Key Takeaways:

- LVR is important because 1) it disincentivizes AMM LPs from providing liquidity to the pool, and 2) it leads to the centralization of block building because searcher-builders who are performing CEX-DEX arbitrages can generate far more valuable blocks than neutral builders due to their exclusive orderflow.

- There are broadly two design philosophies to tackle the LVR problem: LVR minimization/reduction and LVR redistribution. LVR minimization/reduction can either be achieved via a low-latency oracle to inform CFMM LP of the current market price or via an LVR rebate parameter that allows arbitrageurs to execute a portion of the CEX-DEX arbitrage trade. LVR redistribution includes the dynamic fee model and the auction model: the dynamic fee model imposes price discrimination on informed and uninformed order flows based on pattern recognition. The auction model allows CEX-DEX arbitrageurs to pay LPs for the right to be the first trade per block at the application level.

- The primary concern with the LVR reduction approach is the reduced orderflow directed to the LVR minimization/reduction pool. This results in LPs incurring greater losses compared to traditional CFMM pools. Instead of merely minimizing LVR, the future seems to be in redistributing LVR back to the LPs, effectively converting LVR into "Gain-versus-rebalancing(GVR)".

- In the realm of LVR redistribution, both the dynamic fee and auction approaches present distinct trade-offs: the dynamic fee model may cause uninformed flows to suffer from higher swap fees during high market volatility, while the auction model presents greater execution challenges, including building a competitive bidding market in a timely manner to allow uninformed traders to trade with the pool.

LVR is important because 1) it disincentivizes AMM LPs from providing liquidity to the pool, and 2) it leads to the centralization of block building because searcher-builders who are performing CEX-DEX arbitrages can generate far more valuable blocks than neutral builders due to their exclusive orderflow.

There are broadly two design philosophies to tackle the LVR problem: LVR minimization/reduction and LVR redistribution. LVR minimization/reduction can either be achieved via a low-latency oracle to inform CFMM LP of the current market price or via an LVR rebate parameter that allows arbitrageurs to execute a portion of the CEX-DEX arbitrage trade. LVR redistribution includes the dynamic fee model and the auction model: the dynamic fee model imposes price discrimination on informed and uninformed order flows based on pattern recognition. The auction model allows CEX-DEX arbitrageurs to pay LPs for the right to be the first trade per block at the application level.

The primary concern with the LVR reduction approach is the reduced orderflow directed to the LVR minimization/reduction pool. This results in LPs incurring greater losses compared to traditional CFMM pools. Instead of merely minimizing LVR, the future seems to be in redistributing LVR back to the LPs, effectively converting LVR into "Gain-versus-rebalancing(GVR)".

In the realm of LVR redistribution, both the dynamic fee and auction approaches present distinct trade-offs: the dynamic fee model may cause uninformed flows to suffer from higher swap fees during high market volatility, while the auction model presents greater execution challenges, including building a competitive bidding market in a timely manner to allow uninformed traders to trade with the pool.

Introduction

There have been continuous discussions around the profitability of liquidity provider (LP) at Constant Function Market Maker (CFMM) across crypto communities and academic researchers. Late last year, [this thread](#) estimated that Uniswap V3 LPs on ETH/USDC pools lost approximately $100 million dollars after fees on top of their original LP holdings by using markout on 5m/1h/24h/7d horizons (it is worthwhile to notice that the data presented by [@thiccythot_](#) could arguably be improved by using pool prices as markout prices as discussed [in this article,](#) which tends to increase the estimates of LP profitability. Nonetheless, it is not a myth that most passive LPs in the CFMM V3 pool are losing money currently). The thread went viral in the DeFi community with many voices [discouraging retail users from passively providing AMM liquidity](#).

Uniswap V3 LPs markout on ETH/USDC pools: https://dune.com/thiccythot/uniswap-markouts

Meanwhile, many researchers took the initiatives to unveil the fundamental reasons behind passive LPs' money-losing game by introducing the concepts of Loss-versus-rebalancing (LVR) and fee liquidity-adjusted instantaneous returns (FLAIR). LVR looks at the problem from the liquidity pool's perspective, quantifying LPs' adverse selection cost due to the sophisticated counterparty (informed trader). FLAIR, on the other hand, looks at the problem from an individual LP's perspective to explain how the competition among LPs affects individual LP performance. These two concepts provide us with a more holistic understanding of the factors hindering passive LPs from making sustainable profits.

In this article, I will pinpoint my focus on LVR. Specifically, I want to explain how LVR leads to an unsustainable DeFi ecosystem due to misaligned incentives and how we can potentially realign those incentives to build a healthier DeFi architecture.

Understanding LVR

LVR refers to CFMM LPs' adverse selection cost due to the information asymmetry between the maker (LP) and taker (informed swapper).

