Università degli Studi di Camerino

School of Advanced Studies

Doctoral course in

Blockchain and Distributed Ledger Technology

Curriculum

Social Systems and Smart Societies

Cycle **XXXVIII**

The Impact of Maximal Extractable Value Strategies and Attacks in Proof of Stake Blockchains

PhD Student

Davide Mancino

Supervisor

Prof. Giovanni Denaro

Co-supervisor

Prof. Alberto Leporati

Prof. Marco Viviani

Coordinator of the PhD Programme

Prof. Flavio Corradini

Finanziato dall'Unione europea NextGenerationEU

Ministero dell'Università e della Ricerca

Italiadomani

Piano Nazionale di Ripresa e Resilienza

Date of award 2024/2025

Abstract

Maximal Extractable Value (MEV) is often introduced as a purely technical artifact of transaction ordering, yet its real-world significance is determined elsewhere, in the execution pipeline that decides whether an opportunity is merely observable or can be reliably monetized. This dissertation studies MEV as an empirical, measurable phenomenon, and as an organizational outcome shaped by protocol rules, off-chain coordination, and an increasingly professional infrastructure that mediates how transactions reach blocks.

The thesis addresses a simple but consequential question: what changed when Ethereum block production shifted from proof of work to proof of stake, and when proposer builder separation (PBS) and private order flow became commonplace. The central claim is that these shifts not only relocate power among roles, they also increase conversion efficiency, meaning that a larger fraction of theoretical opportunities becomes realized revenue, and that the identity of the beneficiaries becomes more legible in the data.

To test this claim, the dissertation develops measurement pipelines across eras and links monetary flows to the economic relations among searchers,
builders, validators, and users. On Ethereum, it quantifies how relay-mediated block building reshapes reward composition and the observed concentration of extraction, then it moves from prevalence to behavior, asking whether ordering under MEV infrastructure departs from a fee-based "neutral" baseline. By contrasting what priority should look like under simple price-time logic with what blocks actually exhibit, we surface systematic patterns consistent with preferential inclusion for specific submission channels and explicit payment paths.

Beyond Ethereum, the dissertation connects these findings to the broader proof of stake landscape, using Solana as a motivating reference point for high-throughput execution and specialized MEV routing infrastructure, and framing cross-domain extraction as a sequence-level phenomenon. The cross-chain component introduces an n-hop detector for sequence-dependent arbitrage across swaps and bridge transactions, and empirically bounds multihop extraction over proof of stake systems and rollups, where bridging latency, confirmation rules, and fee schedules act as first-order feasibility constraints.

A network perspective reframes MEV from isolated transactions to persistent relations. The dissertation constructs MEV interaction graphs and uses community detection algorithms to identify stable, role-structured clusters that emerge as statistical evidence of organized extraction without assuming the existence of "cartels" a priori.

Overall, the dissertation offers a cross-era and cross-chain view of realized MEV, clarifying how modern proof of stake execution architectures can simultaneously improve efficiency and sharpen trade-offs among decentralization, fairness, and governance.

Key words: Blockchain; Ethereum; MEV; Maximal Extractable Value; proposer builder separation; private order flow; cross-chain arbitrage; Solana.

Scientific field of the dissertation (SSD): *INF/01*

School the Ph.D. Student belongs to at UNICAM:
School of Advanced Studies, University of Camerino

University where the Ph.D. student carried out his program:
University of Milano-Bicocca, Department of Informatics, Systems and Communication (DISCo)

Other (e.g., Co-funding Programs or Institutions):
Italian National PhD Program in Blockchain and Distributed Ledger Technology, hosted at the University of Camerino, carried out at the University of Milano-Bicocca

Contents

List of Figures

List of Tables

1

Introduction

1.1 Context

Blockchains did not emerge in an intellectual vacuum. They grew out of a longer trajectory in which cryptography and networking were imagined as tools for reshaping power in digital environments. From the 1980s onward, David Chaum, Timothy C. May, Eric Hughes and others described privacy preserving digital cash, anonymous communication, and market-based coordination without centralized intermediaries [21, 68, 84]. Proposals such as b-money [28] and Bit Gold [115] sketched architectures with replicated ledgers, proof of work, and distributed validation that prefigured later cryptocurrencies. In this cypherpunk milieu the slogan "Cypherpunks write code" expressed a programmatic stance: cryptography should be deployed to create new institutional realities, not just to protect existing ones [68].

In October 2008, an author using the pseudonym "Satoshi Nakamoto" published the white paper "Bitcoin: A Peer-to-Peer Electronic Cash System" [88]. Bitcoin instantiated many of these earlier ideas in a concrete protocol: a public ledger of transactions, ordered by proof of work, where miners expend computational effort to propose blocks and the longest valid chain becomes canonical. Digital signatures authorize spending, and consensus about ordering and validity emerges from game theoretic incentives [89]. On 3 January 2009, the genesis block was mined, and the system began operating as a live, adversarial network [20, 45].

Bitcoin's success prompted broader experiments. Ethereum generalized Bitcoin's design by adding a virtual machine and smart contracts, turning the blockchain into a general purpose platform for stateful computation rather than only a payments ledger [17, 89, 124]. A global account model, a Turing-complete execution environment, and a gas-based fee mechanism enabled decentralized exchanges, lending, collateralized stablecoins, governance tokens, and many other applications [62, 124]. In parallel, an ecosystem of alternative platforms emerged, including high throughput chains such as Solana that couple proof of stake with a cryptographic clock known as proof of history to pre order transactions before consensus [125].

Fifteen years after the launch of Bitcoin, this family of ideas underpins a large and heterogeneous ecosystem. Industry surveys estimate that in 2024, roughly 6.8 percent of the world population, more than 560 million individuals, own or use cryptoassets [120]. Earlier benchmarking by the Cambridge Centre for Alternative Finance already found an increase from about 35 million identity verified users in 2018 to roughly 101 million by the third quarter of 2020 [9]. Aggregate market capitalization peaked around 2.8 trillion USD in late 2021, fell sharply during the subsequent correction, and returned to the 2 to 3 trillion USD range by 2024 and 2025 [7, 26]. Central banks and regulators now track this sector as part of their monitoring of financial stability, payment systems, and monetary policy transmission, and many institutions are experimenting with or studying central bank digital currencies in parallel [6, 43].

Within this broader arc, permissionless blockchains evolved from peer-to-peer electronic cash into general-purpose platforms for stateful, verifiable computation. Bitcoin introduced a permissionless ledger secured by proof-of-work, where miners extend the chain by expending computational effort and transaction inclusion is governed by a decentralized fee market, with network latency influencing which transactions miners observe first [88, 89]. Ethereum generalized this model with a virtual machine and smart contracts, enabling decentralized finance and large-scale on-chain coordination [17, 89, 124].

Crucially, once applications manage shared, price-sensitive state, value does not depend only on which transactions are included, but also on *how they are sequenced within the block*. Because transactions execute sequentially and each one updates global state, running a transaction earlier or later can change prices, balances, and execution outcomes, creating profit opportunities or imposing losses on others. This ordering-dependent value is known as *Maximal Extractable Value* (MEV), i.e., the value that an actor with influence over ordering can capture by inserting, removing, or reordering transactions within a block [29, 96]. As MEV opportunities became systematic and competitive, they ceased to be merely an incidental artifact of sequential execution and instead formed an explicit market, with specialized actors coordinating and paying for priority and inclusion through dedicated on-chain and off-chain mechanisms.

