maximum) Extractable Value, refers to the widespread phenomenon whereby a block builder leverages its unilateral control over the inclusion and ordering of transactions within a block to extract profit from this privileged position. MEV is harmful in multiple respects. It not only imposes additional costs on ordinary users, but also fosters opaque off-chain markets in which block builders, traders, and specialized searchers compete to identify and exploit MEV opportunities. These off-chain ecosystems, in turn, create strong forces toward centralization. Indeed, a recent measurement study by Yang et al. [YNZ25] found that the top two block builders produce more than 85% of Ethereum blocks, severely undermining the decentralized vision of blockchain systems.

In this work, we focus on Automated Market Makers (AMMs), one of the most widely used applications in decentralized finance (DeFi) and a major source of MEV. Numerous prior works [ZQT+21, Par23, KDC22] have demonstrated how a strategic block builder can exploit its control over transaction inclusion and ordering to launch front-running and back-running attacks, often collectively referred to as *sandwich attacks*. Through such attacks, the block builder can extract essentially risk-free profits at the expense of ordinary users.

In a typical two-asset AMM, users trade one of two assets, denoted X and Y, against a smart contract, which we henceforth also refer to as the *pool*. Initially, the state of the pool is denoted by (x_0, y_0), meaning that the pool holds x_0 units of X and y_0 units of Y. The exchange rate for each trade is determined by a predefined pricing curve. For example, the widely adopted constant-product rule requires

$$(x_0 - x)(y_0 + y) = C$$

for some positive constant C. Thus, starting from state (x_0, y_0), a user who contributes y units of Y to the pool receives x units of X. Following the trade, the pool transitions to state $(x_0 - x, y_0 + y)$, preserving the invariant that the product of the reserves remains equal to C.

1.1 Our Results and Contributions

Our main contribution is a new two-asset AMM mechanism called Otter (short for "Optimal Truthful Trading with Excess Redistribution") that achieves provable MEV resilience, assuming that the underlying consensus protocol provides censorship resilience and that block space is uncongested (i.e., there is ample capacity to include all relevant transactions). Specifically, compared with today's AMMs, our mechanism satisfies several desirable properties:

1. **Provable MEV resilience.** We prove that a block builder (henceforth simply a *builder*) cannot benefit from any strategic manipulation within its control, including misreporting its true type, manipulating transaction ordering, or creating pseudonymous identities to inject sybil bids. Since we assume that the underlying consensus protocol provides censorship resilience, censoring transactions lies outside the builder's strategy space. Moreover, this guarantee continues to hold even when the builder is itself a user with an intrinsic demand to trade; we refer to such an entity as a *builder-as-user*. In particular, the profit-maximizing strategy for a builder or builder-as-user is to truthfully report its type (if it has an intrinsic demand to trade) and refrain from injecting sybil bids. Our mechanism is inherently *order-independent*: its outcome does not depend on the ordering of transactions within a block. Consequently, the builder gains no advantage from reordering transactions and may order them arbitrarily without affecting the outcome.

2. **Truthful reporting and sybil-proofness.** We prove that every user's profit-maximizing strategy is to truthfully report its type, including both its valuation for the asset being purchased and its budget, where both quantities may be denominated in units of the other asset. Furthermore, no user can benefit by creating pseudonymous identities and submitting additional bids through them—a property we refer to as *sybil-proofness*.

   This stands in sharp contrast to today's AMMs. Under common MEV attacks, most notably sandwich attacks [ZQT+21, Par23, KDC22], users may be forced to trade at the worst possible price permitted by their reported slippage tolerance. Consequently, users face an inherently strategic choice when setting their slippage tolerance: it must be sufficiently permissive for the transaction to execute, yet sufficiently restrictive to limit their exposure to MEV extraction.

   Today's users may also have an incentive to delay submitting their bids, hoping to first observe others' bids and then choose their own actions strategically. By contrast, the incentive-compatibility guarantees of our mechanism hold *ex post*: even after observing all other bids, each user, builder, or builder-as-user still maximizes its own profit by behaving truthfully. Thus, observing others' bids before submitting one's own confers no strategic advantage, eliminating the incentive for last-minute bid submission.

3. **Social welfare optimization.** Since truthful behavior constitutes an equilibrium for all participants, including users, builders, and builders-as-users, we show that our AMM maximizes social welfare in a strong sense at equilibrium.

   By contrast, today's AMMs generally incentivize both users and builders to behave strategically, and their resulting equilibrium behavior is poorly understood. Without a well-characterized equilibrium, it is difficult for a mechanism designer to provide rigorous guarantees about economic objectives such as social welfare.

Last but not least, our new AMM is intuitive, conceptually simple, and computationally efficient, making it a promising candidate for practical implementation and deployment.

Surplus redistribution as a new paradigm. To simultaneously achieve the above desiderata, we introduce a new paradigm that we call *surplus redistribution*. Most popular AMMs today [uni], as well as a large body of the literature [FP23, CWS25, LSZ26], impose two requirements: (1) all tokens output by the pool must be completely distributed to the users, and no surplus is allowed to remain; and (2) the resulting state of the pool, after each trade or batch of trades, must land on the prescribed pricing curve (e.g., the constant-product curve).

We depart from this conventional model. While we still require the end state of the pool to conform to the pricing curve, we allow only a portion of the tokens output by the pool to be paid to the users. Rather than allowing the remaining surplus to be captured as MEV by privileged participants, our mechanism redistributes it for the benefit of the broader community. For example, the surplus may be transferred to a smart contract governed through a decentralized mechanism and subsequently allocated in various ways that benefit the community, such as:

(1) as additional rewards for liquidity providers, thereby incentivizing greater liquidity provision and, in turn, reducing price impact and slippage for users;

(2) back to the pool itself, thereby deepening liquidity and reducing the cost of future trades;

(3) to subsidize or offset users' transaction fees; and/or

(4) to the underlying layer-one blockchain or its stakeholders, where it can support the continued development and maintenance of the blockchain, or provide economic incentives that strengthen the security and robustness of the broader ecosystem.

We summarize our main contribution in the following theorem:

Theorem 1 (MEV-resilient AMM mechanism)**.** *Suppose that the underlying consensus layer ensures censorship resilience, and that we allow surplus redistribution. Then, there exists a computationally efficient AMM mechanism that incentivizes truthful reporting, achieves strategy-proofness for a block builder who may or may not have intrinsic demand, and maximizes social welfare in a strong sense (to be formally defined later).*

Mitigating the externalities of an already centralized builder economy. We argue that deploying our new AMM *alongside, and in direct competition with*, legacy AMMs can significantly mitigate the negative externalities arising from today's highly centralized builder economy, in the following ways:

- **Eliminating builders' privileged control over economic surplus.** Today, MEV extraction has contributed to a severe centralization of the builder economy, which has effectively become a duopoly [YNZ25]. This concentration allows a small number of dominant builders to capture a disproportionate share of the economic surplus generated through MEV. By contrast, our new AMM prevents this surplus from being captured as MEV and instead redistributes it to the broader decentralized ecosystem.
- **Pressuring builders to offer competitive pricing.** The first three avenues for surplus redistribution described above can all, directly or indirectly, reduce users' effective trading costs, making our AMM economically competitive with legacy AMMs. Under today's AMMs, users targeted by MEV attacks effectively suffer from the worst execution price permitted by their reported slippage tolerance. By contrast, in our new AMM, increases in the execution price arise naturally from competition among users, and provably never from MEV extraction.

  Deploying our AMM alongside legacy AMMs therefore creates competitive pressure on the existing builder economy. Builders may be compelled to offer users more favorable pricing, effectively returning a greater share of the economic surplus to users and liquidity providers rather than capturing it as MEV.

Necessity of the censorship-resilience assumption. To establish the desirable properties of our AMM, we assume that the underlying consensus layer provides censorship resilience. We argue that this assumption is both necessary and reasonable. We first establish its necessity through the following impossibility result: if block space is finite and the builder has the ability to censor bids, then no non-trivial AMM can simultaneously guarantee truthful reporting for individual users and strategy-proofness for a builder-as-user.

Theorem 2 (Necessity of censorship resilience)**.** *Suppose that block space is finite and that the block builder can censor bids. Then, no non-trivial AMM mechanism can simultaneously incentivize truthful reporting by every individual user and be strategy-proof for a builder-as-user.*

The censorship-resilience assumption is also well motivated, as numerous blockchain projects are actively pursuing censorship resilience at the consensus layer, most notably Ethereum through FOCIL [foc]. Beyond FOCIL, a variety of other approaches to achieving censorship resilience have

been proposed [GNR25, AER25]. More broadly, motivated in part by the urgency of mitigating MEV and its negative externalities, many blockchain projects are exploring enhancements to their consensus protocols that provide stronger guarantees over transaction inclusion and ordering. These efforts include censorship resilience [foc, GNR25, AER25], fair sequencing [dec, esp, KZGJ20, KDK22, KDL+21, Kur20, ZSC+20], and private transaction submission [Dra, BO22, AFP, FPTX25, BLT25, GWWW26, CGPW25, BFOQ25].

Despite this growing effort at the consensus layer, relatively little is understood about how such guarantees can expand the possibilities for application-layer mechanism design from a game-theoretic perspective. In particular, can stronger consensus-layer guarantees enable fundamentally new mechanism-design feasibility results that would otherwise be impossible?

Our results provide a concrete affirmative answer to this question. Together, our AMM construction and Theorem 2 jointly establish a separation between consensus layers with and without censorship resilience: the desired incentive guarantees are achievable in the former setting, whereas no non-trivial AMM can achieve them in the latter. Thus, our results mathematically confirm the community's wide-spread intuition that censorship resilience at the consensus layer can play a fundamental role in mitigating MEV at the application layer.

Additional results. As a by-product of our analysis, we also characterize the permissible forms of builder compensation when a portion of the surplus is allocated to the builder. Specifically, we prove the following:

Theorem 3 (Constraints on surplus returned to the builder)**.** *For any AMM mechanism that incentivizes truthful reporting and is strategy-proof for a builder-as-user, the amount of surplus returned to the builder must be constant and independent of the outcome of the AMM.*

Interestingly, the proofs of both Theorem 2 and Theorem 3 draw on techniques from the recent literature on transaction fee mechanisms (TFMs) [Rou21, CS23]. Our results thus uncover a useful connection between TFM and AMM mechanism design, demonstrating how techniques and insights developed in the study of one can inform the other.

Last but not the least, we also prove that allowing surplus redistribution is essential for obtaining a "dream" AMM mechanism that simultaneously satisfies our desired incentive compatibility and efficiency guarantees.

Theorem 4 (The necessity of surplus redistribution)**.** *If we do not allow surplus redistribution, then no deterministic AMM mechanism can simultaneously incentivize truthful reporting and achieve Pareto optimality for orders with "reasonable" asks (to be formally defined later).*

Theorem 4 also explains why recent works on AMM mechanism design that do not allow surplus redistribution [CWS25, LSZ26] cannot achieve reasonable notions of social welfare maximization.

We stress, however, that we introduce surplus redistribution not merely as a technical device to circumvent the impossibility of Theorem 4, but more importantly, as a desirable paradigm itself to provably defend against MEV and substantially reduce the negative externalities of the already centralized builder economy [YNZ25].

Both Theorem 3 and Theorem 4 hold in a strong sense, even when the underlying consensus offers strong properties including censorship resilience and fair sequencing.

1.2 Comparison with Existing Approaches of MEV Mitigation

Capturing MEV through builder auction. A line of work [Dra23, Dom22, tim] has explored capturing MEV through auctions for block-building or transaction-ordering rights. Such auction-based approaches have also been deployed by major blockchain ecosystems, including Ethereum's

Proposer-Builder Separation (PBS) framework [pbs] and Arbitrum's Timeboost [tim]. Broadly, these approaches allocate privileged block-building or transaction-ordering rights through competitive bidding, allowing some of the associated MEV to be captured by the protocol or returned to the broader community. Unfortunately, such approaches may exacerbate builder centralization by favoring participants with superior access to high-value order flow or other MEV opportunities, thereby disadvantaging smaller builders with fewer resources or customers [YNZ25, BCS24]. As the builder economy becomes increasingly concentrated, diminished competition may in turn allow dominant builders to exercise market power, extract monopoly rents, and retain a larger share of the MEV, ultimately undermining the goal of returning MEV to the broader community.

Application-specific MEV mitigation or recovery at the consensus layer. Several academic works, as well as discussions in the blockchain community, have proposed implementing application-specific MEV mitigation or recovery directly at the consensus layer. For example, Ferreira and Parkes [FP23] propose enforcing application-specific sequencing rules at the consensus layer that restrict how a block builder may order transactions, thereby limiting its ability to extract MEV. In particular, they design sequencing rules for AMM contracts that provide weak but provable guarantees against MEV extraction. Beyond enforcing verifiable sequencing rules, the cryptocurrency communityhave also informally explored the possibility of estimating the amount of MEV extracted by a builder and requiring the builder to return, or pay as a tax, an amount proportional to the estimated MEV to the underlying blockchain.

A fundamental drawback of this general approach is that it pushes application-specific logic into the consensus layer, blurring the architectural boundary between consensus and applications. As the ecosystem grows to support an increasingly diverse collection of applications, incorporating application-specific MEV rules into consensus could substantially increase the complexity of the consensus layer. In the extreme, the consensus layer would need to understand increasingly rich aspects of application semantics in order to determine which builder behaviors are permissible.

Consensus-level MEV recovery faces an additional challenge: it requires the consensus layer to identify and quantify MEV across potentially arbitrary smart contracts. Such an approach may be inherently difficult to sustain in a permissionless environment, where developers can continually deploy new or deliberately obfuscated contracts. Each new application may require the consensus protocol or its maintainers to analyze the contract's semantics and devise an application-specific method for detecting and quantifying its MEV. This creates an inherently reactive dynamic: as new strategies and contracts emerge, the consensus layer must continually adapt its MEV-detection mechanisms to keep pace.

Our approach instead preserves a clean separation of concerns. We ask the consensus layer to provide only generic, application-agnostic primitives—most importantly, censorship resilience—while leaving application-specific incentive design to the application layer. Our results demonstrate that even such a generic consensus-layer guarantee can be sufficient to enable strong, provable MEV resilience at the application layer.

The approach of Ferreira and Parkes [FP23] also faces additional limitations specific to their model. Their framework requires trades to be executed sequentially, with the pool state satisfying the prescribed pricing curve after every individual trade. Under this restriction, they establish strong impossibility results. In particular, even arbitrage resilience—a guarantee strictly weaker than the incentive-compatibility properties we seek—is impossible in their model. They therefore consider a substantially weaker MEV-resilience notion: roughly speaking, if the builder earns risk-free profit, then each user must receive an execution price at least as favorable as the price it would have received had its order been the only order in the block.