This concept was first brought forth in this paper by Milionis, Moallemi, Roughgarden, and Zhang Adverse selection is a phenomenon that arises when one party has information that the counterparty does not have. In the context of CFMM, adverse selection occurs because high-frequency traders (HFTs) have access to the most up-to-date market prices on centralized exchanges (the most up-to-date market prices, for short-tail tokens at least, are at centralized exchange currently because centralized exchange has the largest amount of volumes) that the passive LPs do not have. Consequently, HFTs are able to exploit passive LPs' information failure by making top-of-the-block arbitrages between CFMM stale prices and CEX prices. From the passive LPs' perspective, they are consistently providing liquidity (quoting) to HFTs at worse-than-market prices, bearing losses when the price slippage is higher than the swap fees they charge to the HFTs. As HFTs perform CEX-DEX arbitrages, asset prices at CFMM will eventually converge to the market prices in a perfectly competitive market disregarding fees. Therefore, LVR can be measured as the gap between CFMM LPs executing risk assets at CFMM stale prices versus at market prices (equal to the top-of-the-block arbitrage profits extracted by HFTs), which essentially is the cost of price discovery for CFMM.

LVR Formula

The following example from the first LVR paper mentioned previously can help us better contextualize LVR: suppose the x and y axes of the graph denote the quantity of X (risky asset) and quantity of Y (USD). Currently, CFMM LPs (the black invariant curve) are holding reserves at point A, corresponding to an instantaneous price Pt represented by the red line. Similarly, there is a rebalancing strategy that holds the same amount of reserves as CFMM LPs but trades at the centralized exchange along the brown line crossing point A. Assume that the CEX price for X is now at Pt + dPt, the arbitrageurs are thus incentivized to trade with CFMM, buying dxt amount of X from CFMM and selling dxt amount of X in CEX, moving CFMM from A to B along the invariant curve. From CLMM LPs' perspective, they sold dxt amount of X at Pt+ dPt/2 (the slope of the scant purple line connecting A and B). If the rebalancing strategy makes the same trade as CFMM LPs but at CEX prices, they will sell dxt amount of X at Pt + dPt without slippage (assuming there is deep enough liquidity on CEX), eventually ending up at point B. *Therefore, the LVR can be figuratively captured by the height between B and B*, quantitatively, it would equal to dxt * ((Pt + dPt) - Pt+ dPt/2), which is also the (maximum) profits captured by arbitrageurs (GVR).

LVR Explanation Graph: https://arxiv.org/abs/2208.06046

Why is this a problem?

In traditional finance, adverse selection is a known risk to market makers. Market makers facilitate trading by always being ready to buy (bid) or sell (ask) securities at publicly quoted prices, thereby providing liquidity to the market. In return, market makers profit from the bid-ask spread but bear the risk of not updating their quoted prices faster than the counterparty either due to information failure or execution speed. Because of this risk, the current market-making landscape has become centralized with a few high-frequency market making firms

that are fast enough and informed enough to not get picked off by other traders.

Therefore, If the LVR issue continues, retail LPs will realize that their strategies become unprofitable and exit. This will cause CFMM to mirror the centralized market-making seen in TradFi, dominated by a few sophisticated LPs using just-in-time (JIT) liquidity. The absence of abundant, affordable liquidity from retail LPs will reduce overall liquidity, impairing trade execution and causing DeFi to lose significant volume to CEX. In essence, for DeFi to remain competitive, it requires efficient trade execution supported by profitable retail LPs.

Additionally, LVR also poses a threat to the infrastructure level as it can lead to the centralization of block builders.

For HFTs to successfully capture CEX-DEX arbitrages, they need to ensure that their arbitrage bundle is at the top of the block. Therefore, HFTs become builders themselves (also known as integrated searcher-builder), organizing their CEX-DEX arbitrage to be included at the top of the block. Currently, CEX-DEX arbitrages account for [60% of the total MEV revenue](#), which means that these integrated searcher-builders are able to construct more valuable blocks than neutral builders. Integrated builders can thus dominate the PBS auction to get their block included by the block proposer (this claim can be empirically supported by [this paper](#)). In another word, integrated builders are able to dominate the PBS auctions because of the exclusive orderflow from their CEX-DEX arbitrages. As such, CEX-aware integrated builders exploit the information failure of CFMM LPs to generate disproportionately valuable blocks, which leads to the centralization of block builders

.

LVR Value Chain

Potential Solutions

The nature of the LVR problem lies in the value leaked from the application level to the infrastructure level

. Since integrated builders need to pay bribe fees to the proposer to win the PBS auction, CFMM LP's adverse selection cost (value leaked from the application level) goes to HFTs and proposers (it is estimated that [37-77% of CEX-DEX arbitrage profits go to the validators currently](#)). This creates a misalignment of interest within the application level, namely HFTs and proposers are gaining profits at the expense of LPs, leading to CFMM LPs being unincentivized to participate.

Solution Overview

Thus, the LVR solution design generally takes in two forms:

- LVR reduction/minimization: this form has two lines of approaches:

-
    1. Eliminate the value leaked (LP's adverse selection cost) by cutting the need for informed/toxic flow to help with CFMM's price discovery. This line of approach generally employs a low-latency oracle to inform CFMM LP of the current market price, solving the information failure at its root.

- Eliminate the value leaked (LP's adverse selection cost) by cutting the need for informed/toxic flow to help with CFMM's price discovery. This line of approach generally employs a low-latency oracle to inform CFMM LP of the current market price, solving the information failure at its root.

