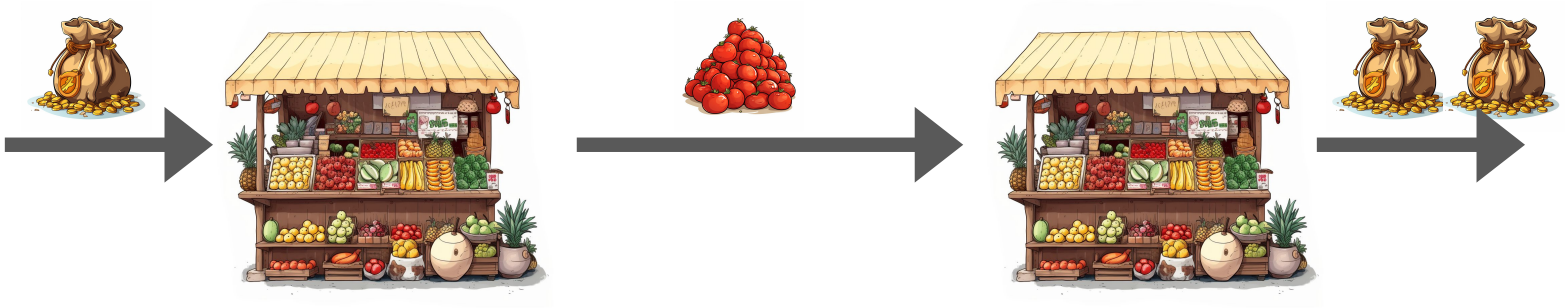
A Large Scale Study of the Ethereum Arbitrage Ecosystem

Robert McLaughlin, Christopher Kruegel, Giovanni Vigna

University of California, Santa Barbara

Primer: What is Arbitrage?

- **Simple definition**: Arbitrage is buying an asset in one market and selling it into another with advantageously differing prices
- This is a *normal and expected* facet of financial markets
- Arbitrage moves assets from abundant to scarce markets

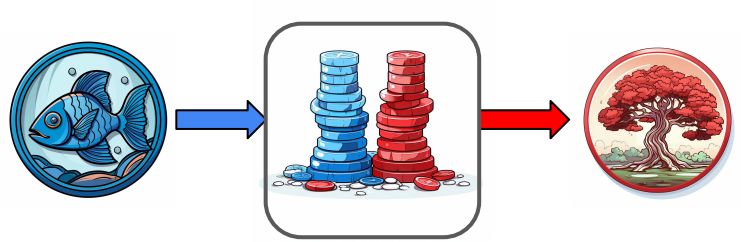


Buy
$2 / tomato

Sell
$4 / tomato

Automated Market Makers: An Automated Exchange

- Smart contracts on the blockchain (Ethereum)
- The contract maintains a *liquidity pool*
  - Reserves of (at least) two tokens
- To *swap* ERC-20 tokens:
  - Pay into the pool
  - The pool calculates a fair price, and subtracts fees
  - Proceeds are sent back to the user
- Users may *provide liquidity* to the pool in order to earn fees



$$(b_{\text{in}} + \delta_{\text{in}}) \cdot (b_{\text{out}} - \delta_{\text{out}}) \geq b_{\text{in}} \cdot b_{\text{out}}$$

swap invariant: constant-product market maker

Atomic Arbitrage: Free Money?

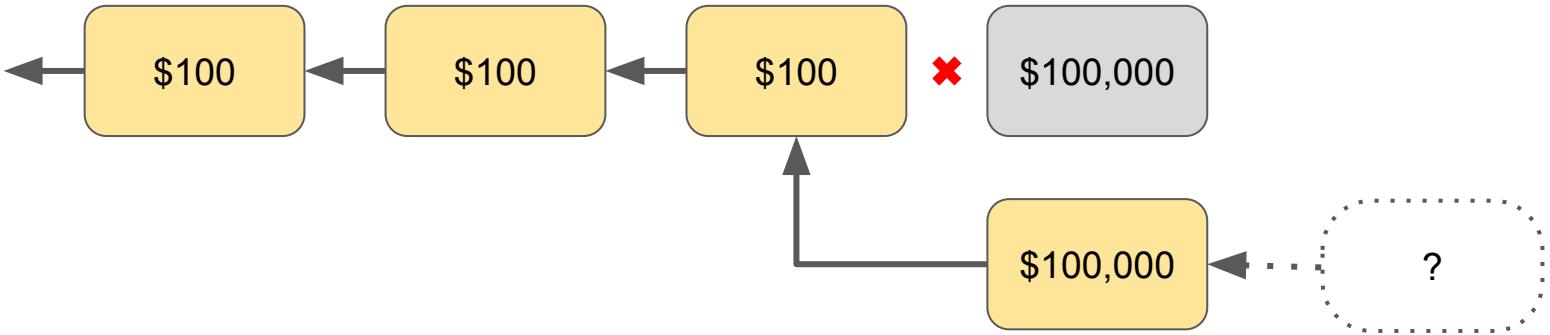
- The Ethereum blockchain has many independent AMM applications
  - Uniswap, SushiSwap, Balancer, Bancor, 1inch, ...
  - Prices update when users swap tokens, independently of each other
- Limited risk: Ethereum transactions are atomic
  - If profit is not achieved the transaction can be *reverted*
- Bots arbitrage between these markets for profit [Daian 19]
- Highly competitive, total annual profits around $100m [Daian 19, Torres 21, Qin 21]