1.2 Problem Statement and Motivation

The phenomenon now called Maximal Extractable Value has deep roots. Well before Ethereum's decentralized finance ecosystem took shape, practitioners were already concerned that miners might observe pending transactions and trade against them. A widely cited 2014 discussion on the r/ethereum subreddit raised the possibility that miners could front-run users by inserting their own transactions just before profitable trades [102]. At that stage, however, these concerns were diffuse and largely anecdotal.

The term *Miner Extractable Value* was popularized by Daian et al.'s "Flash Boys 2.0" study, which analyzed decentralized exchanges on Ethereum and showed how block producers could systematically profit by reordering, inserting, or censoring transactions [29]. In their original formulation, MEV denoted

the extra profit that miners in a proof-of-work setting could capture over and above standard block rewards and fees. Subsequent work broadened the concept to *Maximal* Extractable Value to emphasize that any entity with power over ordering, not only miners and validators[1], but also specialized builders, or centralized sequencers, can extract such value [3, 96].

A first era of MEV on Ethereum emerged already in its proof-of-work period and was largely shaped by *public gas auctions* (PGAs) in the public mempool. The mempool is the peer-to-peer pool of pending transactions that have been broadcast but not yet included in a block. In this setting, transaction ordering was effectively allocated through a fee-bidding race: users attached a gas price (the per-unit fee they were willing to pay for execution), and miners had an incentive to prioritize transactions offering higher gas prices, thereby granting earlier intra-block execution. Opportunistic traders, often implemented as automated *searchers* (bots that scan pending transactions for profitable patterns), monitored this public queue for mispriced trades, liquidations, and other on-chain events. When a profitable opportunity appeared, searchers competed by resubmitting transactions with increasing gas prices to outbid rivals and obtain priority in the miner's ordering. "Flash Boys 2.0" documented frontrunning and sandwich attacks on decentralized exchanges in this environment and showed how large MEV opportunities could destabilize consensus by making short reorgs attractive [29]. Later measurement work quantified the scale of this activity and estimated hundreds of millions of USD in extractable value over multi-year windows [95].

As MEV became more visible, the ecosystem began to organize it. One important development was the emergence of private "bundle" relays such as Flashbots. Rather than broadcasting transactions to the public mempool, users or searchers could submit bundles of transactions directly to miners via an encrypted channel, often with a specified ordering and with a side payment to the miner if the bundle was included. This shifted part of the competition from open gas auctions in the mempool to off-chain negotiations over private order flow [47, 49]. In this second era, miners still constructed blocks, but a growing share of the most profitable opportunities bypassed the public mempool and were sold through opaque, private channels.

Ethereum's transition from proof of work to proof of stake in September 2022, commonly referred to as the Merge, reshaped this landscape [34]. Under proof of stake, validators take the role previously played by miners: each slot has a scheduled validator, known as the proposer, who is responsible for proposing a block, while other validators attest to the head of the chain. Leadership is thus predictable at the slot level, and energy-intensive mining is retired [18,35]. Around the same time, the ecosystem converged on an explicit market design known as proposer builder separation (PBS). Under PBS, the validator who is scheduled to propose a block sells the right to construct that block to specialized *builders* via an auction. Builders collect bundles from searchers and transactions from the mempool, assemble candidate blocks that maximize total

[1]Here we use *miners* (PoW) and *validators* (PoS) interchangeably to denote the consensus-selected block producer.

value, and submit sealed bids to the proposer via *relays*, i.e., software infrastructure components that receive builders' execution payloads and forward only the highest-paying one to the proposer, typically keeping its contents private until selection [52, 122]. MEV Boost is the main out of protocol implementation of PBS on Ethereum: it is a client side middleware that validators run to receive bids from a network of builders through trusted relays and to pick the highest paying block [48, 53].

In this third era, MEV is no longer an incidental byproduct of fee markets; it is sold in near real time through specialized block construction markets. Searchers identify opportunities and assemble bundles; builders batch and optimize these bundles into full blocks; relays intermediate between builders and proposers; validators, acting as proposers, are paid for selling their slot. Empirically, and consistent with the measurements reported in this thesis, a large fraction of Ethereum blocks in the post-Merge period are built via MEV-Boost. This implies that most MEV is now realized through this explicit auction pipeline rather than through ad hoc gas wars in the public mempool [3, 44, 122].

A parallel evolution occurs on other blockchains. High-throughput proof-of-stake systems such as Solana combine a fast base layer with mechanisms like proof of history to pre-order transactions [125]. On Solana, searchers and block producers interact through specialized infrastructure such as the Jito block engine and tip accounts, which similarly channel MEV opportunities into structured markets rather than leaving them entirely in the public gossip layer. Recent systematization work on cross-domain and cross-chain MEV emphasizes that ordering power and extractable value are not specific to Ethereum; they arise whenever multiple venues, bridges, and on-chain markets interact under shared settlement constraints [86].

The central problem examined in this thesis is that, while MEV has become industrialized and explicitly traded, its realized scale and distribution under modern proof of stake and PBS-based architectures remain imperfectly understood. Existing empirical work leaves several gaps.

First, most quantitative studies focus either on pre-Merge, proof of work Ethereum or on relatively narrow post-Merge windows [29, 96, 98]. This complicates cross era comparisons and makes it difficult to assess how much the move to proof of stake with PBS has increased the fraction of theoretical MEV that is actually realized as side payments, tips, or builder profits. Second, measurements are largely protocol-centric: they quantify opportunities and profits at the level of transactions, trading pairs, and DeFi venues, but offer limited visibility into the longer-lived actors behind these flows. In particular, there is relatively little work that systematically distinguishes and links the roles of searchers, builders, relays, validators, and users across time, or that maps how these agents are connected in persistent economic relationships.

Third, research and public dashboards tend to concentrate on Ethereum [32, 37, 92]. Yet, the same structural issues, namely order sensitivity, cross-venue arbitrage, liquidations, and sandwich attacks, also arise on other chains. High-throughput architectures such as Solana have different fee markets, consensus parameters, and ordering mechanisms, which may change the form and preva-

lence of MEV but not its existence. A comparative picture across architectures is still fragmentary [86].

Fourth, as MEV markets mature, concerns about negative externalities grow. MEV is often realized by imposing costs on other users: toxic arbitrage exploits stale prices; sandwich attacks worsen slippage and execution quality for ordinary users; liquidation sniping extracts surplus from borrowers on lending protocols such as Aave and Compound, where the first transaction that liquidates an undercollateralized position captures the liquidation bonus, creating a race condition with systemic risk implications if competing liquidators delay action [4, 96]; and gas bidding wars or overfitting of private order flow can reduce effective throughput and raise costs [3, 29, 96]. In extreme cases, large MEV opportunities can create incentives for short reorgs or chain instability. Concentration of extraction in a small number of highly connected actors raises questions about centralization, bargaining power, and fairness in permissionless financial infrastructures [3, 44]. Recent surveys and policy reports treat MEV as a potential source of market manipulation, unequal access, and opacity in on-chain markets [44, 60, 83].