-
    1. Minimize the LP's adverse selection cost by only allowing arbitrageurs to execute a portion of the CEX-DEX arbitrage trade by introducing an LVR rebate parameter. The designated AMM pool price still aligns with the external price by removing corresponding tokens from the pool to a vault. The attained tokens in the vault will be gradually re-added back to the pool in subsequent blocks.

- Minimize the LP's adverse selection cost by only allowing arbitrageurs to execute a portion of the CEX-DEX arbitrage trade by introducing an LVR rebate parameter. The designated AMM pool price still aligns with the external price by removing corresponding tokens from the pool to a vault. The attained tokens in the vault will be gradually re-added back to the pool in subsequent blocks.

-
    1. Eliminate the value leaked (LP's adverse selection cost) by cutting the need for informed/toxic flow to help with

CFMM's price discovery. This line of approach generally employs a low-latency oracle to inform CFMM LP of the current market price, solving the information failure at its root.

- Eliminate the value leaked (LP's adverse selection cost) by cutting the need for informed/toxic flow to help with CFMM's price discovery. This line of approach generally employs a low-latency oracle to inform CFMM LP of the current market price, solving the information failure at its root.

-
    1. Minimize the LP's adverse selection cost by only allowing arbitrageurs to execute a portion of the CEX-DEX arbitrage trade by introducing an LVR rebate parameter. The designated AMM pool price still aligns with the external price by removing corresponding tokens from the pool to a vault. The attained tokens in the vault will be gradually re-added back to the pool in subsequent blocks.

- Minimize the LP's adverse selection cost by only allowing arbitrageurs to execute a portion of the CEX-DEX arbitrage trade by introducing an LVR rebate parameter. The designated AMM pool price still aligns with the external price by removing corresponding tokens from the pool to a vault. The attained tokens in the vault will be gradually re-added back to the pool in subsequent blocks.

- LVR redistribution: let LPs capture part of the LVR to realign the interests of LPs. This line of approach generally takes in two forms:

-
    1. A dynamic fee model that charges more fees when there is more informed flow (similar to how in traditional finance market makers often widen the bid-ask spread during periods of high volatility).

- A dynamic fee model that charges more fees when there is more informed flow (similar to how in traditional finance market makers often widen the bid-ask spread during periods of high volatility).

-
    1. An auction that requires HFTs to pay a fee to be the first trade per block, redirecting the value leaked to the proposer back to the LPs.

- An auction that requires HFTs to pay a fee to be the first trade per block, redirecting the value leaked to the proposer back to the LPs.

-
    1. A dynamic fee model that charges more fees when there is more informed flow (similar to how in traditional finance market makers often widen the bid-ask spread during periods of high volatility).

- A dynamic fee model that charges more fees when there is more informed flow (similar to how in traditional finance market makers often widen the bid-ask spread during periods of high volatility).

-
    1. An auction that requires HFTs to pay a fee to be the first trade per block, redirecting the value leaked to the proposer back to the LPs.

- An auction that requires HFTs to pay a fee to be the first trade per block, redirecting the value leaked to the proposer back to the LPs.

LVR reduction/minimization: this form has two lines of approaches:

-
    1. Eliminate the value leaked (LP's adverse selection cost) by cutting the need for informed/toxic flow to help with CFMM's price discovery. This line of approach generally employs a low-latency oracle to inform CFMM LP of the current market price, solving the information failure at its root.

- Eliminate the value leaked (LP's adverse selection cost) by cutting the need for informed/toxic flow to help with CFMM's price discovery. This line of approach generally employs a low-latency oracle to inform CFMM LP of the current market price, solving the information failure at its root.

-
  1. Minimize the LP's adverse selection cost by only allowing arbitrageurs to execute a portion of the CEX-DEX arbitrage trade by introducing an LVR rebate parameter. The designated AMM pool price still aligns with the external price by removing corresponding tokens from the pool to a vault. The attained tokens in the vault will be gradually re-added back to the pool in subsequent blocks.

- Minimize the LP's adverse selection cost by only allowing arbitrageurs to execute a portion of the CEX-DEX arbitrage trade by introducing an LVR rebate parameter. The designated AMM pool price still aligns with the external price by removing corresponding tokens from the pool to a vault. The attained tokens in the vault will be gradually re-added back to the pool in subsequent blocks.

- Eliminate the value leaked (LP's adverse selection cost) by cutting the need for informed/toxic flow to help with CFMM's price discovery. This line of approach generally employs a low-latency oracle to inform CFMM LP of the current market price, solving the information failure at its root.

- Minimize the LP's adverse selection cost by only allowing arbitrageurs to execute a portion of the CEX-DEX arbitrage trade by introducing an LVR rebate parameter. The designated AMM pool price still aligns with the external price by removing corresponding tokens from the pool to a vault. The attained tokens in the vault will be gradually re-added back to the pool in subsequent blocks.

LVR redistribution: let LPs capture part of the LVR to realign the interests of LPs. This line of approach generally takes in two forms:

-
  1. A dynamic fee model that charges more fees when there is more informed flow (similar to how in traditional finance market makers often widen the bid-ask spread during periods of high volatility).

- A dynamic fee model that charges more fees when there is more informed flow (similar to how in traditional finance market makers often widen the bid-ask spread during periods of high volatility).

-
  1. An auction that requires HFTs to pay a fee to be the first trade per block, redirecting the value leaked to the proposer back to the LPs.