1.3 Additional Related Work

Comparison with closely related work. Our work is directly inspired by that of Chan, Wu, and Shi [CWS25], who construct an AMM that processes trades in batches while ensuring that the pool's end state conforms to the pricing curve. Their AMM achieves the same game-theoretic guarantees as ours, but has two important limitations in comparison. First, their construction relies on a substantially stronger assumption about the consensus layer—namely, fair sequencing. Moreover, the idealized fair-sequencing model assumed in their work is stronger than existing guarantees in the literature [dec, esp, KZGJ20, KDK22, KDL+21, Kur20, ZSC+20], which provide only approximate sequencing fairness. It remains unclear whether their idealized notion of fair sequencing can be realized in practice. Second, unlike our AMM, their mechanism does not achieve optimal social welfare.

Automated market makers. Milionis, Moallemi, and Roughgarden [MMR23] study the design of a market maker's demand curve with the dual objectives of maximizing profit and eliciting truthful reports from traders. Their focus is orthogonal to ours, as their model does not seek to capture or mitigate MEV. In particular, traders submit orders directly to the market maker, and their model does not capture strategic transaction ordering and associated arbitrage opportunities, such as front-running or back-running other traders' orders. Bartoletti et al. [BCLL22] study how to determine an adversary's optimal MEV-extraction strategy in an AMM. In particular, they develop an algorithmic procedure for strategically ordering user transactions and inserting adversarial trades so as to maximize the value extracted from users. Their work seeks to characterize and optimize MEV extraction, rather than to design mechanisms that prevent or mitigate MEV.

A line of work has explored batch trading, particularly batch clearing at a uniform price [CF23, CF24, cow, RGGM24]. However, uniform-price batch clearing alone does not guarantee incentive compatibility: when only a subset of eligible orders can be executed, selecting orders based on reported valuations may incentivize users to misreport. For example, Canidio and Fritsch [CF23, CF24] propose a uniform-price batch AMM with a modified potential function that achieves arbitrage resilience, but does not provide incentive compatibility even with an honest miner. Zhang et al. [ZLS+25] likewise observe this limitation and study optimal strategic behavior under batch clearing. Other batch-trading mechanisms [CH24] similarly do not satisfy our notion of incentive compatibility. A recent work by Li, Shi, and Zhang [LSZ26] establishes a somewhat counterintuitive tension between uniform pricing and incentive compatibility. Although uniform pricing eliminates internal arbitrage, they show that it is fundamentally incompatible with the stronger goal of incentive compatibility: in the absence of surplus redistribution, no efficient AMM mechanism can simultaneously achieve uniform pricing and incentive compatibility.

Other related work. Chitra et al. [CK22] proposes to redistribute MEV to stake-holders to enhance the blockchain's economic security; however, their work does not suggest any secure mechanism for capturing the MEV without leading to builder centralization. In this sense, their work is orthogonal to ours. In particular, the surplus accumulated through our Otter mechanism can also be (in part) redistributed to stake-holders as they suggest.

Hartline and Roughgarden [HR08] study routing, scheduling, anti-spam, and resource-allocation settings in which conventional monetary transfers are undesirable or technologically unavailable. They design mechanisms that instead require agents to destroy resources — for example, by expending computation or accepting degraded service — and use such resource destruction as an instrument for providing incentives. Although our work shares the high-level idea of burning, it differs fundamentally in both setting and objective. In our mechanism, the burnt surplus is redistributed to

the broader community and therefore contributes to social welfare; in their framework, by contrast, burnt resources constitute a social loss. Burning has also been used as a mechanism design tool in the recent transaction fee mechanism literature [Rou21, CS23].

The elegant work of Yokoo et al. [YSM04] considered combinatorial auctions with a finite set of indivisible items. They showed that, under a suitable submodularity condition on the welfare function, the VCG auction is sybil-proof, even though VCG is not sybil-proof in general. While our approach to achieving Sybil-proofness is reminiscent of that of Yokoo et al. [YSM04], we stress that our auction setting differs substantially from theirs, and thus their proof does not directly carry over to our setting. Specifically, in our setting, the tokens being traded are infinitely divisible, trades with the pool are governed by a pricing curve, and agents are subject to budget caps. Moreover, our Sybil strategy space is broader: Sybil identities may not only misreport both valuations and budgets, but may also submit bids in the opposite trading direction.

2 Model

Imagine a pool that holds two crypto-assets (also called tokens) named X and Y, respectively. Users can trade with the pool through two types of orders, called sell-Y and sell-X orders, respectively.

Sell-Y orders. A real user i of sell-Y profile has type $(v_i, q_i) \in \mathbb{R}^2_{\geq 0}$, indicating that the user has a *budget* of $q_i \geq 0$ units of Y to spend on purchasing X, and its reservation value of each unit of Y is $v_i \geq 0$, measured in units of X. A sell-Y identity reports a per-unit ask and a declared budget $(\widetilde{v}_i, \widetilde{q}_i) \in \mathbb{R}^2_{\geq 0}$. Truthful reporting means $(\widetilde{v}_i, \widetilde{q}_i) = (v_i, q_i)$.

Sell-X orders. Symmetrically, a real user j of sell-X profile has type $(u_j, r_j) \in \mathbb{R}^2_{\geq 0}$, where r_j is its budget in units of X, and u_j is its reservation value for one unit of X, measured in units of Y. A sell-X identity reports $(\widetilde{u}_j, \widetilde{r}_j) \in \mathbb{R}^2_{\geq 0}$, and truthful reporting means $(\widetilde{u}_j, \widetilde{r}_j) = (u_j, r_j)$.

AMM mechanism. An AMM mechanism receives as input a list of orders, each of either sell-Y or sell-X type, and decides an outcome (x_i, y_i) for each order:

- For a sell-Y order $(\widetilde{v}_i, \widetilde{q}_i)$, it means that the user obtains x_i units in X in exchange for y_i units of Y. It is required that $0 \leq y_i \leq \widetilde{q}_i$.
- For a sell-X order, it means that the user obtains y_i units in Y in exchange for x_i units of X. It is required that $0 \leq x_i \leq \widetilde{r}_i$.

Unless otherwise noted, we require that an AMM mechanism satisfy the following natural properties:

- *Individual rationality.* Under truthful reporting, every user obtains nonnegative utility.
- *Feasibility.* Every token paid to an order is funded by tokens supplied by opposite-side orders or by an AMM trade. No outside subsidy is required. Any surplus compensation token is *burnt* (i.e., redistributed the decentralized community).

Remark 5. As formalized in Section 2.1, the builder, who determines the order in which transactions appear in a block, is also a strategic player. We assume that the AMM mechanism pays no fee to the builder. Section 7 justifies this assumption by showing that our desired incentive guarantees preclude outcome-dependent builder fees. This restriction does not rule out a fixed exogenous payment,

funded, for example, by surplus that the mechanism previously redistributed to the community. Provided that this payment is independent of the current batch's outcome, it has no effect on our game-theoretic analysis.

2.1 Strategy Space and Utility

Strategy space. A truthful user always reports its true type. A strategic user, however, can not only misrepresent its valuation and budget, but also create an arbitrary finite number of identities, and inject sybil bids. The declared valuation and budget associated with each sybil identity can be arbitrary finite nonnegative reals. Moreover, the bids submitted by a strategic user need not be in the same direction as its true type. For example, a sell-Y user may submit sell-X bids, and vice versa.

A strategic builder can arbitrarily order the bids in the block, and inject sybil bids. If the strategic builder also has intrinsic demand, then it can also misreport its valuation and budget. Note that the mechanism observes only the declared valuations and budgets, and it is not aware which identities belong to the same real user or builder-as-user.

Utility. Let $\mathcal{Y}$ and $\mathcal{X}$ be the sets of identities that have sell-Y and sell-X profiles, respectively. For the finite set $\mathcal{I}$ of identities controlled by one real user, define aggregate token flows

$$\begin{aligned} S_Y(\mathcal{I}) &:= \sum_{i \in \mathcal{I} \cap \mathcal{Y}} y_i, \quad R_X(\mathcal{I}) := \sum_{i \in \mathcal{I} \cap \mathcal{Y}} x_i, \\ S_X(\mathcal{I}) &:= \sum_{i \in \mathcal{I} \cap \mathcal{X}} x_i, \quad R_Y(\mathcal{I}) := \sum_{i \in \mathcal{I} \cap \mathcal{X}} y_i. \end{aligned} \tag{1}$$

Here S_T denotes the amount of token $T \in \{X, Y\}$ supplied by the user and R_T denotes the amount of token T received by the user. Its net token changes are

$$\Delta_X(\mathcal{I}) := R_X(\mathcal{I}) - S_X(\mathcal{I}), \qquad \Delta_Y(\mathcal{I}) := R_Y(\mathcal{I}) - S_Y(\mathcal{I}).$$

A true sell-Y user of type (v, q) has utility

$$\mathsf{Util}^Y(\mathcal{I}; v, q) := \begin{cases} -\infty, & \text{if } -\Delta_Y(\mathcal{I}) > q \text{ or } \big(-\Delta_Y(\mathcal{I}) < 0 \text{ and } \Delta_X(\mathcal{I}) < 0\big) \\ \Delta_X(\mathcal{I}) + v\Delta_Y(\mathcal{I}), & \text{o.w.} \end{cases} \tag{2}$$

Thus a true sell-Y user receives catastrophic utility if its realized net sale in Y exceeds the true budget q, or if its aggregate trade is strictly in the opposite direction of its desire, namely, if it obtains a strict net gain in Y while incurring a strict net loss in X. Otherwise its utility is computed in the normal quasilinear way from its aggregate net trade. In particular, the utility function does not penalize a sell-Y user merely for obtaining a net gain in Y when it does not lose X; this includes a potential free-lunch outcome with weak net gains in both tokens. Nor does it penalize the user if it gross-sells more than q units of Y through some identities provided opposite-direction identities buy back enough Y that its final net sale does not exceed q. This makes the possible strategy larger and will later make our game-theoretic definitions stronger.

Symmetrically, for a true sell-X user of type (u, r), its utility is defined as:

$$\mathsf{Util}^X(\mathcal{I}; u, r) := \begin{cases} -\infty, & \text{if } -\Delta_X(\mathcal{I}) > r \text{ or } \big(-\Delta_X(\mathcal{I}) < 0 \text{ and } \Delta_Y(\mathcal{I}) < 0\big) \\ \Delta_Y(\mathcal{I}) + u\Delta_X(\mathcal{I}), & \text{o.w.} \end{cases}$$

Thus a true sell-X user receives utility $-\infty$ if it net-sells more than r units of X, or if it obtains a strict net gain in X while incurring a strict net loss in Y.

2.2 A Concave AMM Pricing Curve

Suppose the pool initially holds finite reserves $x_0 > 0$ and $y_0 > 0$ of tokens X and Y, respectively. Let $\widehat{y}$ denote the curve's effective Y-payout capacity, where $0 < \widehat{y} \leq y_0$. Trades are governed by a pricing curve

$$F : (-\widehat{y}, +\infty) \longrightarrow (-\infty, x_0).$$

A signed amount $y \in (-\widehat{y}, +\infty)$ denotes net Y sent into the AMM, and $F(y)$ denotes net X sent out of the AMM. Thus $y > 0$ describes selling Y to the AMM and receiving X, while $y < 0$ describes withdrawing Y while depositing X. We assume:

(A1) $F(0) = 0$;

(A2) F is continuous, strictly increasing, and concave;

(A3) F is differentiable at 0 with $0 < F'(0) < \infty$;

(A4) F has the lower-endpoint behavior $\lim_{p \downarrow -\widehat{y}} F(p) = -\infty$.

Define the initial prices

$$\sigma_0 := F'(0) \quad (X \text{ per } Y), \qquad \rho_0 := \frac{1}{F'(0)} = \frac{1}{\sigma_0} \quad (Y \text{ per } X).$$

For the sell-Y direction, the mechanism uses the restriction of F to $[0, +\infty)$. In our mechanism description later, whenever we use F as a compensation curve for the sell-Y direction, we are implicitly using F restricted to $[0, +\infty)$. The pool pays out $F(y) \in [0, x_0)$ units of X upon receiving $y \geq 0$ units of Y, and the initial marginal compensation is σ_0. For the sell-X direction, let

$$H : [0, +\infty) \longrightarrow [0, \widehat{y}), \qquad H(x) := -F^{-1}(-x).$$

The pool pays out $H(x)$ units of Y upon receiving $x \geq 0$ units of X. Its initial marginal compensation is $\rho_0 = 1/\sigma_0$. Assumptions (A2) and (A4) ensure that $F^{-1}(-x)$ exists for every $x \geq 0$. Since F is increasing and concave, H is increasing and concave. Thus the mechanism uses $F|_{[0,+\infty)}$ and H as its two one-sided compensation curves. Both are increasing and concave and have unbounded nonnegative input domains.

Constant product as a special case. For a constant-product AMM with initial reserves (x_0, y_0) and $k = x_0 y_0$, take $\widehat{y} = y_0$. Then

$$F(y) = x_0 - \frac{k}{y_0 + y}, \qquad y > -y_0.$$

It follows that

$$F'(y) = \frac{k}{(y_0 + y)^2} > 0, \qquad F''(y) = -\frac{2k}{(y_0 + y)^3} < 0,$$

and

$$\sigma_0 = \frac{x_0}{y_0}, \qquad \rho_0 = \frac{y_0}{x_0}.$$

In other words, the pool pays out $x_0 - \frac{k}{y_0+y}$ units of X for receiving $y \geq 0$ units of Y, and it pays out $-F^{-1}(-x) = y_0 - \frac{k}{x_0+x}$ units of Y for receiving $x \geq 0$ units of X. Hence constant product is a special case of our generalized concave pricing curve formulation.

2.3 Desired Properties

We would like the AMM mechanism to satisfy the following desired properties.

(1) **User incentive compatibility (a.k.a. truthfulness).** User incentive compatibility (UIC) requires that truthfully reporting both the reservation value and the budget be a dominant strategy over the entire admissible strategy space. In particular, a user cannot benefit from misreporting either its valuation or its budget, nor from creating additional identities and injecting sybil bids.

(2) **Strategy proofness for the builder or builder-as-user.** For either a builder or a builder-as-user, acting honestly is a dominant strategy. Specifically, the builder or builder-as-user is incentivized to truthfully report its valuation and budget (if there is any intrinsic demand), refrain from injecting any sybil bids, and order the bids in the prescribed manner[1].

(3) **Pareto optimality for eligible bids.** A bid is considered *eligible* if its ask is no greater than the initial spot compensation. In other words, a sell-Y order $(\widetilde{v}, \widetilde{q})$ is eligible iff $\widetilde{v} \leq \sigma_0$; and a sell-$X$ order $(\widetilde{u}, \widetilde{r})$ is eligible if $\widetilde{u} \leq \rho_0$. We require the following Pareto optimality condition for eligible bids: under truthful reporting,

   (a) no subset of users can trade among themselves (subject to available budgets) in a way that achieves a Pareto improvement in their joint utilities; and

   (b) no individual user can make an additional trade direclty with the pool that strictly improves its utility.

Remark 6 (Our definitions imply sybil-proofness). Since our strategy space allows the user or builder(-as-user) to create sybil identities and inject sybil bids, our user incentive compatibility and strategy proofness for the builder(-as-user) definitions directly imply sybil-proofness against these strategic players. Sybil-proofness is also referred to as false-name-proofness in some prior literature [YSM04, SSIY08, GLT20].