It is worth emphasizing that ordering sensitivity is not a marginal feature of DeFi but a structural property of any protocol whose state transitions depend on the sequence of transactions within a block. Lending protocols (Aave, Compound) rely on timely oracle price updates and liquidation execution whose outcome changes if an adversarial transaction is inserted before or after the oracle call [4]. Automated Market Makers (Uniswap, Curve) are sensitive to the sequence of swaps within a block because each swap shifts the price, creating opportunities for front-running and sandwiching [29, 96]. Understanding and mitigating MEV is therefore not optional for the long-term health of these protocols; it is a prerequisite for their reliable operation [60].

The structural parallels between MEV and *insider trading* in traditional finance deserve explicit attention, as they frame several of the normative concerns explored in this thesis. In equity and futures markets, insider trading refers to the practice of executing trades on the basis of material non-public information, typically information about pending orders or corporate events that is unavailable to other market participants [70]. In the MEV setting, block producers and searchers observing the public mempool gain analogous information advantages: they see pending user transactions before execution and can insert, reorder, or suppress them to extract profit. Front-running, submitting a transaction ahead of a known pending trade [29], mirrors the classical securities law prohibition on trading ahead of a client order. Sandwich attacks are structurally equivalent to quote-stuffing around a large order. Crucially, a key legal distinction applies: in traditional finance, exploiting non-public order book information is actionable under securities regulation, whereas in public blockchains the mempool is technically *public*, making the same information advantage permissible by protocol design. The European Securities and Markets Authority (ESMA) has explicitly flagged this asymmetry, noting that MEV extraction may constitute market manipulation under existing EU frameworks even when it is not prohibited by the underlying protocol [44]. Ji and Grim-

melmann propose an "intent-respecting" test for MEV liability, arguing that regulatory exposure arises when actors profit by routing transactions contrary to the user's expressed intent [70]. This framing is directly relevant to the ordering analysis in Chapter 5, where systematic prioritization of private flows over public transactions can be interpreted as a mechanism for redirecting execution outcomes away from what fee-paying users would expect under neutral block production.

These gaps motivate the present thesis. The core aim is to move from anecdotal, protocol-level, and Ethereum-centric accounts of MEV to a systematic, cross-era, and cross-chain measurement of realized MEV, grounded in the economic relations among the actors who extract and redistribute value. The thesis takes as its starting point the hypothesis that the transition from proof of work to proof of stake, together with proposer-builder separation (PBS) and the growing use of "private order flow", increases the efficiency with which MEV is converted from theoretical opportunities into realized revenues, thereby raising realized MEV. Here, "private order flow" denotes transactions routed outside the public mempool, for example, via private RPC endpoints or relay and builder-mediated channels, which restrict pre-inclusion visibility to a limited set of intermediaries. The thesis then asks how the higher conversion efficiency interacts with decentralization, fairness, and governance, and how it propagates across architectures.

In this sense, the work is both descriptive and normative. Descriptively, it aims to quantify how much MEV is realized in different eras, who receives it, and how the underlying network of relationships is structured. Normatively, it uses these measurements to engage with questions of democratic ordering and credible neutrality. In particular, we ask whether the combination of PBS, relays, and private order flow mainly formalizes and channels behaviors that would otherwise occur in more chaotic ways, or whether it undermines fee-based meritocracy for ordinary users.

1.3 Research Questions and Objectives

The thesis addresses four main research questions.

1. **RQ1. Quantitative prevalence of MEV.** How prevalent is MEV across eras and architectures. What is the extent of MEV in Ethereum before and after the Merge, how does it compare to high-throughput designs such as Solana, and what share of blocks and transactions show MEV-linked patterns.

2. **RQ2. Ordering and favoritism.** Does observed intra-block ordering systematically favor MEV-related activity. Do private channel submissions and direct payment to the block builder flows[2] obtain priority beyond

[2]In this thesis, "direct payment to the block builder flows" denotes direct payments from a searcher to a builder in exchange for inclusion and ordering guarantees, implemented either as a dedicated payment transaction or as a cryptocurrency transfer executed within a smart contract.

what fees would imply. Are there signals consistent with collusion among searchers, builders, and validators.

3. **RQ3. Social structure and centralization.** Which stable groups and interest clusters form within the MEV ecosystem. How do PBS-mediated relationships differ from direct transfers, and what is the degree of centralization among extracting actors.

4. **RQ4. Cross-chain dynamics.** How does MEV manifest across chains. What is the feasibility and frequency of multi-hop cross-chain arbitrage, and how do bridge latency and costs shape profitability and execution time.

The overarching objective is to provide a comprehensive empirical account of realized MEV that (i) spans Ethereum's pre-Merge and post-Merge eras, (ii) compares Ethereum to an alternative architecture, (iii) maps the social structure of extracting actors, and (iv) characterizes cross-chain strategies. A central theme is to evaluate how proof of stake with PBS affects realized MEV relative to proof of work, and to quantify the consequences for decentralization and democratic, fee-based ordering. In pursuing these objectives, the thesis also develops methods for role classification, sender-centric graph construction, and multi-hop path detection that may be of independent interest for the study of on chain economic networks.

1.4 Contributions

This work makes empirical, methodological, and conceptual contributions to the study of realized MEV under proof of stake and to the comparative analysis of MEV markets across architectures. On the empirical side, a cross-era measurement that spans more than 6.3 million Ethereum blocks documents the post-Merge dominance of relay-mediated MEV markets. Lower bound estimates for Solana complement the Ethereum analysis and clarify the design trade-offs that arise when high-performance chains adopt different execution and ordering architectures. The empirical results further show systematic deviations from fee-based ordering, with strong and repeated prioritization for transactions that carry explicit builder payments. When these flows are mapped into MEV graphs, the resulting network structure exhibits high modularity and clear differences between edges that are mandated by PBS and edges that reflect direct, discretionary ties. Finally, the study characterizes cross-chain arbitrage and shows that multi-hop strategies are rare in practice and subject to heavy frictions, with execution time increasing in path length.

Methodologically, the work introduces a principled framework for classifying actors into searcher, builder, validator, and related roles, together with precedence rules that resolve ambiguous or overlapping behaviors. It develops a sender-centric transformation pipeline that converts contract-centric traces into sender-centric value flows so that hidden relationships between economic agents become visible. Building on this representation, it specifies an algorithm

for multi-hop detection that imposes explicit temporal, value, token, actor, and feasibility constraints to distinguish plausible cross-chain arbitrage from noise. These components support a graph construction methodology that yields typed, weighted, and directed MEV graphs in which PBS obligated edges are separated from direct edges, enabling structural comparisons and community-based analyses.

Conceptually, the thesis offers an organizing perspective on realized MEV under proof of stake and provides evidence that PBS and private order flow raise realized extraction relative to proof of work [64, 78]. It proposes a democratic triad for evaluating MEV practices, based on the principles of no collusion, meritocracy, and legitimate MEV, and uses this triad to interpret observed behaviors. Finally, it advances a PBS versus direct dichotomy that distinguishes procedural, obligated relationships from voluntary, potentially collusive ties, and shows how this distinction helps to reason about power, accountability, and fairness in contemporary MEV markets.

1.5 Structure of the Dissertation

The dissertation is organized in eight chapters that move from context and theory to measurement, network analysis, and cross-chain evidence. Chapter 2 provides the background. It traces the philosophical and cultural origins of decentralization, introduces Bitcoin, and then presents Ethereum, Ethereum's account model, gas mechanics, and the transition from proof of work to proof of stake.