- An auction that requires HFTs to pay a fee to be the first trade per block, redirecting the value leaked to the proposer back to the LPs.

- A dynamic fee model that charges more fees when there is more informed flow (similar to how in traditional finance market makers often widen the bid-ask spread during periods of high volatility).

- An auction that requires HFTs to pay a fee to be the first trade per block, redirecting the value leaked to the proposer back to the LPs.

LVR Solution Design Landscape

LVR Minimization: Diamond Protocol

The Diamond Protocol was initially presented in this paper, followed by a proof-of-concept demonstration in an Ethresearch post that differs in implementation. Recently, Arrakis Finance announced plans to adopt this concept using hooks. Given the variations in implementation across these sources, my discussion will center on the Arrakis Diamond Hook, accessible on this GitHub page.

In the Diamond Protocol, builders must first commit to the end-of-the-block price at the start of the block, which should equate to the LVR maximizing price (i.e., the external market price), before any swaps can occur. Then, the Diamond protocol will update the pool price to the committed price via the openPool()

function in the Hook contract. Specifically, this function will calculate the amount of trade size X required to move the starting pool price to the committed price. However, the actual trade size the builder can execute is discounted based on the LVR-rebate parameter, β, which ranges between 0 and 1. For instance, if β is 0.75, the builder can only execute a trade size

of (1-0.75)X = 0.25X. To move the price in the pool to the committed price as if the entire trade size X is executed, the diamond protocol will remove some tokens from the pool to a vault contract. The tokens in the vault will be gradually re-added back to the pool per block at a percentage between 1-5%. It decides to slowly rebate the vault token back to the pool because re-adding tokens from the pool to the vault creates an expected arbitrage opportunity in the next block. By re-adding fewer tokens, the losses will be reduced, which increases the profitability of the pool.

After the openPool()

function is completed, builders are required to post collateral to the hedger contract before any swap can occur. The deposited collateral must be sufficient to adjust the pool's price back to the committed price at any given point. In other words, the swap size should not exceed the tokens available in the hedger

contract. This is important because If, by the end of the block, the pool's price doesn't return to the committed price, there are enough tokens in the hedger

contract to bring the pool back to balance. The tokens in the hedger

are used to adjust the pool's price in the next block when the pool is accessed.

The crux of how diamond protocol redistributes LVR back to the LPs lies in the collection and gradual readdition of LVR rebates back to the pool. When the vault fills up with LVR rebates, these rebates represent retained tokens from LP positions that would otherwise have been adversely selected against by the top-of-the-block arbitrage.

Every time the vault accumulates a certain amount of tokens, denoted as Xi for blocki, the percentages of Xi owed to the respective LPs are directly proportional to the percentage of liquidity that each LP provided over the pool move in blocki. Over time, these rebated tokens are gradually reintroduced into the pool.

However, there are also potential concerns and issues stemming from the design of the Diamond Protocol (some of these observations are based on comments from [this Twitter post](#)):

- Bounded order size in Diamond Protocol leading to LP's unprofitability

: Orders executed in the diamond pool must be smaller than the available tokens in the hedger contract to ensure rebalancing to the committed price. This limits the order size, potentially reducing order flow and fee earnings for LPs compared to a vanilla CFMM pool.

- Volatility throughout the block can go adversely to builder's committed price at the start of the block

: because the builder need to commit to the end-of-the-block price at the start of the block, any external market price fluctuations during the block can cause the pool price to diverge from the committed price. Given the volatility of crypto market, the likelihood of this happen will be quite high. In such cases, builders will be forced to take the other side of the trade, as they have an obligation to align the pool price with their committed price at the end of the block. This could result in losses for the builders, potentially disincentivizing them from participating in the diamond pool in the first place.

- Potential execution delay due to the exclusive reliance on a single block builder

: If the designated builder fails to secure a win in the PBS auction, it could result in transaction delays within the Diamond pool. This becomes even more concerning considering the committed builder's reduced competitiveness in the PBS auction, given that they can only capture a fraction of the implied LVR.

- Technical specification including which builder will be granted the exclusive right to perform the executing and ordering of the block.

The document did not specify which builder would be granted the exclusive right to perform the executing and ordering of the block. Will this depend on how much collateral they are willing to put? (this is also another question that remains to be answered which is how much collateral the builder needs to put).

Bounded order size in Diamond Protocol leading to LP's unprofitability

: Orders executed in the diamond pool must be smaller than the available tokens in the hedger contract to ensure

rebalancing to the committed price. This limits the order size, potentially reducing order flow and fee earnings for LPs compared to a vanilla CFMM pool.

Volatility throughout the block can go adversely to builder's committed price at the start of the block

: because the builder need to commit to the end-of-the-block price at the start of the block, any external market price fluctuations during the block can cause the pool price to diverge from the committed price. Given the volatility of crypto market, the likelihood of this happen will be quite high. In such cases, builders will be forced to take the other side of the trade, as they have an obligation to align the pool price with their committed price at the end of the block. This could result in losses for the builders, potentially disincentivizing them from participating in the diamond pool in the first place.