Remark 7 (Why restrict Pareto optimality only to eligible bids). Later in Section A, we prove an impossibility result showing that Pareto optimality cannot be achieved, subject to our incentive-compatibility notions, if non-eligible users are also taken into account. In light of this impossibility, excluding ineligible users from consideration is a natural design choice. Intuitively, the initial spot price can be viewed as setting a baseline price for eligible trades, so that only users whose asks are at least as favorable as this baseline are considered. A similar approach has been adopted in prior work on AMM mechanism design [LSZ26].

Remark 8 (Regarding social welfare maximization). In our problem formulation, the two tokens have no common numeraire, and users may value them at different exchange rates. Therefore, it is tricky to define a global notion of social welfare since utilities denominated in different tokens cannot be canonically aggregated. To argue that our mechanism achieves a reasonable notion of social welfare maximization, we proceed in two steps.

- In this section, we first define Pareto optimality for eligible users. Any reasonable notion of social-welfare maximization restricted to eligible users should, at a minimum, preclude a Pareto improvement for those users.

[1] If the AMM mechanism treats the input bids as a set like ours, then the ordering has no effect on the execution outcome.

- Later in Section 6, we show that our game-theoretic requirements impose strong restrictions on the structure of a mechanism. Given these structural restrictions, we can subsequently introduce a natural welfare criterion for eligible users and prove that our mechanism maximizes it.

Thus, Pareto optimality provides a numeraire-free efficiency requirement at the model level, while welfare maximization is established later within the structure implied by our incentive requirements.

2.4 Basic Facts

UIC implies no arbitrage. Some earlier works [CWS25, LSZ26] on AMM mechanism design also defined an extra property called *no free lunch*. No free lunch requires that for any subset $\mathcal{I}$ of identities, let $\Delta_X := \Delta_X(\mathcal{I})$ and $\Delta_Y := \Delta_Y(\mathcal{I})$ be their net gain in the tokens X and Y, respectively. Then, it must be that $\Delta_X > 0 \implies \Delta_Y < 0$ and $\Delta_Y > 0 \implies \Delta_X < 0$. Intuitively, no free lunch implies that no user or builder(-as-user) can make risk free profit, i.e., the mechanism is *arbitrage-free*. It also means that for a user or builder with no intrinsic demand and budget, playing honestly without placing any bids is a profit-maximizing strategy.

The following fact shows that our notion of UIC implies the no free lunch property. A similar observation was made by Chan et al. [CWS25].

Fact 9 (UIC implies no free lunch)**.** *Suppose that a deterministic AMM mechanism satisfies UIC as defined above. Suppose also that a* truthful *zero-budget user of either direction receives zero utility. Then the mechanism satisfies no free lunch.*

Proof. Fix all reports outside an arbitrary finite collection $\mathcal{I}$ of identities, and write $\Delta_X := \Delta_X(\mathcal{I})$ and $\Delta_Y := \Delta_Y(\mathcal{I})$. Suppose, towards a contradiction, that $\Delta_X > 0$ and $\Delta_Y \geq 0$. For the other direction where $\Delta_Y > 0$ and $\Delta_X \geq 0$, the proof is symmetric.

Interpret the identities in $\mathcal{I}$ as the sybil identities created by a strategic sell-Y user with no intrinsic demand and budget, i.e., its type is $(v, q) = (0, 0)$. Observe that the strategic deviation does not violate its true budget, and its utility is $\mathsf{Util}^Y(\mathcal{I}; 0, 0) = \Delta_X > 0$. By assumption, truthful reporting by the zero-budget user gives utility zero, so this sybil strategy is strictly profitable, contradicting user incentive compatibility. □

For our mechanism, the zero-budget normalization in Theorem 9 is automatic: a truthful report with $q = 0$ (or $r = 0$) has zero allocation, and hence zero compensation. Therefore, once the full cross-direction incentive guarantee is established in Section 5.3, Theorem 9 yields no free lunch as a corollary.

Fact 10 (Order-insensitive mechanism + UIC $\implies$ strategy-proofness for or builder(-as-user))**.** *Suppose that the AMM mechanism treats the input bids as an* unordered *set. Then, under our strategy space, UIC implies strategy proofness for the builder or builder-as-user.*

Proof. Under our strategy space, the only privilege of the builder(-as-user) over an ordinary user is its ability to reorder bids within the block. However, if the mechanism is insensitive to ordering, this ability is useless. Therefore, if the builder(-as-user) with intrinsic type (v, q) has a profitable deviation, then so does an ordinary user with the same type. □

Since our proposed mechanism is order-insensitive, Theorem 10 implies that establishing UIC immediately yields strategy-proofness for the builder (or builder-as-user).

3 Warmup: A One-Sided VCG Auction

As a warmup, we first present an AMM mechanism assuming that there are only sell-Y bids and no sell-X bids are allowed. For the opposite case, in which there are only sell-X bids, a symmetric mechanism can be defined. The one-sided construction applies more generally to any compensation curve

$$F : [0, +\infty) \longrightarrow \mathbb{R}_{\geq 0},$$

that is continuous, strictly increasing, and concave, with

$$F(0) = 0, \qquad \sigma_0 := F'_+(0) \in (0, +\infty).$$

The restriction of the signed AMM curve from Section 2.2 to $[0, +\infty)$ is one such curve.

A reported sell-Y user i has per-unit ask $\widetilde{v}_i \geq 0$ and a budget of $\widetilde{q}_i \geq 0$. Let $0 \leq y_i \leq \widetilde{q}_i$ be the amount of Y user i ends up spending. For a finite set S of bids, define the social welfare under the outcome $\{y_i\}_{i \in S}$ as

$$\mathsf{SW}_F(S, \{y_i\}_{i \in S}) := \underbrace{F\left(\sum_{i \in S} y_i\right)}_{X \text{ tokens output by pool}} - \underbrace{\sum_{i \in S} \widetilde{v}_i y_i}_{\substack{\text{cost of } Y \text{ tokens} \\ \text{sent to pool}}},$$

Specifically, the social welfare is defined using X as a numeraire, and accounts for the total utility of the sell-Y users, as well as the burnt amount that is effectively redistributed to the community. In the above definition of $\mathsf{SW}_F(S)$, the first term is the amount of X tokens output by the pool, and the second term can be viewed as a production cost of the Y tokens sent to the pool.

The maximum achievable social welfare given a finite set S of sell-Y bids is defined as follows:

$$W_F(S) := \max_{0 \leq y_i \leq \widetilde{q}_i \ (i \in S)} \{\mathsf{SW}_F(S, \{y_i\}_{i \in S})\}. \tag{3}$$

We run a VCG-style auction [Vic61, Cla71, Gro73] defined as follows.

- **Allocation rule.** The allocation rule decides how much Y each user spends. Among welfare maximizers, first choose one maximizing total Y sold and then apply an arbitrary exogenous tie-breaking rule. Let y_i^* be the resulting allocation chosen for seller i, and let $Y^* = \sum_i y_i^*$.
- **Compensation rule.** The compensation rule decides the amount of X each user obtains as compensation. Each seller i gets the Clarke pivot compensation defined as follows:

$$\text{let} \quad x_i^* := F(Y^*) - \sum_{j \neq i} \widetilde{v}_j y_j^* - W_F(S \setminus \{i\}). \tag{4}$$

Equivalently, let

$$\phi_i := W_F(S) - W_F(S \setminus \{i\}),$$

then user i's compensation can be written as

$$x_i^* = \widetilde{v}_i y_i^* + \phi_i \tag{5}$$

where the first term represents the cost to user i for spending y_i^* units of Y, and the second term represents a rebate to user i that is equal to the increase in welfare due to user i's existence.

Efficient implementation. The allocation in (3) also has a conceptually simple and efficient implementation. Imagine we start at the initial price and in each step, we extract an infinitesimally small unit dX from the pool, and among users with remaining budget, provide dX to one with the smallest ask $\widetilde{v}_i \le \frac{dX}{dY}$, where $\frac{dX}{dY}$ is the current marginal rate. The seller supplies the corresponding amount dY and receives dX[2]. Repeat the above until the cheapest remaining user has an ask above the current marginal price. At equality, use the same quantity-maximizing and exogenous tie-breaking convention as above.

In particular, the allocation rule can be computed by 1) sorting the users in nondecreasing order based on their asks, 2) for each user in sorted order until the stopping criterion is triggered, compute the point on the AMM's pricing curve at which either the user's budget is exhausted, or its ask exceeds the marginal price (in which case the user's budget is partially spent). Thus the allocation requires $O(|S| \log |S|)$ comparisons, and the computation of $O(|S|)$ crossing points on the pricing curve.

Fact 11 (Equivalence of the efficient implementation)**.** *The efficient implementation described above results in exactly the same allocation as the social-welfare maximizing allocation of Equation* (3)*.*

Proof. For $y \in [0, \sum_i \widetilde{q}_i]$, let

$$C(y) := \min \left\{ \sum_i \widetilde{v}_i y_i : 0 \le y_i \le \widetilde{q}_i,\ \sum_i y_i = y \right\}.$$

An exchange argument shows that $C(y)$ is obtained by filling users in nondecreasing order of ask: if a higher-ask user supplies a positive amount while a lower-ask user has unused capacity, moving any common positive amount from the former to the latter weakly lowers cost (strictly when their asks differ). Hence Equation (3) is equivalent to the one-dimensional problem

$$\max_{0 \le y \le \sum_i \widetilde{q}_i} \{F(y) - C(y)\}. \tag{6}$$

The function C is convex and piecewise linear, and its right marginal cost at y is the ask of the cheapest user that still has capacity. Concavity of F makes its marginal output nonincreasing. Consequently the objective in Equation (6) is nondecreasing while the curve's marginal output is at least the next ask and is nonincreasing once the inequality is reversed. The marginal-auction rule therefore stops at a maximizer of Equation (6). Its convention at equality chooses the maximal maximizing quantity, and filling cheapest capacity with the fixed tie-break then gives precisely the selected maximizer in Equation (3). □

Theorem 12 (One-sided VCG auction)**.** *Let $F : [0, +\infty) \to \mathbb{R}_{\ge 0}$ be a compensation curve satisfying the one-sided curve conditions stated at the beginning of this section. Suppose that only sell-Y bids are allowed. Then, the above one-sided VCG auction satisfies the following properties:*

(a) **Well-formedness***: every user's allocation and compensation are within the desired ranges:*

$$0 \le y_i^* \le \widetilde{q}_i \quad \textit{and} \quad 0 \le x_i^* \le F(Y^*) - F(Y^* - y_i^*) \le \sigma_0 y_i^*; \tag{7}$$

(b) **Individual rationality***: every user's truthful utility is nonnegative;*

[2]We stress that this description *concerns the efficient allocation; the actual compensations are subsequently determined by the Clarke pivot formula* as in Equation (4).

(c) **Feasibility***: aggregate VCG compensation is covered by the concave output:*

$$\sum_i x_i^* \le F(Y^*); \tag{8}$$

(d) **Submodularity of welfare***: the set function W_F is normalized, monotone, and submodular;*

(e) **UIC***: the mechanism satisfies user incentive compatibility (UIC).*

Proof. The allocation bounds in part (a) follow directly from the feasible region in (3). Under truthful reporting,

$$\mathsf{Util}_i = x_i^* - v_i y_i^* = W_F(S) - W_F(S \setminus \{i\}) \ge 0,$$

because adding an optional user cannot reduce the optimum. This proves (b) and also shows $\phi_i \ge 0$ in (5). To obtain the upper bound, delete user i but retain every other user's selected allocation. This is feasible for $S \setminus \{i\}$, so

$$W_F(S \setminus \{i\}) \ge F(Y^* - y_i^*) - \sum_{j \ne i} \widetilde{v}_j y_j^*.$$

Substitution into (4) yields

$$x_i^* \le F(Y^*) - F(Y^* - y_i^*).$$

Because a concave function has nonincreasing marginal slope and initial right slope σ_0,

$$F(Y^*) - F(Y^* - y_i^*) \le \sigma_0 y_i^*.$$

Nonnegativity follows from (5). Together with the allocation bound proved above, this completes the proof of (a).

Summing the above upper bound on x_i^* gives

$$\sum_i x_i^* \le \sum_i \big(F(Y^*) - F(Y^* - y_i^*)\big).$$

If $Y^* = 0$, the claim is immediate since it must be the case that for every i, $y_i^* = 0$ and $x_i^* \le F(Y^*) - F(Y^* - y_i^*) = 0$. Otherwise concavity and $F(0) = 0$ imply

$$F(Y^* - y_i^*) \ge \left(1 - \frac{y_i^*}{Y^*}\right) F(Y^*).$$

Therefore

$$F(Y^*) - F(Y^* - y_i^*) \le \frac{y_i^*}{Y^*} F(Y^*).$$

Summing over i proves (8), and hence (c).

We next prove submodularity. For $t \in [0, \sigma_0]$, let

$$D_S(t) := \sum_{i \in S} \widetilde{q}_i \mathbf{1}\{\widetilde{v}_i \le t\}$$

be aggregate capacity of Y whose ask is at most t, and let

$$K_F(t) := \sup\{y \ge 0 : F'_+(y) \ge t\},$$

where the supremum is allowed to be $+\infty$. Let $\alpha_S(y)$ be the inverse aggregate supply curve: the minimum reported per-unit cost of the yth infinitesimal unit of the capacity. For a fixed total input Y, the minimum reported cost is

$$\int_0^Y \alpha_S(y)\,dy,$$

while concavity gives

$$F(Y) = \int_0^Y F'_+(y)\,dy.$$

Since F'_+ is nonincreasing and α_S is nondecreasing, their difference is nonincreasing. Maximizing over Y therefore integrates exactly the region where marginal compensation is at least marginal cost. A layer-cake/Fubini argument gives

$$W_F(S) = \int_0^{\sigma_0} \min\{D_S(t), K_F(t)\}\,dt. \tag{9}$$

For every fixed t, the map

$$S \longmapsto \min\left\{\sum_{i\in S} \widetilde{q}_i \mathbf{1}\{\widetilde{v}_i \le t\}, K_F(t)\right\}$$

is a capped modular set function and is therefore monotone and submodular. Integration preserves these properties, and $W_F(\varnothing) = 0$. This proves (d).