Chapter 3 turns to measurement and quantifies MEV in Ethereum. We construct a longitudinal dataset that spans the consensus transition, identifies the roles of relays and builders, and analyzes validator rewards and the evolution of MEV strategies over time, with a focus on realized extraction rather than only theoretical opportunities.

Chapter 4 extends the analysis across architectures by focusing on Solana. It estimates the prevalence of MEV on a high-throughput, proof-of-history-based chain, studies Jito-mediated extraction, and highlights architectural contrasts with Ethereum that bear on ordering power, private order flow, and realized MEV.

Chapter 5 examines ordering and democratic principles. Using intra-block data, we test whether gas price-based meritocracy holds in practice, measure the treatment of private transactions and builder payments, and search for signals consistent with collusion or preferential treatment among searchers, builders, and validators.

Chapter 6 adopts a network perspective and studies the social structure of MEV. We build and analyze MEV graphs, apply community detection, compare PBS-mediated edges with direct ties, and evaluate temporal stability and the balance between intra- and inter-community flows in the MEV ecosystem.

Chapter 7 moves to cross-chain MEV. We investigate the feasibility and frequency of multi-hop arbitrage paths across blockchains, quantify how bridge latency and fees affect profitability, and document execution time patterns as

path length grows, thereby characterizing frictions that limit realized cross-chain extraction.

Finally, Chapter 8 outlines our conclusions, summarizing the main empirical findings, drawing out implications for protocol design and governance, reflecting on limitations, and outlining directions for future work.

Table 1.1 provides a concise map from research questions to chapters, datasets, and primary methods, as a navigation aid for the reader.

Table 1.1: Research questions, corresponding chapters, datasets, and primary methods.

RQ	Question (summary)	Chapter	Dataset	Primary method
1	MEV prevalence across eras and architectures	3, 4	6.3M ETH blocks (Feb 2021–Sep 2023); 500 SOL blocks (Apr 2023)	Relay linkage; Jito label matching
2	Ordering favoritism beyond fee-based baseline	5	220k ETH blocks, 36M txs (Jan 2024)	Tx_Shift metric; ZeroMEV labels
3	Social structure and centralization	6	MEV interaction graph (Jan 2024)	Leiden community detection; Sankey analysis
4	Cross-chain multihop SDA feasibility	7	2.49B swaps, 34.8M bridge txs, 12 chains (Sep 2023–Aug 2024)	N-hop SDA detector; power-law fit

2

Background

Blockchain technology did not emerge in an intellectual vacuum. Its ethos of decentralization, distributed power, and trustless interaction can be traced to rich philosophical and cultural lineages [31, 82]. Decades before the advent of Bitcoin in 2008, scholars and visionaries were grappling with how technology could reshape power structures, enable new forms of freedom, or conversely reinforce control [82]. This chapter surveys those deep roots (from post structuralist philosophy and critical theory to science fiction and the cypherpunk movement) and then recounts the historical development from Bitcoin's invention [88] to Ethereum's expansion of blockchain into a general purpose platform [17, 124]. In doing so, it establishes the context for understanding later developments. The evolution of blockchain can be viewed as a confluence of ideas: a philosophical impulse to diffuse power, a cultural imagination of networked liberty, and a technical realization in distributed computing [8].

2.1 Philosophical and Cultural Origins

In network engineering, decentralization acquired an early and concrete technical meaning, which later became influential beyond its original domain. In 1964, Paul Baran introduced distributed communications networks that could operate without a central controller [8]. Baran showed that a distributed mesh topology could achieve high robustness against node failures and targeted attacks, since messages could be rerouted around broken links [8]. Baran's work offered not only methods, but also a guiding image of resilient, many-to-many connectivity. Figure 2.1 contrasts three architectures: (A) a centralized star in which a single hub concentrates coordination and failure risk, (B) a decentralized layout with multiple hubs and limited inter-hub connectivity, and (C) a distributed mesh in which connectivity is spread across peers. In Baran's framework, (C) is the most robust configuration, because it minimizes single points of failure and supports rerouting along many alternative paths.

Within philosophy, early post-structuralist thought articulated parallel models for non-hierarchical organization. Gilles Deleuze and Felix Guattari's notion of the rhizome provides a compact frame. Unlike a tree, a rhizome is an open,

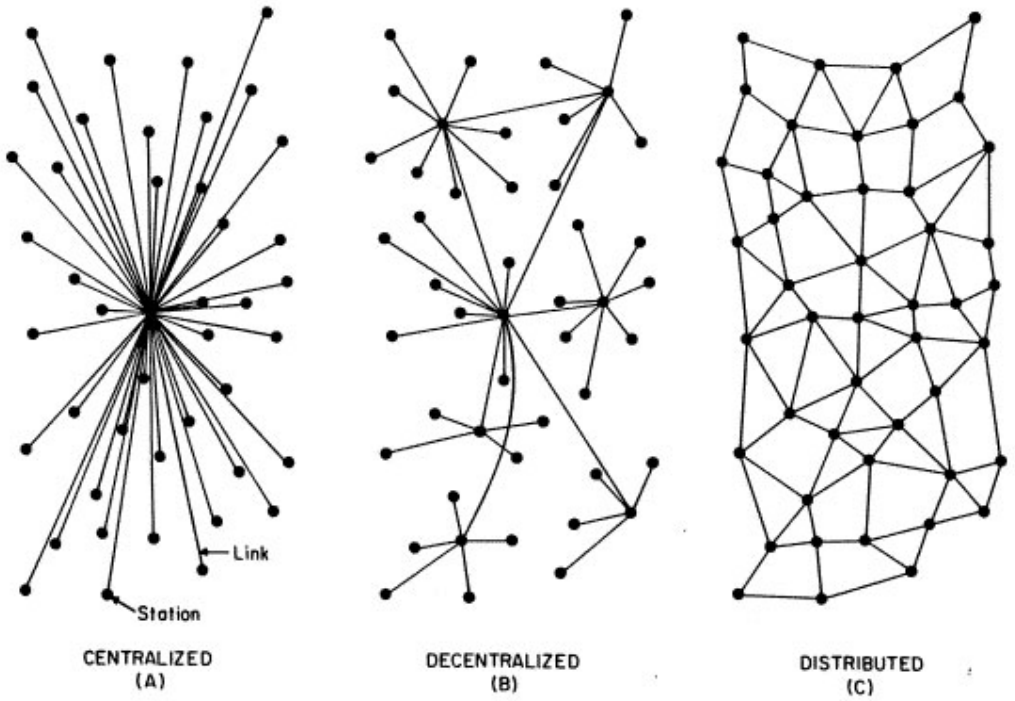


Figure 2.1: Paul Baran's network topologies [8]. (A) Centralized star. (B) Decentralized network with multiple hubs and sparse inter-hub links. (C) Distributed mesh is the most robust topology in Baran's framework, since it minimizes single points of failure and supports extensive rerouting.

subterranean system that spreads in all directions and, in their words, "any point of a rhizome can be connected to anything other, and must be" [31]. In "A Thousand Plateaus" they state that a rhizome "ceaselessly establishes connections between semiotic chains, organizations of power, and circumstances relative to the arts, sciences, and social struggles" and they codify a "principle of connection and heterogeneity" meaning that any element may link to any other, even when they are very different, together with a "principle of asignifying rupture" meaning that breaks do not collapse the system, which can reconfigure around them [31]. Read in this way, the rhizome describes flexible, distributed, resilient assemblages, and these ideas, alongside many distinct developments in computing and network engineering, have been taken as conceptually consonant with later interest in distributed, peer-to-peer architectures.