Potential execution delay due to the exclusive reliance on a single block builder

: If the designated builder fails to secure a win in the PBS auction, it could result in transaction delays within the Diamond pool. This becomes even more concerning considering the committed builder's reduced competitiveness in the PBS auction, given that they can only capture a fraction of the implied LVR.

Technical specification including which builder will be granted the exclusive right to perform the executing and ordering of the block.

The document did not specify which builder would be granted the exclusive right to perform the executing and ordering of the block. Will this depend on how much collateral they are willing to put? (this is also another question that remains to be answered which is how much collateral the builder needs to put).

LVR Reduction: Oracle-Informed AMM

Oracle-informed AMM can employ a low-latency oracle to inform CFMM LPs of the current market price before the CEX-DEX arbitrage. Specifically, it can work by including an active liquidity manager module that adjusts its liquidity concentration dynamically to the updated oracle price (like an automated JIT) before the CEX-DEX arbitrages. This is also called a [dealer-like](#) model of liquidity provisioning.

For oracle-informed AMM to prevent toxic orderflow, the oracle needs to be fast enough and accurate enough to prevent CEX-DEX arbitrageurs. Specifically, the speed of the oracle price and the subsequent adjustment of liquidity provision price should be before the CEX-DEX arbitrage. Additionally, the accuracy of the oracle price needs to be smaller than the pool's swap fee. This is because an arbitrage is unprofitable whenever fees made from the trade are smaller than the price different from the market price. In this model, instead of acting as passive market makers who are waiting to get their stale price picked up, CFMM LPs become active market makers who have the same amount of information as the HFTs to actively provide quote prices compatible with the market prices.

Oracle-Informed AMM

However, this model, if implemented without modification, can lead to unbounded losses for LPs in the event of inaccurate oracle-fed prices due to oracle manipulation or oracle liveness issues.

Since oracle-informed AMM loses its internal price discovery mechanism, it becomes a secondary price discovery platform that is vulnerable to oracle manipulation/shut-down. Oracle manipulation involves intentional interference with the data source, skewing the accuracy of the prices relayed. Additionally, if the oracle experiences downtime or liveness issues, it can't provide real-time price updates. Therefore, when the oracle is down or provides an inaccurate price due to oracle manipulation, LPs will face an unbounded loss because they will aggregate their liquidity based on these flawed prices.

Given this concern, oracle-informed AMM should implement relevant modifications to cap the potential unbounded loss.

One approach to achieve this is by limiting the trade size executable within the pool, similar to [how GMX caps the long/short open interest](#). However, this method introduces a ceiling to the growth of the protocol. An alternative modification could involve using oracle-fed information to guide the pool's fee instead of the price. This brings us to the next topic of discussion: the dynamic fee model.

LVR Redistribution: Dynamic Fee

The underlying philosophy of the dynamic fee model is to exert price discrimination

upon informed and uninformed orderflow based on pattern recognition. It needs to 1) utilize all sorts of signals to derive a pattern that can differentiate informed and uninformed orderflow and then 2) charge a higher fee to toxic orderflow and a lower fee to retail orderflow.

In theory, the dynamic fee model redistributes part of the LVR back to the LPs because LPs can charge a higher fee for CEX-DEX arbitrages. In other words, CFMM LPs under the dynamic fee model can widen their bid-ask spread to informed traders and tighten their bid-ask spread to uninformed traders.

One example of the dynamic fee model is proposed by Alex Nezlobin in this thread. It recognizes an autocorrelated trading pattern of toxic flow that tends to keep picking on the ask side of the AMM if the market price pushed the AMM's ask price in the previous block.

To illustrate this, assume an AMM pool of ETH/USDC and ETH market price increases that jump outside of the AMM's bid-ask spread. This will attract CEX-DEX arbitrages to buy ETH at AMM and sell ETH at CEX, making AMM's best ask price equals to the CEX price. From the arbitragers' perspective, they are more likely to continue buying ETH (pick on the ask side) than selling ETH (pick on the bid side) in the AMM. This is because it is more likely to profit from buying ETH than selling ETH considering the CEX price only needs to increase by a marginal point for buying ETH to be profitable but the CEX price needs to decrease by 2f for selling ETH to profit.

Compared with the autocorrelated trading pattern of informed flow, the trading direction of uninformed flow should be fairly random. Therefore, if the AMM price increases by Δ in block t, the model will raise the fee for buys and lower the fee for sells by $\delta = c\Delta$ for $c > 0$ (to make neither fee negative) at block t+1, making informed flow pay more.

Graphic Representation of Alex Nezlobin's Dynamic Fee Model: https://twitter.com/0x94305/status/1674857993740111872

Empirically, Alex ran a Python simulation that changed fees by 0.75 if the AMM price went up in the previous block in the ETH/USD pool ($50,000 liquidity per basis point, 5 bps fee, 5% daily volume). The result shows that the losses of LPs in the dynamic fee model are around 10% lower

than with the fixed fee model depending on the gas fee per swap.