Finally, we prove UIC. Fix all reports not controlled by one real user, and denote their set by O. Let the user, whose true cost is vy on $[0, q]$, submit an arbitrary finite collection $\mathcal{I}$ of admissible reports — a singleton collection is an ordinary joint misreport of value and cap. Let y_i, x_i be the selected sale and compensation of report $i \in \mathcal{I}$, and let

$$Y_{\mathcal{I}} := \sum_{i\in\mathcal{I}} y_i.$$

The reports may have budget caps whose sum exceeds q. If $Y_{\mathcal{I}} > q$, the real user's utility is $-\infty$, so the deviation is immediately unprofitable. It therefore remains only to analyze the case $Y_{\mathcal{I}} \le q$. Put

$$\phi_i := W_F(O \cup \mathcal{I}) - W_F(O \cup (\mathcal{I} \setminus \{i\})).$$

By (5),

$$x_i = \widetilde{v}_i y_i + \phi_i.$$

Submodularity implies

$$\sum_{i\in\mathcal{I}} \phi_i \le W_F(O \cup \mathcal{I}) - W_F(O).$$

Hence the real user's utility from the possibly sybil deviation is

$$\begin{aligned}\mathsf{Util}^{\mathrm{Syb}} &= \sum_{i\in\mathcal{I}} x_i - vY_{\mathcal{I}} \\ &\le \sum_{i\in\mathcal{I}} \widetilde{v}_i y_i + W_F(O \cup \mathcal{I}) - W_F(O) - vY_{\mathcal{I}}.\end{aligned}$$

At the chosen allocation for $O \cup \mathcal{I}$, the reported costs of the sybil identities $\mathcal{I}$ cancel with the same terms inside $W_F(O \cup \mathcal{I})$. Thus

$$\mathsf{Util}^{\mathrm{Syb}} \le F\left(Y_{\mathcal{I}} + \sum_{j\in O} y_j\right) - \sum_{j\in O} \widetilde{v}_j y_j - vY_{\mathcal{I}} - W_F(O).$$

Because $Y_{\mathcal{I}} \le q$, the first three terms are the welfare of a feasible allocation after merging $\mathcal{I}$ into a truthful user of type (v, q) that sells $Y_{\mathcal{I}} \le q$. Therefore

$$\mathsf{Util}^{\mathrm{Syb}} \le W_F(O \cup \{(v, q)\}) - W_F(O),$$

which is exactly the utility from reporting (v, q) truthfully as a single identity. Hence no deviation is profitable. This proves UIC, including ordinary joint misreports and sybil bids, and establishes (e). □

4 The Full Two-Sided Mechanism

We now describe our full Otter mechanism, which extends the one-sided mechanism in Section 3.

Intuition. The full mechanism first discards all ineligible bids whose asks are less favorable than the initial price. Among the remaining eligible bids, it then computes the spot-eligible supply on both sides at the initial spot price. Henceforth, let D_Y and D_X denote the quantities of Y and X, respectively, available to be sold at the initial spot price or better. Without loss of generality, suppose that $D_Y \ge \rho_0 D_X$, in which case we say that the sell-Y side is *dominant*; the other case is symmetric. The mechanism then fully executes all eligible sell-X orders at the initial spot price. On the sell-Y side, we run the one-sided VCG auction of Section 3 using a modified curve $\widetilde{F}_Y^M(y)$ defined for $y \ge 0$. This curve is obtained by extending $F(y)$ so that the first $M := \rho_0 D_X$ units of Y sold enjoy the initial spot price. Equivalently, we can view the pool as being augmented with D_X units of X at the initial spot price.

4.1 Detailed Description

A sell-Y report is willing to sell at the initial spot price exactly when $\widetilde{v}_i \le \sigma_0$. Define total spot-eligible sell-Y quantity, in units of Y, by

$$D_Y := \sum_{i \in \mathcal{Y}} \widetilde{q}_i \mathbf{1}\{\widetilde{v}_i \le \sigma_0\}.$$

Similarly, define total spot-eligible sell-X quantity, in units of X, by

$$D_X := \sum_{j \in \mathcal{X}} \widetilde{r}_j \mathbf{1}\{\widetilde{u}_j \le \rho_0\}.$$

At the spot price, D_X units of X require $\rho_0 D_X$ units of Y as compensation. We call a profile *sell-Y dominated* if

$$D_Y \ge \rho_0 D_X, \tag{10}$$

and *sell-X dominated* otherwise. Ties are resolved in favor of the sell-Y direction.

Sell-Y dominated profiles. Suppose (10) holds and define

$$M := \rho_0 D_X.$$

Thus M is the number of units of Y needed to buy all spot-eligible sell-X supply at the initial price.

First, fully execute every sell-X order with $\widetilde{u}_j \le \rho_0$ at the initial spot price:

$$x_j = \widetilde{r}_j, \qquad y_j = \rho_0 \widetilde{r}_j. \tag{11}$$

Every sell-X identity with $\widetilde{u}_j > \rho_0$ is assigned zero allocation and zero compensation. These minority-side trades supply D_X units of X and require exactly M units of Y as compensation.

The dominant sell-Y side faces the following effective compensation curve, measured in units of X available as a function of units of Y sold:

$$\widetilde{F}_Y^M(y) := \begin{cases} \sigma_0 y, & 0 \le y \le M, \\ D_X + F(y - M), & y \ge M. \end{cases} \tag{12}$$

Because $D_X = \sigma_0 M$ and $F'_+(0) = \sigma_0$, the two pieces meet continuously with the same initial marginal slope. Since F is concave, $\widetilde{F}_Y^M$ is increasing and concave, and therefore satisfies the one-sided curve conditions.

Run the one-sided VCG mechanism of Section 3 on all sell-Y reports with compensation curve $\widetilde{F}_Y^M$. The interpretation of (12) is intuitive: the first $M = \rho_0 D_X$ units of dominant-side Y can be compensated using the D_X units of X supplied by the sell-X orders at the initial spot price. If more than M units of Y are sold, only the excess Y is sent to the AMM and the original curve F is used thereafter.

Fact 13. *Let Q_Y be the aggregate Y spent by the sell-Y users. Then, $Q_Y \ge M$.*

Proof. By definition, the total reported budget of sell-Y identities with ask at most σ_0 is at least M. On $[0, M]$, the marginal compensation of $\widetilde{F}_Y^M$ is exactly σ_0. If a welfare maximizer sold a total $Q_Y < M$, some unsold spot-eligible capacity would remain. Increasing a seller with ask strictly below σ_0 would strictly increase welfare; increasing one with ask exactly σ_0 would preserve welfare while increasing total sold quantity. Either alternative contradicts optimality together with our quantity-maximizing tie rule. Hence $Q_Y \ge M$. □

Sell-X dominated profiles. Now suppose

$$D_Y < \rho_0 D_X.$$

Define the amount of X needed to buy all spot-eligible sell-Y supply at the initial spot price:

$$M := \sigma_0 D_Y = \frac{D_Y}{\rho_0}.$$

First, fully execute every sell-Y order with $\widetilde{v}_i \le \sigma_0$ at the initial spot price:

$$y_i = \widetilde{q}_i, \qquad x_i = \sigma_0 \widetilde{q}_i.$$

Every sell-Y identity with $\widetilde{v}_i > \sigma_0$ is assigned zero allocation and zero compensation. These minority-side trades supply D_Y units of Y and require exactly M units of X as compensation.

The dominant sell-X side faces the effective compensation curve, now in units of Y available as a function of units of X sold,

$$\widetilde{H}_X^M(x) := \begin{cases} \rho_0 x, & 0 \le x \le M, \\ D_Y + H(x - M), & x \ge M. \end{cases}$$

Because $D_Y = \rho_0 M$ and $H'_+(0) = \rho_0$, this is increasing and concave. It therefore satisfies the one-sided curve conditions. Run the one-sided VCG mechanism of Section 3 on all sell-X reports using $\widetilde{H}_X^M$.

Fact 14. *Let Q_X denote the aggregate X spent by sell-X users. Then, $Q_X \geq M$.*

Proof. The proof is symmetric to that of Theorem 13. □

Theorem 15 (Main theorem: our full two-sided mechanism)**.** *Suppose that the pricing curve F satisfies assumptions (A1)–(A4) defined in Section 2.2. Our full two-sided mechanism is well-formed and satisfies individual rationality, feasibility, user incentive compatibility (UIC), strategy-proofness against a builder(-as-user), and Pareto optimality for eligible users.*

The remainder of the paper proves the theorem. In comparison with our one-sided guarantees in Theorem 12, we must now account for deviations that can change which side is dominant or submit reports in both directions.

5 Proofs for the Two-Sided Mechanism

5.1 Well-Formedness, Individual Rationality, and Feasibility

Without loss of generality, we prove well-formedness, individual rationality, and feasibility only for the sell-Y dominated case, since the other case is symmetric.

Well-formedness and individual rationality. Well-formedness and individual rationality for the minority side follow in a straightforward fashion. Well-formedness and individual rationality for the dominant side follow directly from Theorem 12.

Feasibility. By Lemma 13, the aggregate Y spent by sell-Y users satisfies $Q_Y \geq M$. Recall that the minority sell-X orders supply D_X units of X and must receive $M = \rho_0 D_X$ units of Y. Thus the compensation to sell-X users is fully funded by the sell-Y orders. After using M units of the dominant sell-Y supply to compensate the sell-X users, the residual $y := Q_Y - M \geq 0$ units of Y are sent into the AMM's pool, which returns exactly $F(y)$ units of X. The total amount of X available to compensate dominant sell-Y sellers is therefore

$$D_X + F(Q_Y - M) = \widetilde{F}_Y^M(Q_Y).$$

By Equation (8) of Theorem 12,

$$P_X := \sum_{i \in \mathcal{Y}} x_i \leq \widetilde{F}_Y^M(Q_Y) = D_X + F(Q_Y - M).$$

where x_i is the amount of X tokens compensated to user $i \in \mathcal{Y}$. Thus every dominant-side compensation payment is fully funded, and the burnt surplus $B_X := D_X + F(Q_Y - M) - P_X \geq 0$.

5.2 Pareto Optimality for Eligible Users

We now prove Pareto optimality for eligible users. Without loss of generality, we assume the sell-Y dominated case, since the proof for the other case is symmetric.

Every eligible sell-X user is on the minority side and, by (11), exhausts its input budget. Consequently, after the mechanism executes, no eligible sell-X user has any X left available to sell. Any nontrivial trade among eligible users would require both a user supplying X and a user supplying Y. Thus no subset of eligible users can make an additional admissible trade among themselves, let alone one that is a Pareto improvement.

It remains to rule out a strictly beneficial additional trade with the pool. Let $\widetilde{v}_i \le \sigma_0$ be the ask of an eligible sell-Y user. Let $y := Q_Y - M \ge 0$ be the residual Y input to the pool after compensating the sell-X users where Q_Y is the aggregate Y spent by sell-Y users. Define the final marginal prices by

$$\sigma_e := F'_+(y) \quad (X \text{ per } Y), \qquad \rho_e := \frac{1}{\sigma_e}.$$

An eligible user whose input budget is exhausted cannot make such a trade. Fix, therefore, a dominant-side sell-Y user i that is underfilled, i.e., $y_i < \widetilde{q}_i$. For convenience, let $G(\cdot) := \widetilde{F}_Y^M(\cdot)$. If its ask satisfied $\widetilde{v}_i < G'_+(Q_Y)$, then increasing y_i by a sufficiently small amount would raise reported welfare, contradicting the optimality of social welfare achieved by the allocation rule. Hence

$$\widetilde{v}_i \ge G'_+(Q_Y) = F'_+(Q_Y - M) = \sigma_e$$

where the first equality follows from the fact that $Q_Y \ge M$ and the two pieces of (12) meet with the same slope. For any additional amount $0 < \delta \le \widetilde{q}_i - y_i$, concavity of F gives

$$F(y + \delta) - F(y) \le F'_+(y)\delta = \sigma_e \delta \le \widetilde{v}_i \delta.$$

The additional pool output is therefore no greater than the user's reported cost of supplying the additional input, so the trade cannot strictly improve its utility. This proves both Pareto-optimality conditions in a sell-Y dominated profile.

5.3 Incentive Compatibility

It suffices to prove user incentive compatibility (UIC), since by Theorem 10, UIC directly gives strategy-proofness for a builder(-as-user) for our mechanism which is order-insensitive.

We first prove a technical lemma which will later help us bound the utility of a strategic user who submits opposite-direction sybil orders, effectively extending the curve's available offering at the initial price.

Lemma 16 (Technical lemma)**.** *Fix a finite set O of sell-Y reports other than one comparison seller, and for $M, q \ge 0$ let*

$$\phi_Y(M; v, q) := W_{\widetilde{F}_Y^M}\big(O \cup \{(v, q)\}\big) - W_{\widetilde{F}_Y^M}(O).$$

For every $v \le \sigma_0$, every $\alpha \ge 0$, and every comparison cap $\widehat{q}$ satisfying $0 \le \widehat{q} \le q + \alpha$,

$$\phi_Y(M + \alpha; v, \widehat{q}) \le \phi_Y(M; v, q) + (\sigma_0 - v)\alpha. \tag{13}$$

Symmetrically, if O is a fixed set of sell-X reports and

$$\phi_X(M; u, r) := W_{\widetilde{H}_X^M}\big(O \cup \{(u, r)\}\big) - W_{\widetilde{H}_X^M}(O),$$

then for $u \le \rho_0$, $\alpha \ge 0$, and $0 \le \widehat{r} \le r + \alpha$,

$$\phi_X(M + \alpha; u, \widehat{r}) \le \phi_X(M; u, r) + (\rho_0 - u)\alpha.$$

Proof. We prove (13) since the other direction is symmetric. Apply the layer-cake representation (9). Let $D_O(t)$ be the aggregate capacity in O with ask at most t, and let $K_M(t)$ be the threshold quantity associated with the curve $\widetilde{F}_Y^M$. Extending the initial spot-price segment from M to $M + \alpha$ translates this threshold by α: for almost every $t \in [0, \sigma_0]$,

$$K_{M+\alpha}(t) = K_M(t) + \alpha.$$

For $v \le \sigma_0$, the layer-cake formula gives

$$\phi_Y(M; v, q) = \int_v^{\sigma_0} \min\{q, (K_M(t) - D_O(t))_+\}\, dt.$$

For every real w, every $q, \alpha \ge 0$, and every $0 \le \widehat{q} \le q + \alpha$,

$$\min\{\widehat{q}, (w + \alpha)_+\} \le \min\{q, w_+\} + \alpha. \tag{14}$$

Indeed, replacing $\widehat{q}$ by the larger $q+\alpha$ can only increase the left-hand side, after which the inequality is immediate by considering $w \ge 0$, $-\alpha < w < 0$, and $w \le -\alpha$. Apply Equation (14) pointwise with $w = K_M(t) - D_O(t)$ and integrate over $t \in [v, \sigma_0]$. The interval has length $\sigma_0 - v$, yielding Equation (13). □

5.3.1 A Real Sell-Y User

Fix all reports of other real users and consider one real sell-Y user of type (v, q). Let

$$D_Y^{\text{oth}} := \sum_{\substack{i \text{ other sell-}Y \\ \widetilde{v}_i \le \sigma_0}} \widetilde{q}_i, \qquad D_X^{\text{oth}} := \sum_{\substack{i \text{ other sell-}X \\ \widetilde{u}_i \le \rho_0}} \widetilde{r}_i, \qquad M_Y^{\text{oth}} := \rho_0 D_X^{\text{oth}},$$

where D_Y^{oth} and D_X^{oth} are the total spot-eligible sell-Y and sell-X capacities of the other users, respectively. Thus D_Y^{oth}, D_X^{oth}, and M_Y^{oth} are fixed before the strategic user's sybil identities are introduced.