2.2 The Cypherpunk Movement

In the late 1980s and 1990s, mathematicians, programmers, and cryptography advocates coalesced around discussions of privacy and freedom on the Cypherpunks mailing list, a forum documented in Eric Hughes's writings [68, 74]. The list operated as an electronic space for code, debate, and coordination, with early administrative activity evidenced by Hughes's September 1992 mes-

sage [67]. The term cypherpunk is a portmanteau of cipher and cyberpunk, credited to Jude Milhon in 1992, and the name explicitly draws on cyberpunk literary themes [97]. Preceding the emergence of this community, cyberpunk fiction by William Gibson, Bruce Sterling, and Neal Stephenson in the mid 1980s had depicted networked societies, ubiquitous computation, and countercultural actors operating through code and networks, themes that informed the later coinage and framing of cypherpunk [58, 97, 112, 113].

Eric Hughes articulated the movement's stance in "A Cypherpunk's Manifesto" with the statements "Privacy is necessary for an open society in the electronic age" and "We cannot expect governments, corporations, or other large, faceless organizations to grant us privacy out of their beneficence. We must create it for ourselves", and with the motto "Cypherpunks write code" [68]. Timothy C. May, an early organizer associated with the Cypherpunks list, authored "The Crypto Anarchist Manifesto" in 1988, presenting a programmatic statement on privacy, encryption, and market exchange in networked environments [84]. The manifesto argues that strong cryptography enables private communication and voluntary transactions outside centralized oversight in digital contexts [84].

Researchers associated with this milieu developed concrete tools and proposals, including anonymous communications and public key encryption for private messaging [68, 84]. David Chaum introduced blind signatures to support untraceable payments, later implemented in DigiCash [21]. Wei Dai described b-money, a scheme for online money without a central issuer [28]. Nick Szabo proposed bit gold, outlining chained proofs of work and replicated records, and earlier formalized smart contracts as enforceable relationships in code [114, 115].

By the 2000s these writings and prototypes documented a trajectory emphasizing privacy, pseudonymity, and distributed validation [28, 68, 84, 115]. A public implementation of a peer to peer electronic cash system appeared in 2008 with the publication of "Bitcoin: A Peer to Peer Electronic Cash System" [88].

2.3 The Birth of Bitcoin: Motivation, Structure, and Impact

2.3.1 Motivation and Historical Context

In October 2008, an unknown individual (or group) under the pseudonym "Satoshi Nakamoto" published the white paper "Bitcoin: A Peer to Peer Electronic Cash System" [88]. Released to the Metzdowd cryptography mailing list [75], the paper proposed a way to prevent double spending without a trusted intermediary, namely a public, append-only ledger secured by proof of work (PoW). As Nakamoto put it, "A purely peer-to-peer version of electronic cash would allow online payments to be sent directly from one party to another without going through a financial institution" [88]. The result was a system where digital signatures authorize spending, while consensus about

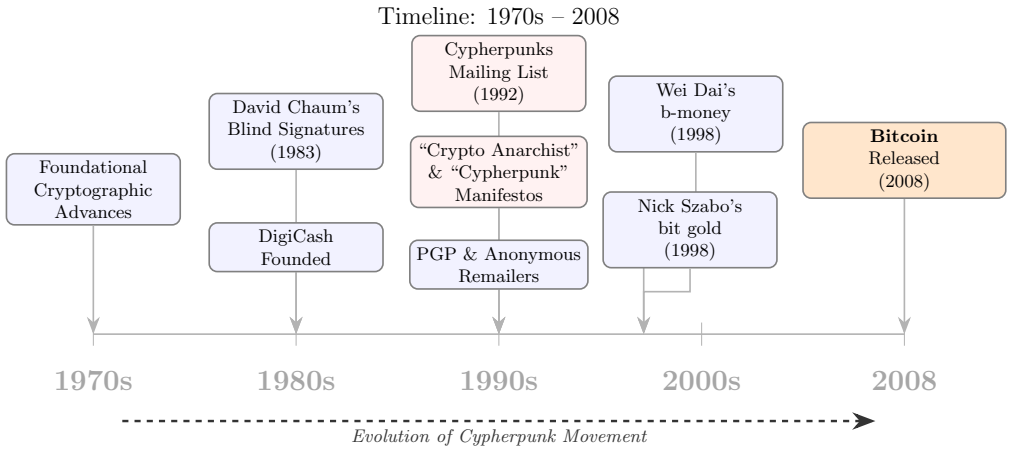


Figure 2.2: Milestones often cited in accounts of the cypherpunk milieu, including Chaum's blind signatures, the Cypherpunks mailing list and manifesto, and late 1990s proposals for digital cash [21, 28, 68, 84, 115].

ordering and validity emerges from game theoretic incentives and the longest chain rule [88, 89]. At a high level, each block carries a header that links to its predecessor and a Merkle root that commits to the block's transactions; this structure is depicted in Figure 2.3, while the Merkle aggregation itself is sketched in Figure 2.4.

On 3 January 2009, Nakamoto mined the genesis block, embedding the headline "The Times 03/Jan/2009 Chancellor on brink of second bailout for banks" in the coinbase data [20, 45]. This served both as a timestamp and as a socio-economic signal in the wake of the 2008 crisis [20]. In Bitcoin's first months, Hal Finney, a renowned cryptographer and early cypherpunk, received the first-ever Bitcoin transaction, namely 10 BTC from Nakamoto in January 2009 as a test [20, 46]. By 2010, exchanges like Mt. Gox gave the asset a dollar price, and by 2011, the broader economy around Bitcoin had grown into the millions of USD equivalent [94]. The Silk Road marketplace, launched in 2011, demonstrated both Bitcoin's censorship resistance and the policy controversies it would raise [23]. Beyond payments, researchers and builders recognized that Bitcoin's underlying architecture, that is, a distributed ledger plus economic consensus, could support use cases beyond currency [62, 89].

2.3.2 UTXO Transaction Model: What Is Owned and Spent

Bitcoin represents value as discrete unspent transaction outputs (UTXOs): each transaction spends one or more previously created outputs and produces fresh outputs locked by spending conditions [62, 88, 89]. Because transactions spend entire outputs and create fresh ones, coins are not global balances but discrete UTXOs that move along edges of the transaction graph.