Alex Nezlobin's Dynamic Fee Model Simulation Result: https://twitter.com/0x94305/status/1674858006205571072

Alex's model demonstrates a neat way to exert price discrimination without using an oracle. However, there is still some downside including the arbitrageur who caused the first strong price shift will not be price discriminated against and the uninformed flow also faces this higher fee if they want to pick on the ask side. To further improve upon Alex's model, there are other signals that AMMs can use to differentiate informed and uninformed flow to a greater extent (some of the following signals were discussed in the LVR Reduction Twitter Space):

- Oracle-fed External world information

: this refers to the information that CFMM passive LPs are unaware of, including market price, market volatility, etc. This information can be fed to the AMM using a low-latency oracle. For example, since LVR is greater when volatility is high (CEX-DEX price slippage is higher under high volatility), we can use an oracle that feeds us with updated market volatility signals to increase the swap fees when market volatility is high or vice versa. Ambient Finance currently uses this model. This line of approach minimizes the effect of oracle compared to the oracle-informed AMM model because it still maintains internal price discovery that does not have a liveness issue. The problem with this approach is that high-frequency price movements on CEX can occur intra-block, rendering the oracle's snapshot of the market outdated by the time it's available for the AMM's decision-making. Consequently, this misalignment creates a window of opportunity for arbitrageurs, as they can capitalize prior to an oracle update. Additionally, this could also lead to a fall in overall volumes because it neutralizes the increased fees to both informed and uninformed flow, discouraging them from participating in this pool at high volatility and choosing fixed-price pools instead, leading to a welfare loss to all.

- Source of orderflow:

this refers to the channel in which the orderflow is routed from (e.g., CoWSwap, UniswapX, Metamask, etc.). In traditional finance, Robinhood monetizes by selling its orderflow to high-frequency market makers because most orderflow coming from

Robinhood are retail flow. Similarly, in crypto, the source of the ordeflow can also be an important signal to differentiate informed and uninformed flow. One example is CoWSwap: CoWswap runs a batch auction that introduces a delay that makes them unattractive for toxic flow. Therefore, AMM pools could lower the prices for orderflow routed from CoWSwap solvers as there is a high possibility that the orderflow is uninformed. Currently, Balancer reduces the 50-75% swap fees for all orderflow coming from CoWSwap. In the future, there could be a reputation system for all the relayers based on the toxicity of their orderflow and exert price discrimination based on that

- Existing orderflow

: This refers to the trading pattern observed from the existing orderflow that can be used to differentiate informed and uninformed flow. One example is Alex Nezlobin's dynamic fee model as discussed above: it discovers a trading pattern that informed flow has a higher possibility of picking on the ask side than on the bid side if the market price pushed the AMM's ask price in the previous block, which allows the model to exert price discrimination accordingly. Additionally, one can employ a co-processor to perform computation based on the history of on-chain transactions to differentiate informed and uninformed flow. For example, the protocol can leverage Axiom to analyze the historical record of on-chain transactions. Axiom provides trustless access to all on-chain data and offers verified compute capabilities over it. This means that developers can query historic data from Ethereum via Axiom, and Axiom will trustlessly fulfill ZK-verified results on-chain. By doing so, the dynamic fee model can further differentiate between informed and uninformed transaction flows without introducing additional trust assumptions.

Oracle-fed External world information

: this refers to the information that CFMM passive LPs are unaware of, including market price, market volatility, etc. This information can be fed to the AMM using a low-latency oracle. For example, since LVR is greater when volatility is high (CEX-DEX price slippage is higher under high volatility), we can use an oracle that feeds us with updated market volatility signals to increase the swap fees when market volatility is high or vice versa. Ambient Finance currently uses this model. This line of approach minimizes the effect of oracle compared to the oracle-informed AMM model because it still maintains internal price discovery that does not have a liveness issue. The problem with this approach is that high-frequency price movements on CEX can occur intra-block, rendering the oracle's snapshot of the market outdated by the time it's available for the AMM's decision-making. Consequently, this misalignment creates a window of opportunity for arbitrageurs, as they can capitalize prior to an oracle update. Additionally, this could also lead to a fall in overall volumes because it neutralizes the increased fees to both informed and uninformed flow, discouraging them from participating in this pool at high volatility and choosing fixed-price pools instead, leading to a welfare loss to all.

Source of orderflow:

this refers to the channel in which the orderflow is routed from (e.g., CoWSwap, UniswapX, Metamask, etc.). In traditional finance, Robinhood monetizes by selling its orderflow to high-frequency market makers because most orderflow coming from Robinhood are retail flow. Similarly, in crypto, the source of the ordeflow can also be an important signal to differentiate informed and uninformed flow. One example is CoWSwap: CoWswap runs a batch auction that introduces a delay that makes them unattractive for toxic flow. Therefore, AMM pools could lower the prices for orderflow routed from CoWSwap solvers as there is a high possibility that the orderflow is uninformed. Currently, Balancer reduces the 50-75% swap fees for all orderflow coming from CoWSwap. In the future, there could be a reputation system for all the relayers based on the toxicity of their orderflow and exert price discrimination based on that

Existing orderflow

: This refers to the trading pattern observed from the existing orderflow that can be used to differentiate informed and uninformed flow. One example is Alex Nezlobin's dynamic fee model as discussed above: it discovers a trading pattern that informed flow has a higher possibility of picking on the ask side than on the bid side if the market price pushed the AMM's ask price in the previous block, which allows the model to exert price discrimination accordingly. Additionally, one can employ a co-processor to perform computation based on the history of on-chain transactions to differentiate informed and uninformed flow. For example, the protocol can leverage Axiom to analyze the historical record of on-chain transactions. Axiom provides trustless access to all on-chain data and offers verified compute capabilities over it. This means that developers can query historic data from Ethereum via Axiom, and Axiom will trustlessly fulfill ZK-verified results on-chain. By doing so, the dynamic fee model can further differentiate between informed and uninformed transaction flows without introducing additional trust assumptions.