Truthful utility. The following facts about truthful utility are easy to see:

- If $v > \sigma_0$, truthful utility is zero.
- If $v \le \sigma_0$ and $D_Y^{\text{oth}} + q < M_Y^{\text{oth}}$, the truthful user is on the minority side and receives

  $$\mathsf{Util}^{\text{tr}} = (\sigma_0 - v)q.$$

  Moreover, we have

  $$\phi_Y(M_Y^{\text{oth}}; v, q) = (\sigma_0 - v)q = \mathsf{Util}^{\text{tr}}. \tag{15}$$

- If $v \le \sigma_0$ and $D_Y^{\text{oth}} + q \ge M_Y^{\text{oth}}$, truthful sell-$Y$ is dominant and

  $$\mathsf{Util}^{\text{tr}} = \phi_Y(M_Y^{\text{oth}}; v, q) \ge (\sigma_0 - v)\ell, \qquad \ell := (M_Y^{\text{oth}} - D_Y^{\text{oth}})_+. \tag{16}$$

Replace the truthful identity by an arbitrary finite set $\mathcal{I}$ of sybil bids in either direction. Henceforth, we use the aggregate flow notation from Equation (1). The user's net Y sale is $-\Delta_Y = S_Y - R_Y$. If $-\Delta_Y > q$, the attack has utility $-\infty$ by Equation (2). A negative net sale is not catastrophic by itself: under Equation (2) we must also show that it entails a strict net loss in X. We verify this separately in each dominance regime below. After doing so, essentially we only need to consider the regime

$$0 \le -\Delta_Y \le q. \tag{17}$$

Notice that Equation (17) does *not* imply $S_Y \le q$.

The sybil profile is sell-Y dominated. Every spot-eligible sell-X sybil identity is then on the minority side and is fully executed at the spot price; every ineligible one executes zero. Therefore $R_Y = \rho_0 S_X$. These opposite-direction sybils extend the dominant sell-Y crossing parameter from M_Y^{oth} to $M_Y^{\text{oth}} + R_Y$.

The individual payment bound (7), summed over the user's dominant-side sell-Y sybils, gives $R_X \le \sigma_0 S_Y$. If net Y sale $-\Delta_Y = S_Y - R_Y < 0$, then $\Delta_X = R_X - S_X = R_X - \sigma_0 R_Y \le \sigma_0(S_Y - R_Y) < 0$. In this case, the second catastrophic condition in (2) applies, and the attack has $-\infty$ utility. Moreover, if $-\Delta_Y = S_Y - R_Y > q$, the attack also has $-\infty$ utility. Thus we may assume

$$R_Y \le S_Y \le q + R_Y. \tag{18}$$

The user's utility is

$$\begin{aligned} \mathsf{Util}^{\text{syb}} &= R_X - S_X - v(S_Y - R_Y) \\ &= R_X - \sigma_0 R_Y - v(S_Y - R_Y). \end{aligned} \tag{19}$$

If $v > \sigma_0$, using again $R_X \le \sigma_0 S_Y$ together with $S_Y - R_Y = -\Delta_Y \ge 0$,

$$\mathsf{Util}^{\text{syb}} \le (\sigma_0 - v)(S_Y - R_Y) \le 0 = \mathsf{Util}^{\text{tr}}.$$

Now suppose $v \le \sigma_0$. Freeze the curve $\widetilde{F}_Y^{M_Y^{\text{oth}}+R_Y}$ and merge the user's sell-Y sybils only for purposes of comparison. The UIC guarantee in part (e) of Theorem 12, with a virtual sell-Y user of true cost vz and budget equal to the realized gross sale S_Y, gives

$$R_X - vS_Y \le \phi_Y(M_Y^{\text{oth}} + R_Y; v, S_Y). \tag{20}$$

By (18), $S_Y \le q + R_Y$. Combining (19), (20), and Lemma 16,

$$\begin{aligned} \mathsf{Util}^{\text{syb}} &\le \phi_Y(M_Y^{\text{oth}} + R_Y; v, S_Y) - (\sigma_0 - v)R_Y \\ &\le \phi_Y(M_Y^{\text{oth}}; v, q). \end{aligned}$$

The last quantity equals truthful utility both when truth is dominant, by the VCG marginal-contribution identity, and when truth is minority, by (15). Thus this attack cannot help.

The sybil profile is sell-X dominated. The user's active sell-Y sybil identities are now minority identities, so they receive $R_X = \sigma_0 S_Y$. The dominant-side payment bound gives $R_Y \le \rho_0 S_X$, or equivalently $S_X \ge \sigma_0 R_Y$. If $-\Delta_Y = S_Y - R_Y < 0$, then we also have $\Delta_X = \sigma_0 S_Y - S_X \le \sigma_0(S_Y - R_Y) < 0$. Hence Equation (2) assigns the attack utility $-\infty$. Moreover, if $-\Delta_Y = S_Y - R_Y > q$, the attack also has $-\infty$ utility. We may therefore restrict finite-utility attacks in this regime to $0 \le S_Y - R_Y \le q$. Their utility is

$$\mathsf{Util}^{\text{syb}} = \sigma_0 S_Y - S_X - v(S_Y - R_Y) \le (\sigma_0 - v)(S_Y - R_Y) = (\sigma_0 - v) \cdot (-\Delta_Y). \tag{21}$$

If $v > \sigma_0$, this is nonpositive and cannot beat truthful utility zero. If $v \le \sigma_0$ and truthful reporting makes sell-Y the minority side, then $-\Delta_Y \le q$ makes Equation (21) at most the truthful value $(\sigma_0 - v)q$.

It remains to consider the case $v \le \sigma_0$ in which truthful reporting makes sell-Y dominant. Rewrite the user's utility as

$$\mathsf{Util}^{\text{syb}} = (\sigma_0 - v)(S_Y - \rho_0 S_X) + v(R_Y - \rho_0 S_X). \tag{22}$$

The second term is nonpositive because $R_Y \le \rho_0 S_X$. Let S_X^{oth} be the amount of X sold, in the sybil outcome, by all other sell-X identities. Theorem 14 gives

$$S_X + S_X^{\mathrm{oth}} \ge \sigma_0(D_Y^{\mathrm{oth}} + S_Y).$$

Given that $S_X^{\mathrm{oth}} \le D_X^{\mathrm{oth}} = \sigma_0 M_Y^{\mathrm{oth}}$, we have

$$S_Y - \rho_0 S_X \le M_Y^{\mathrm{oth}} - D_Y^{\mathrm{oth}} \le \ell.$$

Since $\sigma_0 - v \ge 0$, Equation (22) and (16) yield

$$\mathsf{Util}^{\mathrm{syb}} \le (\sigma_0 - v)\ell \le \mathsf{Util}^{\mathrm{tr}}.$$

Thus no arbitrary finite cross-direction sybil attack benefits a true sell-Y user under the utility function defined in Equation (2).

5.3.2 A Real Sell-X User

Fix all reports of other real users and consider one real sell-X user of type (u, r). Let

$$D_Y^{\mathrm{oth}} := \sum_{\substack{i \text{ other sell-}Y \\ \widetilde{v}_i \le \sigma_0}} \widetilde{q}_i, \qquad D_X^{\mathrm{oth}} := \sum_{\substack{i \text{ other sell-}X \\ \widetilde{u}_i \le \rho_0}} \widetilde{r}_i, \qquad M_X^{\mathrm{oth}} := \sigma_0 D_Y^{\mathrm{oth}}.$$

Thus M_X^{oth} is the other users' spot-eligible sell-Y capacity measured in units of X. The proof is obtained from the proof for a real sell-Y user by making the substitutions

$$\begin{gathered} X \longleftrightarrow Y, \qquad F \longleftrightarrow H, \qquad \sigma_0 \longleftrightarrow \rho_0, \qquad (v, q) \longleftrightarrow (u, r), \\ D_Y^{\mathrm{oth}} \longleftrightarrow D_X^{\mathrm{oth}}, \qquad M_Y^{\mathrm{oth}} \longleftrightarrow M_X^{\mathrm{oth}}, \qquad \phi_Y(M; v, q) \longleftrightarrow \phi_X(M; u, r), \\ S_Y \longleftrightarrow S_X, \qquad R_Y \longleftrightarrow R_X, \qquad S_X^{\mathrm{oth}} \longleftrightarrow S_Y^{\mathrm{oth}}, \qquad -\Delta_Y \longleftrightarrow -\Delta_X, \end{gathered}$$

and replacing the dominant-side curve $\widetilde{F}_Y^M$ by $\widetilde{H}_X^M$.

The only modifications concern boundary classification, because ties are resolved in favor of sell-Y dominance. For a truthful eligible sell-X user, truth is on the minority side when

$$D_X^{\mathrm{oth}} + r \le M_X^{\mathrm{oth}},$$

where equality is included, and its utility is $(\rho_0 - u)r = \phi_X(M_X^{\mathrm{oth}}; u, r)$. Truth is on the dominant side only when

$$D_X^{\mathrm{oth}} + r > M_X^{\mathrm{oth}},$$

in which case its utility is $\phi_X(M_X^{\mathrm{oth}}; u, r)$ and is at least

$$(\rho_0 - u)(M_X^{\mathrm{oth}} - D_X^{\mathrm{oth}})_+.$$

Likewise, a sybil profile is sell-X dominated only under a strict imbalance $D_Y < \rho_0 D_X$. The reverse branch, $D_Y \ge \rho_0 D_X$, includes equality and is sell-Y dominated. Thus, at equality, the user's spot-eligible sell-X sybils are minority orders and execute fully at the spot rate ρ_0, while its ineligible sell-X sybils execute zero. With these changes to the weak and strict inequalities, every bound in the sell-Y proof carries over under the substitutions above.

6 Social Welfare Maximization

Defining natural. We call a deterministic AMM mechanism *natural* if it satisfies the following properties:

(Drop) *Drop ineligible bids.* Every ineligible bid receives zero allocation and zero transfer. Thus a sell-X report with $\widetilde{u}_j > \rho_0$ receives $(x_j, y_j) = (0, 0)$, and a sell-Y report with $\widetilde{v}_i > \sigma_0$ receives $(y_i, x_i) = (0, 0)$.

(NPR) *No price reversal.* The residual trade with the AMM does not reverse the direction implied by the spot-eligible imbalance. Thus, at a sell-Y-dominated profile the AMM weakly receives Y, and at a sell-X-dominated profile it weakly receives X. Additionally, the final marginal price moves weakly against the dominant side. In particular, if the profile is sell-Y dominated, then

$$\sigma_e \leq \sigma_0, \qquad \text{and} \qquad \rho_e \geq \rho_0.$$

If the profile is sell-X dominated, then

$$\rho_e \leq \rho_0, \qquad \text{and} \qquad \sigma_e \geq \sigma_0.$$

Here σ_e and ρ_e are the final marginal compensation rates in units of X per Y and Y per X, respectively. At a nondifferentiable point, the inequalities refer to the marginal rate for an additional trade in the relevant dominant direction.

Justification of naturalness. The (Drop) condition stipulates that the mechanism simply ignores ineligible bids. This assumption is justified in Section A, where we prove that no UIC mechanism can achieve Pareto optimality for all bids including ineligible ones. In fact, Section A establishes a stronger impossibility result: even an efficiency notion strictly weaker than Pareto optimality is incompatible with UIC.

(NPR) formalizes the conventional price-impact intuition that aggregate selling pressure on an asset should not cause that asset to appreciate. This is analogous to the classical Walrasian tâtonnement principle that prices adjust in the direction of excess demand [BAH59]. (NPR) imposes only this directional requirement and makes no assumption about the magnitude of the resulting price movement.

Lemma 17 (Structure of minority-side execution)**.** *Fix a report profile and a deterministic natural mechanism satisfying UIC and Pareto optimality for eligible users. Every executed minority-side order is compensated at the initial spot price. Moreover, every minority-side order whose ask is strictly below the initial spot price exhausts its reported budget.*

In particular, at a sell-Y-dominated profile, define M by

$$\sigma_0 M = \sum_{j \in \mathcal{X}} x_j, \tag{23}$$

so that $\sigma_0 M$ is the total X spent by the minority side. The minority side receives exactly M units of Y, and the dominant side spends at least M units of Y:

$$\sum_{i \in \mathcal{Y}} y_i \geq M. \tag{24}$$

The symmetric statements hold at a sell-X-dominated profile.

Proof. We prove the result for a sell-Y-dominated profile. Consider an eligible minority-side sell-X user j. If $\widetilde{u}_j < \rho_0$ and $x_j < \widetilde{r}_j$, then (NPR) gives $\rho_e \geq \rho_0 > \widetilde{u}_j$. The user could sell a sufficiently small additional amount of X directly to the pool at a marginal rate strictly above its reservation value, contradicting Pareto optimality for eligible users. Hence

$$x_j = \widetilde{r}_j \qquad \text{whenever } \widetilde{u}_j < \rho_0. \tag{25}$$

Fix the other reports and bidder j's reported budget r. For an ask $t < \rho_0$, let $h(t)$ denote its compensation in Y. By (25), every such report sells r. UIC applied to any two asks $s, t < \rho_0$, in both directions, implies $h(s) = h(t) = C$. A type (t, r) can report an ineligible ask and obtain zero by (Drop), so $C - tr \geq 0$; letting $t \uparrow \rho_0$ gives $C \geq \rho_0 r$. Conversely, a type (a, r) with $a > \rho_0$ obtains zero truthfully and can report $t < \rho_0$, so UIC gives $0 \geq C - ar$; letting $a \downarrow \rho_0$ gives $C \leq \rho_0 r$. Thus $C = \rho_0 r$.

For a boundary report $\widetilde{u}_j = \rho_0$, UIC relative to an ineligible report gives $y_j - \rho_0 x_j \geq 0$. A type $a > \rho_0$ that deviates to this boundary report must obtain utility at most zero, so $0 \geq y_j - a x_j$. Letting $a \downarrow \rho_0$ gives the reverse inequality. Consequently,

$$y_j = \rho_0 x_j \tag{26}$$

for every eligible minority-side user (and hence for every executed one, by (Drop)). Summing (26) and using $\rho_0 = 1/\sigma_0$ shows that the minority side receives M units of Y. If $y := \sum_{i \in \mathcal{Y}} y_i$ is the dominant side's total expenditure, token balance leaves $y - M$ units of Y as the residual input to the AMM. By (NPR) this residual trade is in the dominant direction, so $y - M \geq 0$. This proves (24). The sell-X-dominated case follows by symmetry. □

Definition of dominant-side welfare. By Theorem 17, at a sell-Y-dominated outcome the minority side spends $\sigma_0 M$ units of X and receives M units of Y, while the dominant side spends some total $y \geq M$ units of Y. The first M units of dominant-side expenditure are crossed with the minority supply, and the remaining $y - M$ units are sent to the pool. We therefore define dominant-side social welfare as the X produced by these two sources minus the reported production cost of the dominant-side Y:

$$\mathsf{DSW}_Y := \sigma_0 M + F(y - M) - \sum_{i \in \mathcal{Y}} \widetilde{v}_i y_i, \qquad y := \sum_{i \in \mathcal{Y}} y_i. \tag{27}$$

For a sell-X-dominated outcome, define M by $\rho_0 M = \sum_{i \in \mathcal{Y}} y_i$ and, writing $x := \sum_{j \in \mathcal{X}} x_j \geq M$, define symmetrically

$$\mathsf{DSW}_X := \rho_0 M + H(x - M) - \sum_{j \in \mathcal{X}} \widetilde{u}_j x_j.$$

Ineligible bids make zero contribution by (Drop). The compensation token of the dominant direction is the numeraire; no arbitrary common numeraire is imposed across the two dominance regimes.