Motivating Concerns: MEV

- [**M**iner / **M**aximal / Block] **E**xtractable **V**alue, MEV
  - Enhances a block producer's reward
  - Selectively censoring, inserting, and/or reordering transactions
  - Arbitrage contributes to MEV
- Excessive MEV incentivizes a *Time Bandit Attack* [Daian 19, Qin 21]

Motivating Concerns: MEV

- [**M**iner / **M**aximal / Block] **E**xtractable **V**alue, MEV
  - Enhances a block producer's reward
  - Selectively censoring, inserting, and/or reordering transactions
  - Arbitrage contributes to MEV
- Excessive MEV incentivizes a *Time Bandit Attack* [Daian 19, Qin 21]



Motivating Concerns: Price Oracle Manipulation

- AMMs are also used as *Price Oracles*
  - Other smart contracts query the AMM for spot price quotes
  - Inaccurate quotes risk financial loss: bad loans, currency conversion, etc
- AMMs typically include *Time-Weighted Average Price* (TWAP) oracle
  - Manipulation must be maintained for a period of time [Mackinga 22]
- Arbitrageurs profit when *de-manipulating* the price
  - In this sense, arbitrage is good and desirable
  - Attacker must pay every time to re-manipulate the spot price
- Profitable arbitrages may exist for several blocks [Wang 22]
  - If so, this weakens the TWAP defense!

Ecosystem Study: Two Parts

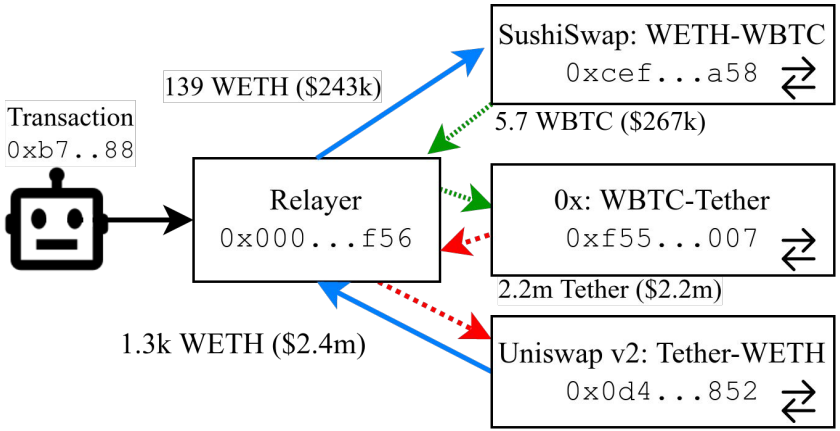
1. Real-world arbitrage activity
    - What patterns of activity do we see bots perform?
    - What are the trends?

2. Opportunity detection
    - What arbitrages could have been taken?
    - How long do they persist?
    - How much profit can be made?