Ideally, the most efficient dynamic fee model should combine all these signals to differentiate informed and uninformed flow to the highest probability.

However, the dynamic fee model, even if it intakes all sorts of signals, is probabilistic in nature in differentiating informed and uninformed flow. Therefore, it still neutralizes its increased fees to both the informed and uninformed traders to a certain extent. The consequence of this is that uninformed flow might be discouraged from participating in these pools at times of high volatility (higher fees). This could be a concern under the adoption of intent-based architecture because solvers will be competing to generate the lowest possible price for the users, begging the question of how much orderflow will be routed to dynamic-fee AMM pools at times of higher fees.

Additionally, a dynamic fee model naturally introduces additional friction

to informed traders because the cost of arbitrage is higher at times of higher volatility (swap fee is higher), which can make the price discovery of the AMM pool less efficient.

LVR Redistribution: Auction

Different from the dynamic fee model that extracts value from the top-of-the-block arbitragers, the auction aims to redistribute LVR back to LPs from the hands of proposers

. The reason why the value was leaked to the proposer was that HFT integrated builders need to compete in the PBS auction for block inclusion for their top-of-the-block arbitrage, hence paying priority fees to the proposers. Therefore, the gist of the auction model is to shift the power of being at the top-of-the-block to the AMM application level, thereby HFTs, instead of paying priority fees to the proposers, pay auction to the AMM for the right of the first trade.

One example, as proposed by Josojo in Ethresearch, is to auction off the right to be the first trade per block at the application's smart contract level. The winning bid for the auction is then distributed back to the LPs in the AMM pools. This design idea is similar to CEX provides faster execution speed for their VIP-tier customers.

https://ethresear.ch/t/mev-capturing-amm-mcamm/13336

Specifically, it introduces a global contract that oversees the first execution right for all McAMMs. This contract, optimized for gas efficiency, ideally functions as the standard router for all McAMM pools. The designated party with the first execution right, termed the "leadsearcher," is identified by this global contract. The contract maintains a record, per block, of whether the leadsearcher has traded on any of the deployed AMM contracts. Other participants are permitted to trade with the McAMM only after the leadsearcher's trade; otherwise, trades are reverted.

To acquire the leadsearcher status, one must win the auction for the first execution right, structured similarly to the Eden network's auction for the block's initial slot. Leadsearchers are automatically deselected by the router contract if they miss more than three blocks, preventing prolonged trade blockages. If leadsearchers opt not to trade in a block, they must still signal their interaction with the McAMMs, enabling others to trade against them. This signaling incurs a minimal gas cost for the leadsearcher. To ensure the leadsearcher captures all arbitrage opportunities, they are exempted from standard AMM fees.

Theoretically, block-builders' interests are also aligned with this specification because they are motivated to propose blocks with maximal priority fee gas consumption, enhancing their chances of winning the MEV boost auction.

In other words, block builders want the transactions to consume as much gas as possible so that their bidding price in the PBS auction is higher. Therefore, if the builders put other transactions in front of the leadsearcher's transaction, those transactions would revert, leading to less gas and hence fewer fees. Therefore, block-builders are incentivized to ensure the correct trade sequence, prioritizing the leadsearcher's transaction. Moreover, McAMM users are encouraged to exclusively broadcast their trades to block-builders who respect the McAMM's enforced ordering, avoiding the costs of failed transactions. These combined incentives are expected to drive all block-builders to adhere to the necessary ordering for McAMMs.

On the other hand, there are some potential problems and further design specifications needed in the auction mechanism:

- How to allocate the auction fee to the LPs fairly

: One question with the McAMM design and auction design in general is how to allocate the auction fee among LPs across McAMM pools. Josojo discusses some possibility around accumulating the auction fee to a DAO that subsequently distribute the auction fees to the relevant parties.

- How to construct a competitive market for bidders

: in order for LPs to capture the maximal amount of LVR value possible, we need to construct a competitive bidding market. This depends on when should the auction take place and finish. Currently, there are broadly two times where an auction can take place: [ex-ante where the leadsearcher for BlockN is determined in BlockN-1 prior to the CEX price movement ,and ex-post where the leadsearcher for BlockN is determined after the CEX price movement but before BlockN is committed.](#) Compared with the ex-ante auction, the ex-post auction has a relatively lower barrier of entry

because searchers know how much CEX-DEX arbitrage value they can capture whereas searchers in ex-ante auction need to have option pricing ability because the volatility is not realized. Therefore, an ex-post auction should have a more competitive bidding market than the ex-ante auction. On the other hand, the ex-post auction has a shorter time to determine the leadsearcher compared with the ex-ante auction because the auction begins after the CEX price movement and needs to finish no later than the PBS auction. Thus, there seems to be a tradeoff in establishing a more time-contingent ex-post auction that could lead to a more competitive bidding market versus an ex-ante auction that could have a more buffer time but the higher barrier of entry for bidders. Additionally, how to construct a censorship-resistant auction mechanism (e.g., [first-price sealed auction with a commit reveal scheme](#)) within such a short period of time while attracting a significant amount of bidders could be challenging too.