Theorem 18 (Welfare maximality among natural mechanisms)**.** *Fix a report profile. Among deterministic natural mechanisms satisfying UIC and Pareto optimality for eligible users, our mechanism has the following properties:*

(i) every strictly eligible minority-side user obtains the same truthful utility under every mechanism in this class, while every boundary or ineligible minority-side user obtains zero;

(ii) if the profile is sell-Y dominated, our mechanism maximizes DSW_Y; if it is sell-X dominated, our mechanism maximizes DSW_X.

Proof. We prove the claims for a sell-Y-dominated profile. The sell-X-dominated case follows by symmetry.

By Theorem 17, a strictly eligible sell-X user j sells its full budget and is compensated at the initial price. Its truthful utility is therefore $(\rho_0 - u_j)r_j$. A boundary user has zero utility whether or not it is executed, and an ineligible user has zero utility by (Drop). This proves (i).

Dominant-side welfare. Let

$$M^* := \rho_0 D_X$$

be the crossing length used by our mechanism. It fully executes every spot-eligible sell-X bid, including boundary bids. For another natural mechanism, define M as in (23). The lemma and the reported caps imply $M \le M^*$; the only possible difference is partial execution of bids with ask exactly ρ_0.

For purposes of the comparison, extend the expression in (27) to all $y \ge 0$ by writing

$$\widetilde{F}^M(y) := \begin{cases} \sigma_0 y, & 0 \le y \le M, \\ \sigma_0 M + F(y - M), & y \ge M. \end{cases}$$

On every outcome of a mechanism in the comparison class, the lemma gives $y \ge M$, so the second branch is precisely the welfare expression already defined in (27).

We claim that, for every $y \ge 0$, the modified compensation curve is nondecreasing in its crossing length:

$$M \le M^* \quad \Longrightarrow \quad \widetilde{F}^M(y) \le \widetilde{F}^{M^*}(y). \tag{28}$$

If $y \le M$, the two sides are equal to $\sigma_0 y$. If $M < y \le M^*$, concavity and $F'_+(0) = \sigma_0$ give $F(y - M) \le \sigma_0(y - M)$, so the claim follows. If $y > M^*$, concavity gives

$$F(y - M) - F(y - M^*) \le \sigma_0(M^* - M),$$

which again proves (28).

Let (y_i) be the dominant-side allocation of the other mechanism. It is a feasible candidate for our mechanism's one-sided welfare problem because the reported caps are the same. Our allocation maximizes the one-sided objective with curve $\widetilde{F}^{M^*}$, and (28) therefore gives

$$\begin{aligned} \mathsf{DSW}_Y^{\text{ours}} &\ge \widetilde{F}^{M^*}\left(\sum_i y_i\right) - \sum_i \widetilde{v}_i y_i \\ &\ge \widetilde{F}^{M}\left(\sum_i y_i\right) - \sum_i \widetilde{v}_i y_i = \mathsf{DSW}_Y^{\text{other}}. \end{aligned}$$

This proves part (ii). Exchanging X with Y, F with H, and σ_0 with ρ_0 proves both claims for a sell-X-dominated profile. □

Remark 19 (Weak local efficiency suffices). In fact, Theorems 17 and 18 continue to hold even if Pareto optimality for eligible users is replaced by the weaker notion of *weak local efficiency* defined in [LSZ26], which only requires that eligible users cannot obtain a Pareto improvement by trading with the pool. Indeed, the only use of Pareto optimality in the proofs is to show that a strictly eligible minority-side user with unspent budget could profitably make an additional trade with the pool. Weak local efficiency rules out exactly this possibility.

7 Restrictions on Builder Fee Structure

Our mechanism burns the unused compensation token. A natural alternative is to transfer some of that excess to the block builder and burn only the remainder. In this section, we show that subject to UIC and strategy-proofness for a builder who is simultaneously a user with intrinsic demand, the only possible builder fee structure is to have the residual surplus entirely "burnt", i.e., redistributed to the community, and pay the builder zero fees. In practice, however, one can remunerate the builder with a constant exogenous block subsidy that is independent of the current batch's outcomes, e.g., using the community surplus accmulated so far. Such a payment will not affect the game theoretic guarantees achieved by the mechanism.

Our "zero builder fee" impossibility can be viewed as a seller-side analogue of the "zero miner revenue" impossibility in transaction-fee mechanisms; see, for example, Theorem 4.7 of Chung and Shi [CS23]. Our proof techniques are also inspired by their work.

7.1 Additional Preliminaries and Notations

Because this is an impossibility result, it is enough to restrict attention to a smaller type and strategy space. Specifically, we consider a a finite sell-Y-only batch. Moreover, since our proof will only need strategies that misreport valuation, we may fix arbitrary budget $\bar{q}_i > 0$. Thus we effectively consider a single-parameter environment where only the per-unit reservation values are private. Any mechanism that is incentive compatible on the richer type space of the main model must remain incentive compatible on this single-parameter restriction.

Although our earlier feasibility is a deterministic AMM mechanism, our impossibility result in this section holds even for *randomized* mechanisms. For a report vector $\widetilde{\mathbf{v}} = (\widetilde{v}_1, \ldots, \widetilde{v}_n)$, let

$$\bar{y}_i(\widetilde{\mathbf{v}}) \in [0, \bar{q}_i]$$

denote the *expected* amount of Y sold by user i, and let $\bar{x}_i(\widetilde{\mathbf{v}})$ denote its *expected* compensation in X. Let

$$\bar{\mu}(\widetilde{\mathbf{v}}) \geq 0,$$

and $\bar{\beta}(\widetilde{\mathbf{v}}) \geq 0$ denote the *expected* builder fee and *expected* burn amount, respectively. In the above, all expectations are taken over the mechanism's random coins. Let

$$\bar{Y}(\widetilde{\mathbf{v}}) := \sum_i \bar{y}_i(\widetilde{\mathbf{v}}).$$

In our impossibility result (Theorem 22), we will make use of individual rationality (IR), feasibility, user incentive compatibility (UIC), and strategy-proofness for a builder who is also a user with intrinsic demand. Since we want the impossibility to work even for possibly randomized mechanisms, the UIC and strategy-proofness definitions will be modified to use the expected utility instead. Note that for a builder-as-user, the fees obtained by the builder will be added to its utility. For a feasibility result we would ideally want IR to hold with probability 1. However, for our impossibility proof, we only need IR to hold in the expectation too, that is, $\bar{x}_i - v_i \bar{y}_i \geq 0$. Feasibility requires that the builder fees be entirely funded only from the current batch's available AMM output. Ex-post feasibility and concavity of F imply the following expected accounting constraint:

Fact 20 (Accounting constraint imposed by feasibility)**.** *For a possibly randomized AMM mechanism with a concave pricing curve that satisfies feasibility, the following accounting constraint must hold when there are only sell-Y bids.*

$$\sum_i \bar{x}_i(\widetilde{\mathbf{v}}) + \bar{\mu}(\widetilde{\mathbf{v}}) + \bar{\beta}(\widetilde{\mathbf{v}}) \leq F\big(\bar{Y}(\widetilde{\mathbf{v}})\big)\,. \tag{29}$$

Proof. Feasibility requires that with probability 1, the compensation to all users, the builder's fees, and the burnt amount are all fully funded by AMM pool's output. Taking expectation, and applying Jensen's inequality using the concavity of F, we immediately get the stated accounting constrant. □

Throughout, we assume that the mechanism is not aware which identities are owned by the builder.

7.2 Proof of Zero Builder Fee

Lemma 21 (Seller payment uniqueness)**.** *Fix a sell-Y user i and the reports of all other users. Let $\bar{y}_i(\widetilde{v}_i)$ be user i's expected amount of Y spent when it reports reservation value $\widetilde{v}_i$. Suppose two expected compensation rules $\bar{x}_i(\widetilde{v}_i)$ and $\bar{x}'_i(\widetilde{v}_i)$ both make truthful reporting dominant for every true reservation value v_i, under utilities $\bar{x}_i(\widetilde{v}_i) - v_i\bar{y}_i(\widetilde{v}_i)$ and $\bar{x}'_i(\widetilde{v}_i) - v_i\bar{y}_i(\widetilde{v}_i)$, respectively. Then $\bar{x}'_i(\widetilde{v}_i) - \bar{x}_i(\widetilde{v}_i)$ is constant in $\widetilde{v}_i$.*

Proof. This is the seller form of the usual single-parameter payment-uniqueness lemma. Dominant-strategy truthfulness implies that $\bar{y}_i(\widetilde{v}_i)$ is nonincreasing in $\widetilde{v}_i$. For the truthful compensation rule $\bar{x}_i$, the envelope identity [Mye81, AT01] (i.e., the unique payment identity of Myerson's lemma adapted to a seller or procurement setting) gives

$$\bar{x}_i(v_i) - v_i\bar{y}_i(v_i) = \bar{x}_i(0) - \int_0^{v_i} \bar{y}_i(s)\,ds,$$

and hence

$$\bar{x}_i(v_i) = v_i\bar{y}_i(v_i) + \bar{x}_i(0) - \int_0^{v_i} \bar{y}_i(s)\,ds.$$

The same identity holds for $\bar{x}'_i$ with exactly the same spending rule $\bar{y}_i$. Subtracting the two identities yields

$$\bar{x}'_i(v_i) - \bar{x}_i(v_i) = \bar{x}'_i(0) - \bar{x}_i(0),$$

which is independent of v_i. The argument applies to expected spending and expected compensation for a randomized mechanism. □

Theorem 22 (Zero builder fees)**.** *Let $F : [0, +\infty) \to \mathbb{R}_{\geq 0}$ be continuous, strictly increasing, and concave, with $F(0) = 0$, and let $\sigma_0 = F'_+(0) \in (0, +\infty)$. Consider any possibly randomized mechanism that satisfies individual rationality, feasibility, UIC, and strategy-proofness for a builder-as-user. Then, for every report profile $\widetilde{\mathbf{v}}$,*

$$\bar{\mu}(\widetilde{\mathbf{v}}) = 0. \tag{30}$$

In particular, because the builder fee is nonnegative, zero expectation implies that it is zero with probability 1*.*

As mentioned, this impossibility uses only sell-Y profiles and only deviations that misreport one's valuation. Thus the impossibility continues to hold when allowing profiles in both directions and a broader strategy space, e.g., when users and builders can misreport both their valuation and budget, and when a builder may additionally censor or reorder bids or inject sybil orders.

Proof. (of Theorem 22.) Fix a user identity i and all reports $\widetilde{\mathbf{v}}_{-i}$. For a possible report $\widetilde{v}_i$ of identity i, abbreviate $\bar{y}_i(\widetilde{\mathbf{v}}_{-i}, \widetilde{v}_i)$, $\bar{x}_i(\widetilde{\mathbf{v}}_{-i}, \widetilde{v}_i)$, and $\bar{\mu}(\widetilde{\mathbf{v}}_{-i}, \widetilde{v}_i)$ by $\bar{y}_i(\widetilde{v}_i)$, $\bar{x}_i(\widetilde{v}_i)$, and $\bar{\mu}(\widetilde{v}_i)$, respectively. By user incentive compatibility, $\bar{x}_i(\widetilde{v}_i)$ is a truthful seller compensation rule for spending rule $\bar{y}_i(\widetilde{v}_i)$. Now consider the same numerical bid profile in the admissible role assignment in which identity i belongs to the block builder. Strategy-proofness for a builder-as-user says that

$$\bar{x}_i(\widetilde{v}_i) + \bar{\mu}(\widetilde{v}_i)$$

is also a truthful seller compensation rule for the *same* spending rule $\bar{y}_i(\widetilde{v}_i)$, because the builder's total utility from this restricted deviation is

$$\bar{x}_i(\widetilde{v}_i) + \bar{\mu}(\widetilde{v}_i) - v_i \bar{y}_i(\widetilde{v}_i).$$

Lemma 21 therefore implies that

$$\bar{\mu}(\widetilde{v}_i) \text{ is independent of } \widetilde{v}_i \quad \text{when } \widetilde{\mathbf{v}}_{-i} \text{ is fixed.} \tag{31}$$

Since strategy-proofness for a builder-as-user applies no matter which identity is builder-owned, (31) holds for every coordinate. Thus expected builder revenue is invariant under changing the bids one coordinate at a time. For every finite profile $\widetilde{\mathbf{v}}$,

$$\bar{\mu}(\widetilde{\mathbf{v}}) = \bar{\mu}(\sigma_0, \ldots, \sigma_0). \tag{32}$$

It remains to evaluate the anchor profile in which every user's true reservation value and truthful report equal σ_0. User individual rationality gives, for every i,

$$\bar{x}_i \geq \sigma_0 \bar{y}_i.$$

Summing and writing $\bar{Y} = \sum_i \bar{y}_i$,

$$\sum_i \bar{x}_i \geq \sigma_0 \bar{Y}. \tag{33}$$

Concavity, $F(0) = 0$, and $F'_+(0) = \sigma_0$ imply the global tangent bound

$$F(y) \leq \sigma_0 y \qquad \text{for every } y \geq 0. \tag{34}$$

Using the funding constraint (29) and (33)–(34),

$$\begin{aligned} \sigma_0 \bar{Y} + \bar{\mu} + \bar{\beta} &\leq \sum_i \bar{x}_i + \bar{\mu} + \bar{\beta} \\ &\leq F(\bar{Y}) \\ &\leq \sigma_0 \bar{Y}. \end{aligned}$$

Hence

$$\bar{\mu} + \bar{\beta} \leq 0.$$

Both quantities are nonnegative, so in particular

$$\bar{\mu}(\sigma_0, \ldots, \sigma_0) = 0.$$

Combining this with (32) proves (30). □

8 Impossibility with Censorship

For our feasibility result, we assume that the underlying consensus guarantees censorship resilience, and thus the ability to censor is excluded from the builder's strategy space. In this section, we justify this assumption by proving an impossibility. We show that when the block size is finite and the block builder may censor bids, no AMM mechanism can simultaneously satisfy individual rationality, feasibility, UIC, as well as strategy-proofness for a builder who is simulatenously a user with intrinsic demand. Further, this impossibility holds even for randomized mechanisms. The proofs in this section are inspired by the transaction fee mechanism literature, e.g., see Corollary 4.9 of Chung and Shi [CS23].

A finite block model allowing censorship. More specifically, let k denote the finite block size. The builder is tasked with selecting up to k bids from among all outstanding bids and including them in the block in some prescribed order. Then, some trusted on-chain rule makes allocation and compensation decisions based on the included bids and their relative ordering.

A strategic builder-as-user may observe the other users' bids and then strategically choose which bids to include and how to sequence them. In particular, it may omit any pending bids and may inject and include bids associated with its own sybil identities.