Study period:
Feb. 28, 2020 - Jul 10, 2022

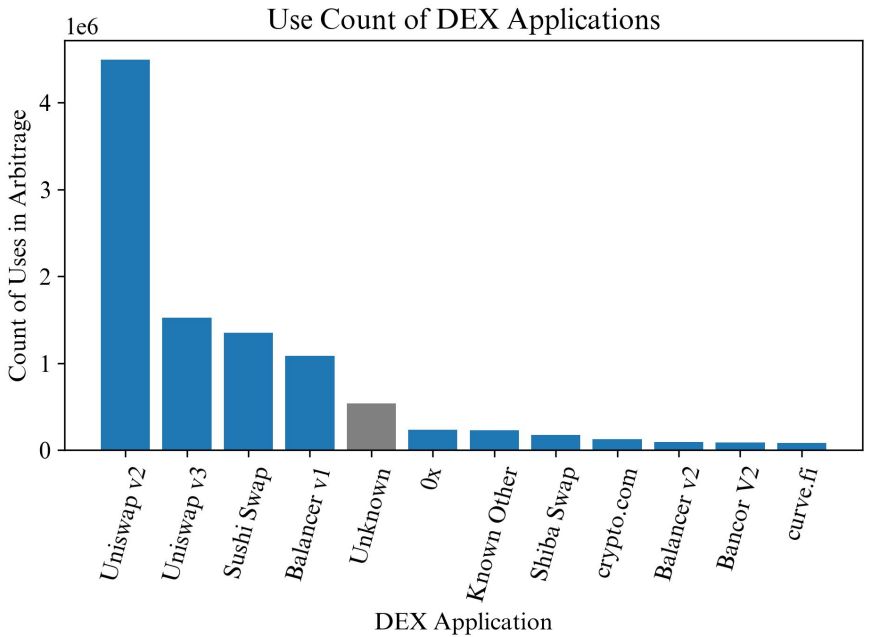
Real-World Activity Detection: Graph Analysis

1. Exchange Inference
   - Find smart contracts that receive one token and emit another
2. Graph Construction
   - Draw a directed graph with ERC-20 Tokens as vertices, exchanges as edges
3. Cycle Detection
4. Cycle Analysis
   - Profit-taking token
   - Amount gained
   - Profiting account address



Selected Results: Overview

- We identify **3.8 million** arbitrages over 28 months
- Total profit after fees: **$321m**



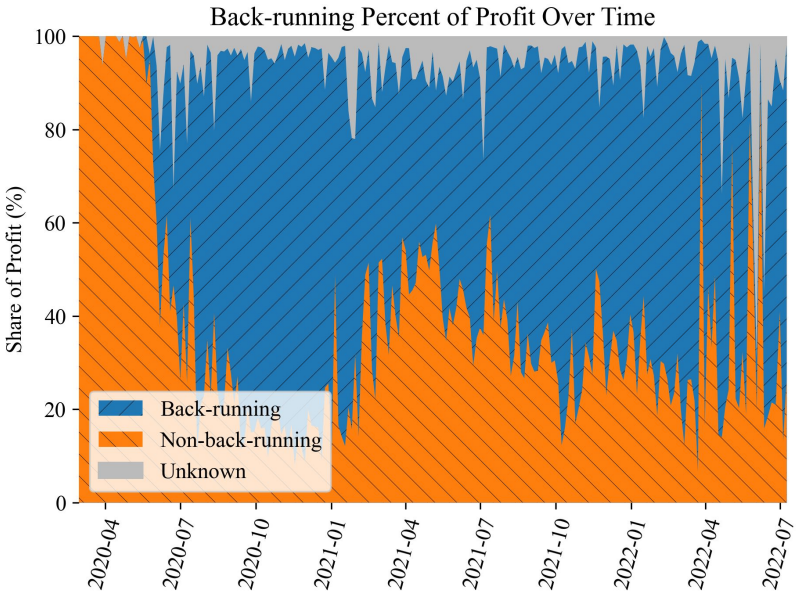
Arbitrage Characteristics

- **98%** perform just one arbitrage cycle
- **Small cycles**:
  - 2 exchanges - 47%
  - 3 exchanges - 44%
- **WETH dominates** profit-taking:
  - 92.4% WETH
  - 2% USDC, 1% Tether, 1% DAI, ...
- Profit is **small**
  - Median: 0.007 WETH (~$10)

Cycle Count	# of Arbitrages	Percent
2	1,817,769	47.29%
3	1,677,920	43.65%
4	286,743	7.46%
5	50,904	1.32%
6	9,802	0.26%
7	480	0.01%
8	77	0.00%
9	13	0.00%
10	6	0.00%
11	1	0.00%
12	1	0.00%