- On-chain auction vs off-chain auction

: whether we would want to run the auction on-chain or off-chain is another important metric that worth further consideration. In general, on-chain auction smart contracts offer the advantage of reduced overhead costs compared to establishing an off-chain trust-minimized entity for auctions. However, they introduce two challenges:1) gas costs and 2) the uncertainty tied to latency, which is influenced by network congestion

. The imposition of gas costs on every bid[effectively adjusts the de facto bid value](#). Since gas prices can fluctuate throughout the auction's duration, a potential highest bidder might be sidelined due to prohibitive gas fees.

The uncertainty tied to latency boils down to the time lag between a bid's broadcast and its blockchain inclusion. If this bid is relayed to a public mempool rather than a private transaction pool, it can trigger a [gas bidding war](#) on top of the auction, leading to further complexity. These reasons are why current on-chain auctions [struggle to scale with an increasing number of bidders, leading to network bottlenecks, higher transaction costs, and extended confirmation times](#). The off-chain auction, on the other hand, suffers from the execution complexity around building a censorship-resistant and fair off-chain entity but benefits from its independence without worrying about the gas price cost and the associated uncertainty in latency due to potential network congestion

. [Sorella Lab](#) is designing an ex-post off-chain auction that strikes a good balance between these tradeoffs.

How to allocate the auction fee to the LPs fairly

competitive bidding market than the ex-ante auction. On the other hand, the ex-post auction has a shorter time to determine the leadsearcher compared with the ex-ante auction because the auction begins after the CEX price movement and needs to finish no later than the PBS auction. Thus, there seems to be a tradeoff in establishing a more time-contingent ex-post auction that could lead to a more competitive bidding market versus an ex-ante auction that could have a more buffer time but the higher barrier of entry for bidders. Additionally, how to construct a censorship-resistant auction mechanism (e.g., [first-price sealed auction with a commit reveal scheme](#)) within such a short period of time while attracting a significant amount of bidders could be challenging too.

On-chain auction vs off-chain auction

: whether we would want to run the auction on-chain or off-chain is another important metric that worth further consideration. In general, on-chain auction smart contracts offer the advantage of reduced overhead costs compared to establishing an off-chain trust-minimized entity for auctions. However, they introduce two challenges:1) gas costs and 2) the uncertainty tied to latency, which is influenced by network congestion

. The imposition of gas costs on every bid[effectively adjusts the de facto bid value](#). Since gas prices can fluctuate throughout the auction's duration, a potential highest bidder might be sidelined due to prohibitive gas fees.

The uncertainty tied to latency boils down to the time lag between a bid's broadcast and its blockchain inclusion. If this bid is relayed to a public mempool rather than a private transaction pool, it can trigger a [gas bidding war](#) on top of the auction, leading to further complexity. These reasons are why current on-chain auctions [struggle to scale with an increasing number of bidders, leading to network bottlenecks, higher transaction costs, and extended confirmation times](#). The off-chain auction, on the other hand, suffers from the execution complexity around building a censorship-resistant and fair off-chain entity but benefits from its independence without worrying about the gas price cost and the associated uncertainty in latency due to potential network congestion

. [Sorella Lab](#) is designing an ex-post off-chain auction that strikes a good balance between these tradeoffs.

Conclusion

Overall, the primary concern with the LVR reduction approach is the reduced orderflow directed to the LVR minimization/reduction pool. This results in LPs incurring greater losses compared to traditional CFMM pools. Instead of merely minimizing LVR, the future seems to be in redistributing LVR back to the LPs, effectively converting LVR into "Gain-versus-rebalancing(GVR)".

In the LVR redistribution realm, both the auction and dynamic fee models present distinct trade-offs. The auction mechanism can more deterministically differentiate between informed and uninformed flows, ensuring that the uninformed flows aren't subjected to higher fees as they would under the dynamic fee model. However, the auction approach comes with non-trivial execution challenges, as outlined above. Additionally, auction design must strike the right balance between building a competitive bidding market and ensuring the auction concludes quickly, allowing uninformed traders enough time to swap in the AMM pool before the next auction begins. I envision the future of auctions as off-chain, ex-post auctions settled on a sidechain or suave.

Another consideration is the amount of value that can be redistributed back to the LPs. In the auction model, the value that LPs can extract is contingent upon the competitiveness of the bidding market. Meanwhile, in the dynamic fee model, the key factor is its ability to differentiate between informed and uninformed flows, ensuring that it doesn't levy higher fees on the latter, which otherwise could deter uninformed traders from interacting with the pool. With the adoption of zk-coprocessor, I believe it can aid the current oracle-based dynamic fee approach to differentiate informed and uninformed flows to a higher extent by querying on-chain data in a trust-minimized way.

All in all, I am very excited to see this new evolution of AMM DEX design to solve the adverse selection cost of LPs. This could further motivate passive LPs to join AMM, potentially offering execution prices that can rival the off-chain CEX platforms.