The strategy space of a strategic user remains the same as before. As in Section 7, we assume that the mechanism is unaware which user identities are owned by the builder.

Anonymity. We assume that the AMM mechanism is anonymous: its on-chain rules make allocation and payment decisions based solely on the valuations and budgets reported in the bids, as well as their relative ordering within the block. Any identity-related information (e.g., public keys) is ignored. This anonymity assumption is both desirable and satisfied by real-world AMMs, as well as by the mechanisms proposed in this paper. In particular, anonymity is consistent with the core decentralization principles underlying modern blockchains.

Non-triviality. We call the mechanism *non-trivial* if there is some block S and an identity $i \in S$ whose expected compensation is strictly positive: $\mathbb{E}[x_i(S)] > 0$. The expectation is over all randomization of the mechanism.

Theorem 23 (Impossibility with censorship)**.** *Assume that the block size is finite. No non-trivial, anonymous AMM mechanism can simultaneously satisfy individual rationality, feasibility, UIC, and strategy-proofness for a builder-as-seller.*

Proof. Suppose, toward a contradiction, that the mechanism is non-trivial. Then some block S, of size at most k, contains a bid j with positive expected compensation. Write

$$T := \mathbb{E}[x_j(S)] > 0, \qquad A := \mathbb{E}[y_j(S)],$$

and let $q \geq 0$ be the public quantity cap of that bid. If $q > 0$, choose

$$0 < \varepsilon < \frac{T}{2q}.$$

If $q = 0$, choose any $\varepsilon > 0$; the cost of every feasible allocation to an added identity is then zero.

Create many low-ask identities. Construct a pending-bid set containing all identities in $S \setminus \{j\}$ and N additional identities, each with cap q, true reservation value ε, and truthful ask ε. Run the prescribed block-construction strategy on this pool. For each added identity i, let Util_i denote its expected trading utility under truthful behavior; an omitted bid has allocation and compensation zero. Individual rationality gives $\mathsf{Util}_i \geq 0$.

One identity has small truthful utility. Finite block capacity bounds the added identities' total compensation independently of their number. Indeed, if Q_S is the total budget of the bids in $S \setminus \{j\}$, a realized block contains at most k bids and hence has total budget at most $Q_S + kq$. Feasibility, nonnegative builder fee and burn, and monotonicity of F therefore give

$$\sum_{i=1}^{N} x_i \leq \sum_{h \text{ included}} x_h \leq F(Q_S + kq) =: C.$$

Since each added identity has nonnegative cost,

$$\sum_{i=1}^{N} \mathsf{Util}_i = \mathbb{E}\left[\sum_{i=1}^{N}(x_i - \varepsilon y_i)\right] \leq C.$$

Consequently some identity i^* satisfies $\mathsf{Util}_{i^*} \leq C/N$.

Censorship gives that identity a profitable deviation. Now imagine that the unlucky user i^* belongs to the builder. Its builder fee is zero with probability 1 by Theorem 22, so the builder's truthful utility is Util_{i^*}. The builder can deviate by censoring the prescribed block, including $S \setminus \{j\}$ together with i^*, and making i^* submit j's report. The on-chain rule then sees the same numerical block S, with i^* in place of j. Hence the deviation gives the builder expected utility

$$T - \varepsilon A \geq T - \varepsilon q > \frac{T}{2},$$

again with zero builder fee. Taking $N > 2C/T$ makes the truthful utility $\mathsf{Util}_{i^*} \leq C/N < T/2$, contradicting builder-as-seller incentive compatibility. Therefore no block can give any seller positive expected compensation. □

References

[AER25] Ittai Abraham, Yuval Efron, and Ling Ren. The latency cost of censorship resistance. *IACR Cryptol. ePrint Arch.*, 2025:2136, 2025.

[AFP] Amit Agarwal, Rex Fernando, and Benny Pinkas. Efficiently-thresholdizable batched identity based encryption, with applications. In *Advances in Cryptology - CRYPTO 2025 - 45th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 17-21, 2025, Proceedings, Part III*, Lecture Notes in Computer Science.

[AT01] Aaron Archer and Éva Tardos. Truthful mechanisms for one-parameter agents. In *Proceedings of the 42nd IEEE Symposium on Foundations of Computer Science*. IEEE Computer Society, 2001.

[BAH59] H. D. Block, Kenneth J. Arrow, and Leonid Hurwicz. On the stability of the competitive equilibrium, ii. *Econometrica*, 27(1):82–109, 1959.

[BCLL22] Massimo Bartoletti, James Hsin-yu Chiang, and Alberto Lluch Lafuente. Maximizing extractable value from automated market makers. In *Financial Cryptography and Data Security: 26th International Conference, FC 2022, Grenada, May 2–6, 2022, Revised Selected Papers*, page 3–19, Berlin, Heidelberg, 2022.

[BCS24] Jonah Burian, Davide Crapis, and Fahad Saleh. Mev capture and decentralization in execution tickets, 2024.

[BFOQ25] Jan Bormet, Sebastian Faust, Hussien Othman, and Ziyan Qu. BEAT-MEV: epochless approach to batched threshold encryption for MEV prevention. In *34th USENIX Security Symposium, USENIX Security 2025, Seattle, WA, USA, August 13-15, 2025*, pages 3457–3476. USENIX Association, 2025.

[BLT25] Dan Boneh, Evan Laufer, and Ertem Nusret Tas. Batch decryption without epochs and its application to encrypted mempools. *IACR Cryptol. ePrint Arch.*, 2025:1254, 2025.

[BO22] Joseph Bebel and Dev Ojha. Ferveo: Threshold decryption for mempool privacy in BFT networks. Cryptology ePrint Archive, Paper 2022/898, 2022. https://eprint.iacr.org/2022/898.

[CF23] Andrea Canidio and Robin Fritsch. Batching trades on automated market makers. In *AFT*, 2023.

[CF24] Andrea Canidio and Robin Fritsch. Arbitrageurs' profits, lvr, and sandwich attacks: batch trading as an amm design response. 2024.

[CGPW25] Arka Rai Choudhuri, Sanjam Garg, Guru-Vamsi Policharla, and Mingyuan Wang. Practical mempool privacy via one-time setup batched threshold encryption. In Lujo Bauer and Giancarlo Pellegrino, editors, *34th USENIX Security Symposium, USENIX Security 2025, Seattle, WA, USA, August 13-15, 2025*, pages 3477–3495. USENIX Association, 2025.

[CH24] Andrea Canidio and Felix Henneke. Fair combinatorial auction for blockchain trade intents: Being fair without knowing what is fair, 2024.

[CK22] Tarun Chitra and Kshitij Kulkarni. Improving proof of stake economic security via mev redistribution. In *Proceedings of the 2022 ACM CCS Workshop on Decentralized Finance and Security*, DeFi'22, page 1–7, New York, NY, USA, 2022. Association for Computing Machinery.

[Cla71] Edward Clarke. Multipart pricing of public goods. *Public Choice*, 11(1):17–33, September 1971.

[cow] https://cow.fi/cow-protocol.

[CS23] Hao Chung and Elaine Shi. Foundations of transaction fee mechanism design. In *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 3856–3899. SIAM, 2023.

[CWS25] T-H. Hubert Chan, Ke Wu, and Elaine Shi. Mechanism design for automated market makers. In *AFT*, 2025.

[dec] Challenging periods reimagined: The key role of sequencer decentralization. https://ethresear.ch/t/challenging-periods-reimagined-the-key-role-of-sequencer-decentralization/15110.

[Dom22] Domothy. Burning mev through block proposer auctions. Ethereum Research Forum, October 2022.

[Dra] Justin Drake. Encrypted mempools. https://www.youtube.com/watch?v=XRM0CpGY3sw.

[Dra23] Justin Drake. Mev burn—a simple design, 2023.

[esp] The espresso sequencer. https://hackmd.io/@EspressoSystems/EspressoSequencer.

[foc] Eip-7805: Fork-choice enforced inclusion lists (focil). https://eips.ethereum.org/EIPS/eip-7805.

[FP23] Matheus Venturyne Xavier Ferreira and David C. Parkes. Credible decentralized exchange design via verifiable sequencing rules. In *STOC*, 2023.

[FPTX25] Rex Fernando, Guru-Vamsi Policharla, Andrei Tonkikh, and Zhuolun Xiang. Trx: Encrypted mempools in high performance BFT protocols. *IACR Cryptol. ePrint Arch.*, 2025:2032, 2025.

[GLT20] Yotam Gafni, Ron Lavi, and Moshe Tennenholtz. VCG under sybil (false-name) attacks - A bayesian analysis. In *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7-12, 2020*, pages 1966–1973. AAAI Press, 2020.

[GNR25] Pranav Garimidi, Joachim Neu, and Max Resnick. Multiple concurrent proposers: Why and how. Cryptology ePrint Archive, Paper 2025/1772, 2025.

[Gro73] Theodore Groves. Incentives in teams. *Econometrica*, 41(4):617–631, 1973.

[GWWW26] Junqing Gong, Brent Waters, Hoeteck Wee, and David J. Wu. Threshold batched identity-based encryption from pairings in the plain model. In *Advances in Cryptology - EUROCRYPT 2026 - 45th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Rome, Italy, May 10-14, 2026, Proceedings, Part V*, volume 16545 of *Lecture Notes in Computer Science*, pages 548–578. Springer, 2026.

[HR08] Jason D. Hartline and Tim Roughgarden. Optimal mechansim design and money burning. In *STOC*, 2008.

[KDC22] Kshitij Kulkarni, Theo Diamandis, and Tarun Chitra. Towards a theory of maximal extractable value I: constant function market makers. *CoRR*, abs/2207.11835, 2022.

[KDK22] Mahimna Kelkar, Soubhik Deb, and Sreeram Kannan. Order-fair consensus in the permissionless setting. In *APKC '22: Proceedings of the 9th ACM on ASIA Public-Key Cryptography Workshop*, 2022.

[KDL+21] Mahimna Kelkar, Soubhik Deb, Sishan Long, Ari Juels, and Sreeram Kannan. Themis: Fast, strong order-fairness in byzantine consensus. 2021.

[Kur20] Klaus Kursawe. Wendy, the good little fairness widget: Achieving order fairness for blockchains. In *Proceedings of the 2nd ACM Conference on Advances in Financial Technologies*, 2020.

[KZGJ20] Mahimna Kelkar, Fan Zhang, Steven Goldfeder, and Ari Juels. Order-fairness for byzantine consensus. In *CRYPTO*, page 451–480, 2020.

[LSZ26] Yuhao Li, Elaine Shi, and Mengqian Zhang. A trilemma in amm mechanism design. In *Financial Cryptography and Data Security (FC)*, 2026.

[MMR23] Jason Milionis, Ciamac C. Moallemi, and Tim Roughgarden. A myersonian framework for optimal liquidity provision in automated market makers. *CoRR*, abs/2303.00208, 2023.

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

[Par23] Andreas Park. The conceptual flaws of decentralized automated market making. *Management Science*, 69(11):6731–6751, 2023.

[pbs] Propose-builder separation. https://ethereum.org/roadmap/pbs/.

[RGGM24] Geoffrey Ramseyer, Mohak Goyal, Ashish Goel, and David Mazières. Augmenting batch exchanges with constant function market makers. In *EC*, 2024.

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

[SSIY08] Yuko Sakurai, Yasumasa Saito, Atsushi Iwasaki, and Makoto Yokoo. Beyond quasi-linear utility: Strategy/false-name-proof multi-unit auction protocols. In *2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, volume 2, pages 417–423, 2008.

[tim] How timeboost works. https://docs.arbitrum.io/how-arbitrum-works/timeboost/gentle-introduction.

[uni] https://app.uniswap.org/.

[Vic61] William Vickrey. Counterspeculation, auctions, and competitive sealed tenders. *Journal of Finance*, 16(1):8–37, March 1961.

[YNZ25] Sen Yang, Kartik Nayak, and Fan Zhang. Decentralization of ethereum's builder market. In *2025 IEEE Symposium on Security and Privacy (SP)*, page 1512–1530. IEEE, May 2025.

[YSM04] Makoto Yokoo, Yuko Sakurai, and Shigeo Matsubara. The effect of false-name bids in combinatorial auctions: new fraud in internet auctions. *Games and Economic Behavior*, 46(1):174–188, 2004.

[ZLS+25] Mengqian Zhang, Yuhao Li, Xinyuan Sun, Elynn Chen, and Xi Chen. *Maximal Extractable Value in Batch Auctions*, page 510. Association for Computing Machinery, New York, NY, USA, 2025.

[ZQT+21] Liyi Zhou, Kaihua Qin, Christof Ferreira Torres, Duc Viet Le, and Arthur Gervais. High-frequency trading on decentralized on-chain exchanges. In *IEEE Symposium on Security and Privacy*, 2021.

[ZSC+20] Yunhao Zhang, Srinath Setty, Qi Chen, Lidong Zhou, and Lorenzo Alvisi. Byzantine ordered consensus without byzantine oligarchy. In *14th USENIX Symposium on Operating Systems Design and Implementation (OSDI 20)*, 2020.

A Why Exclude Ineligible Users?

Recall that our Pareto optimality (PO) notion is required to hold only for eligible users whose asks are no higher than the initial spot price. We now explain why we do not aim to achieve efficiency for ineligible users. Specifically, we prove that the requirement of UIC precludes achieving PO for all users, including ineligible ones. Li, Shi, and Zhang [LSZ26] proved a similar impossibility but their setting requires that the mechanism's outcome land on the pricing curve without burning (or redistribution) of the surplus. Therefore, their proof cannot be easily adapted to our setting where part of the surplus may be burnt. Our new argument below instead uses the information rents forced by UIC, i.e., the extra utility a user obtains because the mechanism must remain incentive compatible across different private reservation values.