Back-running Dominates

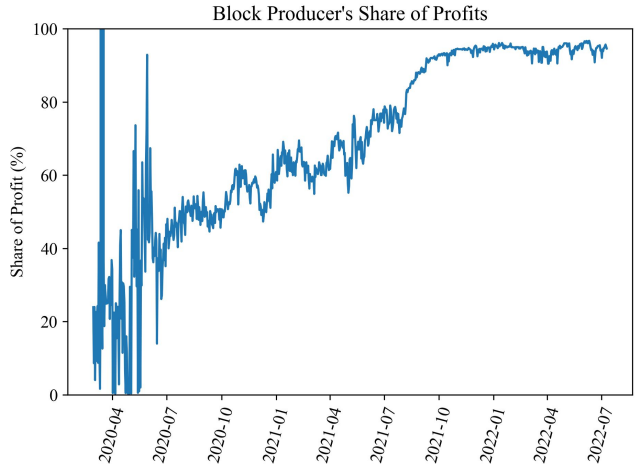
- **Back-running**: strategically placing an arbitrage immediately after the transaction that creates the profit opportunity
- 36% of arbitrages are back-running
- Back-running strategy yields median **5x more** profit

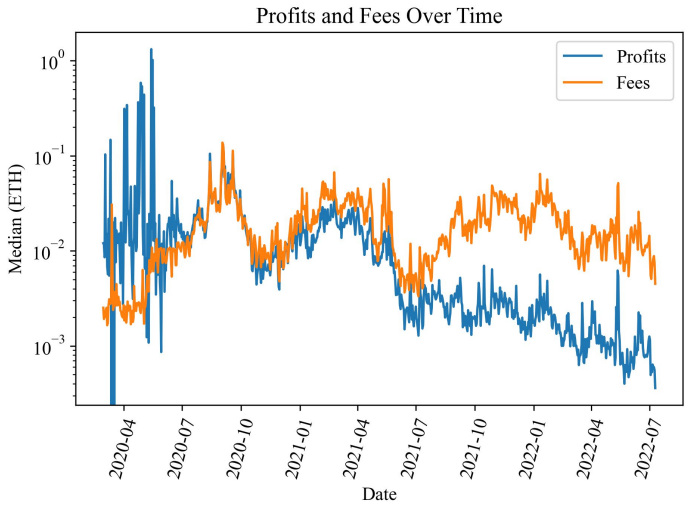


1. …
2. ***price updates***
3. **arbitrage**
…

Block Producer’s Share of Profits

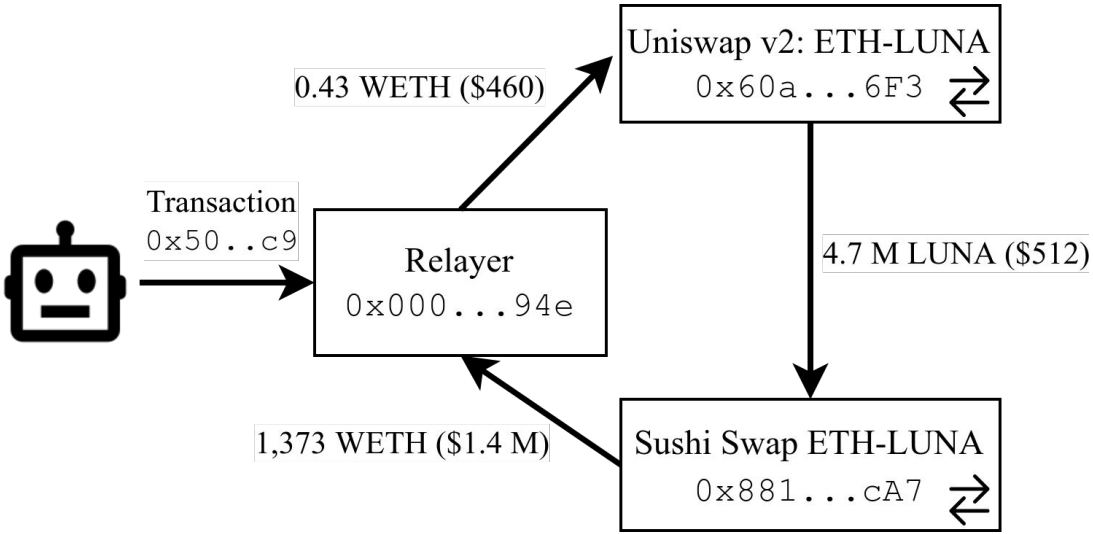
- Block producers are receiving a larger and larger share of the profit
- Profit per arbitrage is decreasing