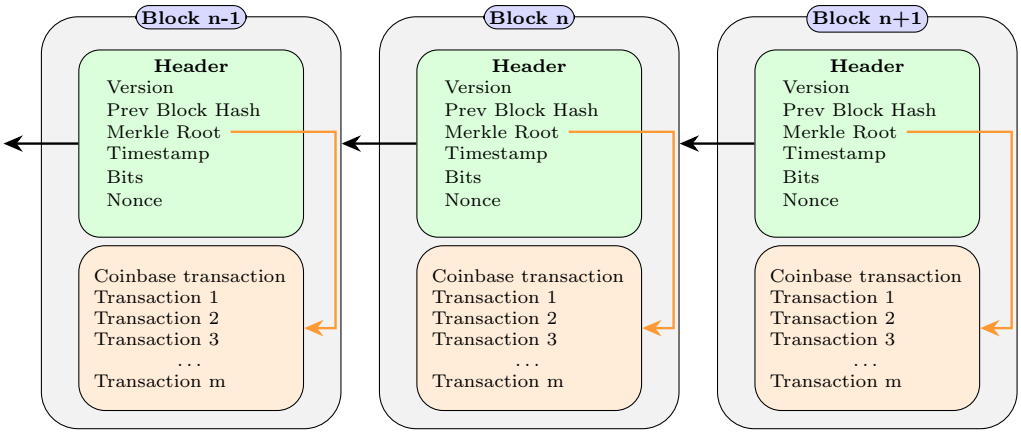


Figure 2.3: Blockchain structure. The *Prev Block Hash* line of each block equals the previous block's hash.

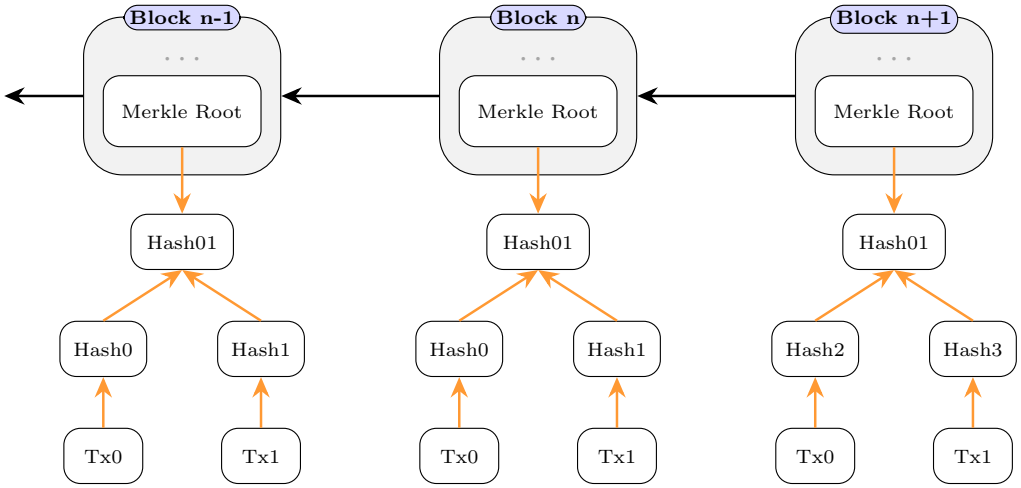


Figure 2.4: A Merkle root is computed by hashing each transaction into a leaf, then repeatedly hashing adjacent pairs to build a binary tree until one hash remains. This single hash, the Merkle root, commits to all transactions and enables compact inclusion proofs (Merkle proofs).

Concrete example (Alice pays Carol and Bob). Suppose a prior transaction $\mathbf{Tx}_0$ credited Alice with two UTXOs: UTXO_{0a} = 0.60 BTC and UTXO_{0b} = 0.20 BTC (so Alice controls 0.80 BTC in total). Alice now wants to pay two recipients at once: 0.50 BTC to Carol and 0.20 BTC to Bob. Because UTXOs are indivisible at spend time, Alice's wallet constructs $\mathbf{Tx}_1$ that spends both inputs UTXO_{0a} and UTXO_{0b}, and creates three new outputs:

to Carol: 0.50 BTC,
to Bob: 0.20 BTC,
change back to Alice: 0.09 BTC.

The remaining 0.01 BTC is the transaction fee paid to the miner. The accounting balances exactly:

$$0.60 + 0.20 \;=\; 0.50 + 0.20 + 0.09 + 0.01.$$

This illustrates three core properties: (i) entire UTXOs are consumed as inputs, (ii) a single transaction can pay multiple recipients via multiple outputs, and (iii) the wallet usually creates a change output returning the excess to an address controlled by Alice. Conceptually, this is the same flow depicted by a UTXO graph (for example, Figure 2.5), where edges denote previously created outputs being consumed by later transactions.

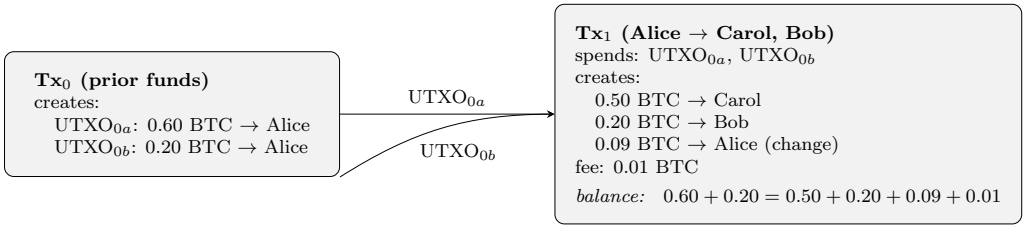


Figure 2.5: UTXO transaction graph for a multi-output payment. Alice spends two inputs (0.60 and 0.20 BTC) in a single transaction to pay 0.50 BTC to Carol and 0.20 BTC to Bob, returning 0.09 BTC as change; the 0.01 BTC remainder is the fee. Entire UTXOs are consumed as inputs, and a single transaction can produce multiple outputs.

Second example (coin selection with several small UTXOs). Assume Alice instead controls four UTXOs worth 0.30, 0.30, 0.15, 0.05 BTC. To pay 0.50 BTC to Carol and 0.20 BTC to Bob with a 0.01 BTC fee, her wallet selects three inputs 0.30 + 0.30 + 0.15 = 0.75 BTC, then creates outputs 0.50 BTC to Carol, 0.20 BTC to Bob, and 0.04 BTC as change back to Alice. The untouched 0.05 BTC remains as a separate UTXO under Alice's control. This shows how multiple inputs can be aggregated to meet a target amount while any remainder returns as change, leaving unselected UTXOs unaffected (Figure 2.6).

2.3.3 PoW Roles, Blocks, Fork Choice, and Reorgs

A miner is a node operator who runs a full validating client, verifies new transactions, and maintains a local memory pool (mempool), which stores valid but unconfirmed transactions waiting for inclusion in a block [89]. To mine a block, the miner constructs a block candidate by selecting transactions from the mempool, usually prioritizing higher fee rates, adds a special coinbase transaction that creates new coins as the block subsidy and collects the fees of the included transactions, and assembles a header that commits to the candidate content [88,89]. The miner then performs proof of work: it repeatedly recomputes the block header hash while varying the nonce and other mutable fields (such as the timestamp) until the double SHA256 of the header is strictly less than