Two-user scenario for the proof. It suffices to consider one sell-Y user i and one sell-X user j. In the notation of the model, user i supplies y_i units of Y and receives x_i units of X, while user j supplies x_j units of X and receives y_j units of Y. Feasibility without an outside subsidy means that there is a signed residual AMM input p and nonnegative surplus-disposal amounts B_Y, B_X such that

$$y_i = y_j + p + B_Y, \qquad x_j + F(p) = x_i + B_X. \tag{35}$$

Here p units of Y are sent to the AMM and $F(p)$ units of X are released. The amounts B_Y and B_X may be burned or redistributed outside the current batch. In particular, (35) only implies

$$y_i \geq y_j + p, \qquad x_i \leq x_j + F(p), \tag{36}$$

and does not require the aggregate user outcome itself to lie on the curve. As elsewhere in the paper, the same residual trade determines the final marginal prices

$$\sigma_e := F'_+(p), \qquad \rho_e := \frac{1}{\sigma_e}.$$

Theorem 24 (UIC precludes Pareto optimality for all users)**.** *Let F be the finite-reserve concave AMM pricing curve defined in Section 2.2. No deterministic AMM mechanism satisfying user individual rationality and the off-curve feasibility condition (35) can simultaneously satisfy user incentive compatibility (UIC) and Pareto optimality for all users, including ineligible users. In fact, the impossibility already holds for two users with fixed, publicly known budgets and with deviations restricted to reservation-value misreports.*

Proof. Consider the two-user scenario mentioned above. Specifically, for each $v \in [v_0, v_1)$, consider a truthful sell-Y user of type (v, Q) and a truthful sell-X user of type (u, R). Denote their respective allocations and compensations by

$$(x_i(v), y_i(v)) \qquad \text{and} \qquad (x_j(v), y_j(v)),$$

and denote the residual AMM input by $p(v)$. We next explain how we choose the parameters v_0, v_1, Q and R. Let

$$0 < v_0 < v_1 < \sigma_0, \qquad \delta := 1 - \frac{v_0}{v_1} > 0.$$

Define the maximum AMM surplus at per-unit cost v_0 by

$$\Gamma(v_0) := \sup_{p \geq 0} \{F(p) - v_0 p\}. \tag{37}$$

Because the pool has finite X reserves, $F(p) < x_0$ for every $p \geq 0$. Consequently $\Gamma(v_0) \leq x_0 < \infty$. Choose the sell-$X$ budget R large enough that the eventual rent lower bound $R\delta$ exceeds this maximum surplus, and then choose the sell-Y budget Q large enough that a full sell-Y allocation cannot be funded (to be shown later):

$$R > \frac{\Gamma(v_0)}{\delta}, \qquad v_0 Q > R + F(Q).$$

The second choice is possible by taking $Q > (R + x_0)/v_0$. Finally, set

$$u := \frac{1}{v_1}.$$

Since $v_1 < \sigma_0$, this choice gives $u = 1/v_1 > 1/\sigma_0 = \rho_0$. Thus the sell-$X$ user with valuation u is ineligible at the initial spot price. Requiring Pareto optimality for this user is precisely what will drive the contradiction.

Finite reserves force the sell-Y user to be underfilled. Suppose, towards a contradiction, that $y_i(v) = Q$. Individual rationality gives

$$x_i(v) \geq vQ \geq v_0 Q. \tag{38}$$

Individual rationality of the sell-X user also gives $y_j(v) \geq u x_j(v) \geq 0$. The first identity in (35) therefore implies $p(v) \leq y_i(v) = Q$. Since $x_j(v) \leq R$ and F is increasing, the second accounting bound gives

$$x_i(v) \leq x_j(v) + F(p(v)) \leq R + F(Q) < v_0 Q,$$

contradicting (38). Consequently,

$$y_i(v) < Q.$$

Pareto optimality fully executes the sell-X order. Because the sell-Y user is underfilled, it could profitably sell an additional amount directly to the pool if $\sigma_e > v$. Pareto optimality for all users rules this out, and hence

$$F'_+(p(v)) = \sigma_e \leq v < \sigma_0.$$

Concavity and differentiability at zero imply $F'_+(p) \geq F'(0) = \sigma_0$ whenever $p \leq 0$. Hence

$$p(v) > 0. \tag{39}$$

Moreover,

$$\rho_e = \frac{1}{F'_+(p(v))} \geq \frac{1}{v} > \frac{1}{v_1} = u.$$

Thus underfilling the sell-Y user pushes the marginal compensation in the reverse direction above the sell-X user's ask. If the sell-X user were underfilled, the strict inequality $\rho_e > u$ would likewise give it a profitable additional trade with the pool, contrary to Pareto optimality for all users. Therefore

$$x_j(v) = R. \tag{40}$$

UIC forces a large sell-X compensation. Fix v and vary only the sell-X user's true and reported ask a over $0 \le a < 1/v$, while keeping its budget equal to R. The preceding underfilling argument does not depend on the numerical sell-X ask. For every such a, the same pool-trade condition therefore forces its allocation to equal R.

Let $y_j(a)$ be its compensation. For any $a, b < 1/v$, the two pairwise UIC inequalities are

$$y_j(a) - aR \ge y_j(b) - aR, \qquad y_j(b) - bR \ge y_j(a) - bR.$$

Because every ask in this interval receives the same full allocation, these inequalities force the compensation to be constant: $y_j(a) = y_j(b)$. Individual rationality gives $y_j(a) \ge aR$ for every $a < 1/v$; taking $a \uparrow 1/v$ shows that this constant is at least R/v. Since the actual ask $u = 1/v_1$ lies in the interval, we obtain

$$y_j(v) \ge \frac{R}{v}. \tag{41}$$

Intuitively, the sell-X user must receive at least R/v units of Y. Those units, together with the positive residual input $p(v)$ sent to the pool, must be supplied by the sell-Y user. Accordingly, combining (36), (39), and (41) gives

$$y_i(v) \ge y_j(v) + p(v) \ge \frac{R}{v} + p(v). \tag{42}$$

UIC forces a large sell-Y information rent. Let

$$U(v) := x_i(v) - v y_i(v)$$

be the sell-Y user's truthful utility. UIC for a user of true type (v_0, Q), comparing truthful reporting with the report (v, Q), implies

$$\begin{aligned} U(v_0) &\ge x_i(v) - v_0 y_i(v) \\ &= U(v) + (v - v_0) y_i(v). \end{aligned}$$

The low-cost type v_0 can mimic type v. UIC therefore requires it to retain the cost advantage $(v - v_0)y_i(v)$ as information rent. Individual rationality gives $U(v) \ge 0$, while (42) gives $y_i(v) \ge R/v$. Hence, for every $v \in (v_0, v_1)$,

$$U(v_0) \ge R\left(1 - \frac{v_0}{v}\right).$$

Taking the supremum as $v \uparrow v_1$ yields

$$U(v_0) \ge R\delta. \tag{43}$$

The forced rent exceeds the AMM's available surplus. Write $p_0 := p(v_0)$ and similarly use a subscript 0 for the other outcomes at v_0. From (42) and (43),

$$\begin{aligned} x_{i,0} &= v_0 y_{i,0} + U(v_0) \\ &\ge v_0 \left(\frac{R}{v_0} + p_0\right) + R\delta \\ &= R + v_0 p_0 + R\delta. \end{aligned}$$

On the other hand, (40) and off-curve feasibility imply

$$x_{i,0} \le R + F(p_0).$$

The lower bound includes the sell-Y user's production cost and its forced information rent, whereas the upper bound is all the X available from the sell-X user and the AMM. Combining them gives

$$F(p_0) - v_0 p_0 \geq R\delta > \Gamma(v_0),$$

whereas $p_0 > 0$ by (39). This contradicts the definition of $\Gamma(v_0)$ in (37).

The argument used only single-identity reservation-value UIC with fixed budgets, which is a restriction of the full UIC requirement. The claimed impossibility follows. □

Remark 25 (The proof only needs local efficiency). The proof does not use the part of Pareto optimality that rules out Pareto improvements through trades among users. It uses only the requirement that no user can trade directly with the pool to obtain a Pareto improvement. This weaker condition is called *local efficiency* (LE) [LSZ26]. If σ_e and $\rho_e = 1/\sigma_e$ are the final marginal compensation rates, LE requires the following to hold for all users i:

$$y_i < \widetilde{q}_i \implies \widetilde{v}_i \geq \sigma_e, \qquad x_j < \widetilde{r}_j \implies \widetilde{u}_j \geq \rho_e.$$

Thus the theorem remains true if Pareto optimality for all users is replaced by LE: UIC is incompatible even with this weaker efficiency requirement.

B Necessity of Surplus Redistribution

We now show that surplus burning is essential: without it, no AMM mechanism can simultaneously provide the full set of guarantees achieved by our mechanism—UIC, strategy-proofness for a builder with or without intrinsic demand, and Pareto optimality for eligible users. We stress, however, surplus redistribution is more than a technical device for circumventing this impossibility. As discussed in Section 1, it offers a compelling new paradigm for provably eliminating MEV in two-asset AMM contracts while substantially reducing the negative externalities of today's highly centralized builder ecosystem [YNZ25].

Li, Shi, and Zhang [LSZ26] showed that, when surplus burning is prohibited, UIC and Pareto optimality (PO) for *all* users cannot be achieved simultaneously. However, as shown in Section A, requiring PO for all users is too stringent: the impossibility persists even when surplus burning is allowed. Our goal here is instead to establish the necessity of surplus burning even under the weaker requirement of PO for *eligible* users only. This distinction calls for a different proof argument described below.

For a concave function, write $F'_-(z)$ and $F'_+(z)$ for its left and right derivatives. Both exist at every interior point, and

$$s < t \quad \implies \quad F'_-(t) \leq F'_+(s). \tag{44}$$

Theorem 26 (Zero burn is impossible)**.** *Let F satisfy assumptions* (A1)–(A4) *of Section 2.2. Suppose that a truthful zero-budget user must receive zero utility, then no deterministic AMM mechanism can simultaneously satisfy individual rationality, feasibility with zero burn, UIC, strategy-proofness for a builder-as-user, and Pareto optimality for eligible bids*

Here zero-burn feasibility means exact token accounting: every token made available by opposite-side orders and the residual AMM trade is paid either to users or to the builder, and the burn amount must be zero.

Proof. Suppose, toward a contradiction, that such a mechanism exists. We use only sell-Y profiles. The key tension is between marginal and average compensation: Pareto optimality determines who

trades by comparing asks with the pool's marginal compensation, whereas zero burn can force a sole trading user to receive the pool's entire output. We will exploit the gap between these two rates to make an otherwise excluded user profit by understating its valuation.

By Theorem 22, strategy-proofness for a builder-as-user, together with IR, feasibility, and UIC, forces the builder fee to be zero on every profile. Since the burn is also zero, a sell-Y-only outcome with total input $Y := \sum_i y_i$ satisfies the exact accounting identity

$$\sum_i x_i = F(Y). \tag{45}$$

Due to Theorem 9, we have the following:

$$y_i = 0 \quad \Longrightarrow \quad x_i = 0 \qquad \text{on a sell-}Y\text{-only profile.} \tag{46}$$

We first record the allocation restrictions imposed by Pareto optimality: cheaper supply must be used first, and users must have no profitable way to increase or unwind their pool trades. Consider a truthful sell-Y-only profile consisting of eligible users and put $Y = \sum_i y_i$.

First, if users i and j have asks $v_i < v_j$, and j sells a positive amount, then i must exhaust its cap. Otherwise, for a sufficiently small $\varepsilon > 0$, user i can give ε units of Y to user j in exchange for $r\varepsilon$ units of X, where $v_i < r < v_j$. Choose ε small enough that user j returns no more than the X and Y involved in its original trade. User i's utility rises by $(r - v_i)\varepsilon$, and user j's utility rises by $(v_j - r)\varepsilon$. The pool position and every other user's outcome are unchanged. This is a Pareto improvement among eligible users. Therefore

$$v_i < v_j,\ y_j > 0 \quad \Longrightarrow \quad y_i = q_i. \tag{47}$$

Second, any user who sells a positive amount must have ask at most the final left marginal compensation:

$$y_i > 0 \quad \Longrightarrow \quad v_i \le F'_-(Y). \tag{48}$$

Indeed, if $v_i > F'_-(Y)$, the user can unwind a sufficiently small amount ε of its trade. It returns $F(Y) - F(Y - \varepsilon)$ units of X and recovers ε units of Y, increasing its utility by

$$v_i\varepsilon - \big(F(Y) - F(Y - \varepsilon)\big) > 0.$$

Exact accounting is preserved and, for small enough ε, the user's net trade remains in its declared direction. Conversely, an underfilled user must satisfy

$$y_i < q_i \quad \Longrightarrow \quad v_i \ge F'_+(Y), \tag{49}$$

because otherwise it could sell a sufficiently small additional amount directly to the pool and receive $F(Y + \varepsilon) - F(Y)$ units of X.

We now choose the witness types using a strict gap between the marginal compensation at input Q and the average compensation over the whole trade. There exists $Q > 0$ such that

$$\sigma := F'_+(Q) < \frac{F(Q)}{Q}.$$

To see this, concavity and $F(0) = 0$ always give $F'_+(Q) \le F(Q)/Q$. If equality held for every $Q > 0$, F would be linear with its positive initial slope on $[0, +\infty)$, contradicting the finite upper reserve bound $F(Q) < x_0$. Moreover, $\sigma > 0$: if the right derivative of a concave increasing function vanished at a finite point, the function would be constant thereafter, contrary to strict increasingness.

Choose numbers

$$0 < v^* < \sigma < v_1 < v_2 < \frac{F(Q)}{Q}. \tag{50}$$

All three asks are strictly eligible, since concavity gives $F(Q)/Q \le F'_+(0) = \sigma_0$. Consider two users, each with cap Q: the first user's true valuation is v_1, and the second user's true valuation is v_2. Write (x_1, y_1) and (x_2, y_2) for their respective compensations and allocations at the profile under consideration. The second user's valuation exceeds the marginal compensation σ, but is below the average compensation $F(Q)/Q$. Thus supplying the whole amount Q would give it positive utility if it received all of $F(Q)$, even though supplying additional units beyond Q would not be worthwhile at the margin. The ask $v^* < \sigma$ will be its dishonest report.

We first show that the second user sells zero under truthful reports: the cheaper first user prevents it from obtaining any allocation. If $y_2 > 0$, then (47) forces $y_1 = Q$. Hence the total input satisfies $Y > Q$, and (44) and (48) give

$$v_2 \le F'_-(Y) \le F'_+(Q) = \sigma,$$

contrary to $v_2 > \sigma$. Thus $y_2 = 0$, and (46) gives $x_2 = 0$. The second user's truthful utility is therefore zero.

Now let the second user deviate only by reporting ask v^*, keeping its cap equal to Q. This reverses the priority: the second user now appears cheaper and, as we show next, becomes the sole supplier. The numerical profile (v_1, v^*) can equally be viewed as a truthful profile, so Pareto optimality applies to its outcome with respect to these reported valuations. If $y_1 > 0$, priority (47) forces $y_2 = Q$. Then $Y > Q$, and the executed first user would have to satisfy

$$v_1 \le F'_-(Y) \le F'_+(Q) = \sigma,$$

again a contradiction. Hence $y_1 = 0$, and (46) gives $x_1 = 0$.

Finally, the second user must sell its full cap under the deviation. Otherwise $Y = y_2 < Q$, while concavity gives

$$F'_+(Y) \ge F'_+(Q) = \sigma > v^*,$$

contradicting the underfilled condition (49). Therefore $y_2 = Q$. At this point, zero burn is decisive: the builder receives no fee, and the inactive first user receives nothing, so the entire AMM output must go to the second user. Exact accounting (45) forces

$$x_2 = F(Q).$$

Evaluated at its true valuation v_2, rather than its report v^*, the second user's deviating utility is

$$F(Q) - v_2 Q > 0$$

by (50), whereas its truthful utility was zero. This contradicts UIC. All reports used in the argument have the same fixed cap, and every ask used in a Pareto comparison is strictly eligible. The claimed impossibility follows. □

Remark 27. If builder fees are prohibited, then the impossibility in Theorem 26 holds without assuming strategy-proofness for a builder-as-user. Indeed, the proof uses builder-as-user strategy-proofness only to invoke the zero-builder-fee theorem (Theorem 22), which is unnecessary when zero builder fees are imposed by assumption.