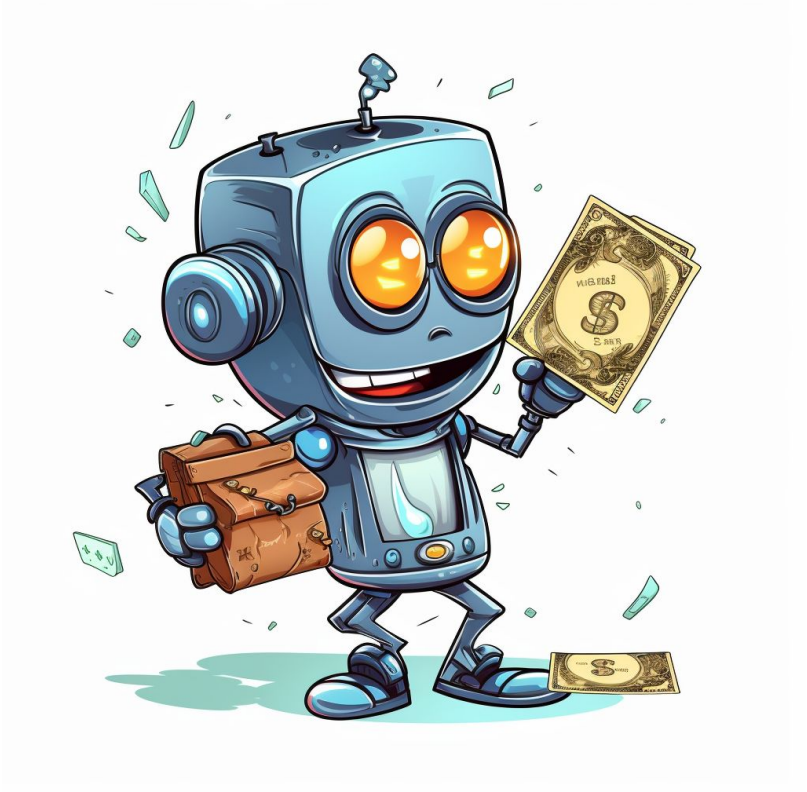


Arbitrage Look-Alikes

- Certain sandwich attacks *look like* arbitrages
- Reduces apparent profit by approx. $5bn



Arbitrage Opportunity Detection



Detection Strategy

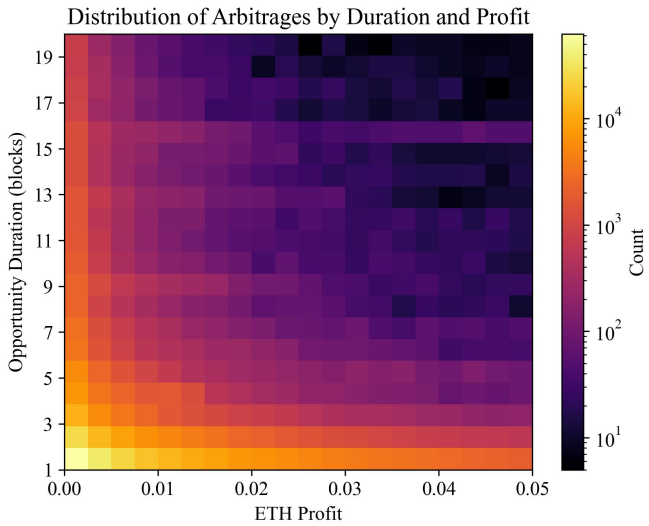
- Scope limitations
  - Only cycles of 2 or 3 exchanges
  - Only take profit in WETH
  - Supported AMM apps
    - Uniswap V2 / V3, Sushi Swap, Shiba Swap, Balancer v1, Balancer v2
- Verify possibility - execute selected transactions via a private fork
- Fee estimate based on prior historical activity

Execution Results: “Large” Arbitrages

- 20.6m potential arbitrages profiting over 1 WETH
- **99.5% failure rate**
  - 55% - Token reverts on transfer
  - Prior work was likely over-estimating arbitrage activity! [excepting Qin ‘19]
- Total profit possibility: $5.7m
- Duration
  - 1 block @ 50th percentile; 4 blocks @ 75th percentile
  - 6 blocks mean

Execution Results: Random Selection

- Executed 126m potential arbitrage opportunities
- **5.7% failure rate**; 119m arbitrages succeed
- After fees, only 1.3 million (~1%) are profitable
- Weekly profit estimate: 395 ETH ($42,600)



Conclusions

- Most arbitrage activity is among a handful of popular exchanges
  - Popular strategies are simple
- The block producers' share is marching upward
- Arbitrage opportunities are quickly taken
  - *But not too quickly!*
- Verifying an arbitrage by execution is essential

Project code:

github.com/ucsb-seclab/goldphish

Robert McLaughlin

robert349@ucsb.edu

SEC LAB