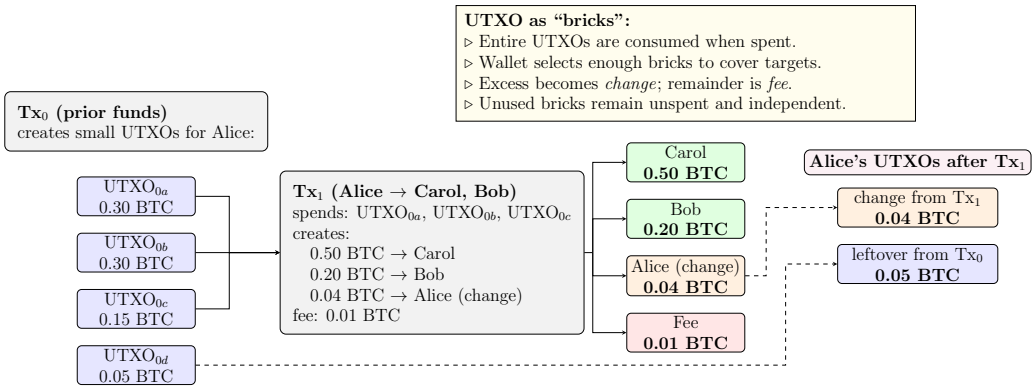


Figure 2.6: Alice controls several small inputs (0.30, 0.30, 0.15, 0.05 BTC). To pay 0.50 BTC to Carol and 0.20 BTC to Bob in a single transaction, her wallet selects three entire bricks (0.30+0.30+0.15). The transaction creates outputs (Carol 0.50, Bob 0.20, change 0.04) and pays 0.01 as a fee. Rightmost: Alice's UTXOs after Tx_1 are the new 0.04 BTC change and the untouched 0.05 BTC brick.

the current difficulty target. Bitcoin specifies hashing the header twice with SHA 256 and the resulting 256 bit number must be below the target threshold implied by the network's current difficulty; a lower target makes success rarer [88, 89]. Each trial is independent and miners with higher hash rate (the number of hashes attempted per second) attempt more trials per second [89]. When a valid header is found, the miner immediately broadcasts the block to its peers; the network verifies the proof and the block contents, and the finder becomes the proposer for that height [88].

Proof of work secures the chain because finding a valid header requires expected work proportional to the difficulty, while verifying the proof requires only a single hash check. The hash rate of a miner is the expected number of header hashes that the miner can compute per second; the network hash rate is the aggregate across all miners, and it determines the expected inter-block time for a given difficulty [89]. Bitcoin adjusts difficulty periodically so that the expected time between blocks remains close to the protocol target despite changes in total hash rate [89].

Consensus follows the longest valid chain rule, often described as the heaviest chain by cumulative proof of work. If two blocks are found close in time, the network may observe a temporary fork; honest nodes extend the branch that they received first, and the branch that accumulates more work becomes canonical [88, 89]. A reorganization, or reorg, occurs when a competing branch gathers more cumulative work and replaces one or more tips; the probability that a buried block is replaced decreases quickly with depth and it depends on network latency and the variance of block discovery [57, 89]. Propagation delays also shape which transactions are visible to miners when a block is assembled, which in turn affects ordering and fee revenue [30, 57].

A typical transaction lifecycle is as follows. One, a wallet signs and broadcasts. Two, peers validate the basic policy and gossip about the transaction.

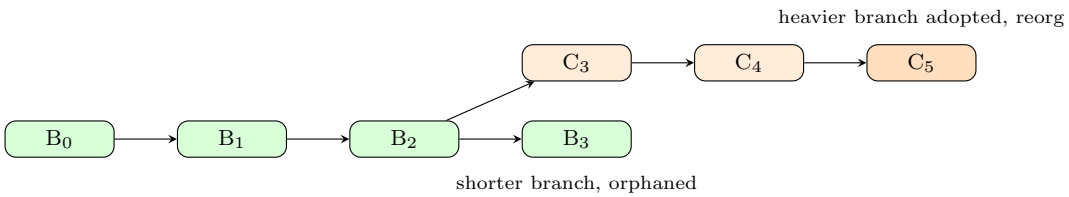


Figure 2.7: Fork choice and reorg. Two competing tips appear, and the branch with more accumulated work becomes canonical.

Three, the transaction sits in local mempools and may be updated via replace-by-fee. Four, a miner selects from its mempool view, chooses ordering, and mines a block. Five, the block propagates. Six, the network adopts the longest valid chain. Seven, with depth, reorg risk declines [30, 88, 89].

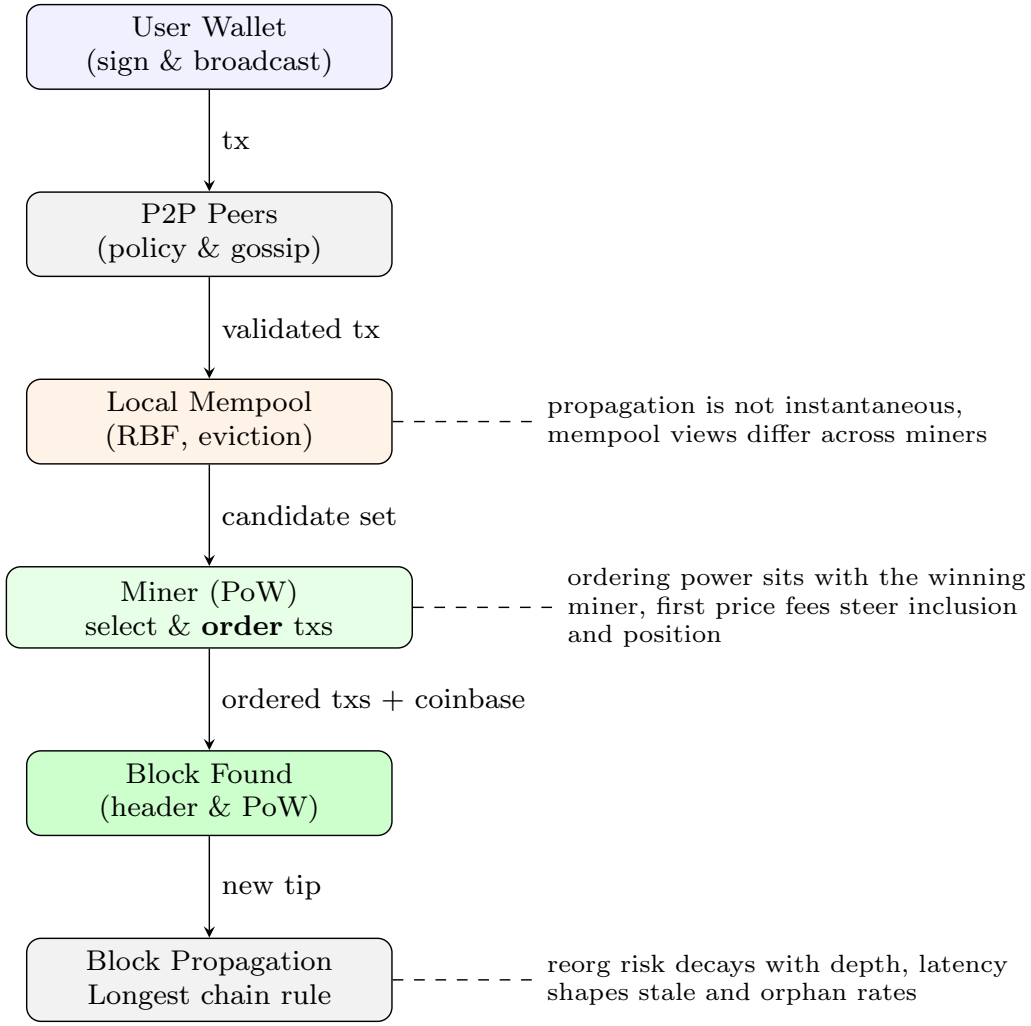


Figure 2.8: Bitcoin transaction lifecycle.

2.3.4 Historical Impact

Early adoption and socio-economic impact include the Finney transaction in January 2009 [20, 46], the emergence of exchanges and price discovery in 2010, and the Silk Road era in 2011 [23]. These events contextualize the shift from a niche experiment to a live, adversarial network, conditions under which inclusion, ordering, and finality became economically meaningful [89,94]. This breakthrough opened the imagination to other use cases that traditionally required central authorities. Could the blockchain paradigm be extended to contracts, property records, web infrastructure, and more. Bitcoin was deliberately limited in scope (serving primarily as a ledger of currency transactions, with only a simple scripting language for conditional payments). By around 2013, the community was actively asking whether a more expressive blockchain could be built, one that would support arbitrary programs and not just monetary transfers. That question led directly to the creation of Ethereum [17, 124].

2.4 Ethereum as a General Purpose Platform

By late 2013, Vitalik Buterin, already known as co-founder of "Bitcoin Magazine", proposed Ethereum, namely a platform that generalizes the blockchain concept beyond currency [17]. Buterin argued that Bitcoin's scripting was too limited for many ideas that the community envisioned, for example decentralized exchanges, derivatives, identity, and complex multiparty agreements [17, 89]. His vision was a world computer, that is a decentralized platform capable of running arbitrary programs, later called smart contracts, so that many forms of coordination could be automated in a trust minimized way [17].

In the whitepaper, Buterin explicitly framed Bitcoin as an early proof of decentralized trust, but one offering only a weak version of the smart contract concept in Nick Szabo's sense [17, 114]. Ethereum integrates a Turing-complete virtual machine, namely the Ethereum Virtual Machine, and a transaction fee mechanism called gas, thereby enabling on-chain programs with bounded computation [124]. The architectural leap, from a single-purpose ledger to a general state transition system, launched what many describe as Blockchain 2.0 [89].

2.4.1 Account Model, EVM, Gas, and State

Ethereum adopts an account-based execution model, in contrast with Bitcoin's UTXO set. The global world state maps address externally owned accounts and contract accounts. Each account holds a nonce, a balance, an optional code, and persistent storage as a key-value mapping. Transactions are signed by externally owned accounts and either transfer Ether or invoke contract code. Under identical inputs, nodes execute the same bytecode on the Ethereum Virtual Machine and apply the same state transition rules, so replicas that start from the same pre state converge to the same post state [89, 124]. Gas metering prices computation and storage in a uniform unit and enforces bounded execution [124].

The account model concentrates value in accounts and permits internal value movements within a single transaction through contract calls. Fees are paid as a base fee that is burned and as a priority tip to the fee recipient under EIP 1559, which separates a predictable base component from a strategic tip component [27]. Figure 2.9 summarizes balances, internal transfers, and fee components in the account-based setting.

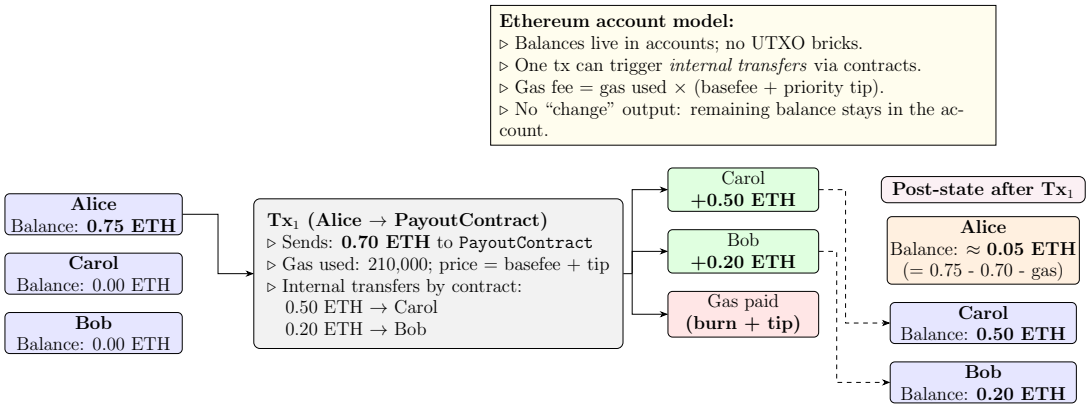


Figure 2.9: Account-based execution and balances. A user transaction can trigger internal value transfers via contract calls. Fees consist of a base fee, a burned, and a priority tip to the fee recipient. There is no explicit change output, and the sender's remaining funds stay in the account balance after execution [27, 124].

From pre-state to post-state, a node validates a transaction for signature correctness, nonce matching, and sufficient balance to cover the upfront gas cost. Then the EVM interprets bytecode as opcodes that may read or write storage and may create or call contracts. Each opcode consumes gas, so the gas remaining decreases as execution progresses. If execution completes, writes and balance updates are applied, and a receipt with logs is recorded. If gas is exhausted or a revert is triggered, state writes are discarded while gas consumed is charged as specified. Determinism follows because all correct nodes execute the same rules on the same inputs [124]. Figure 2.10 depicts this validation and execution pipeline and the role of gas accounting.

Execution outcomes are committed in block headers that link the chain and authenticate data. Each header links to its parent, identifies the fee recipient, and commits to state, transactions, and receipts tries. After Shanghai, withdrawals are also committed, and with EIP 1559, the header records the base fee per gas together with gas accounting fields. Data availability upgrades introduce blob-related fields in the header [35, 124]. These commitments allow clients to verify inclusion and state evolution without replaying all computation. Figure 2.11 presents the main header fields and their roles.

Under EIP 1559, users specify *maxFeePerGas* and *maxPriorityFeePerGas*. The base fee adjusts with demand and is burned, the priority tip is paid to the fee recipient, and any unused portion of the maximum fee is refunded. The elastic block target smooths congestion and reduces volatility in user costs [27, 104].

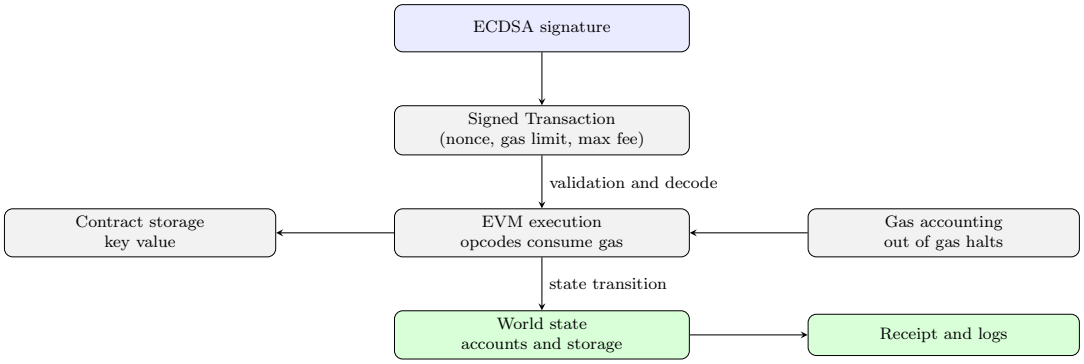


Figure 2.10: Deterministic state transition on the EVM. Validation and decoding precede opcode execution. Opcodes consume gas, which bounds resource usage. On success, world state updates and logs are committed; on out of gas or on revert, writes are discarded while gas used is charged [124].

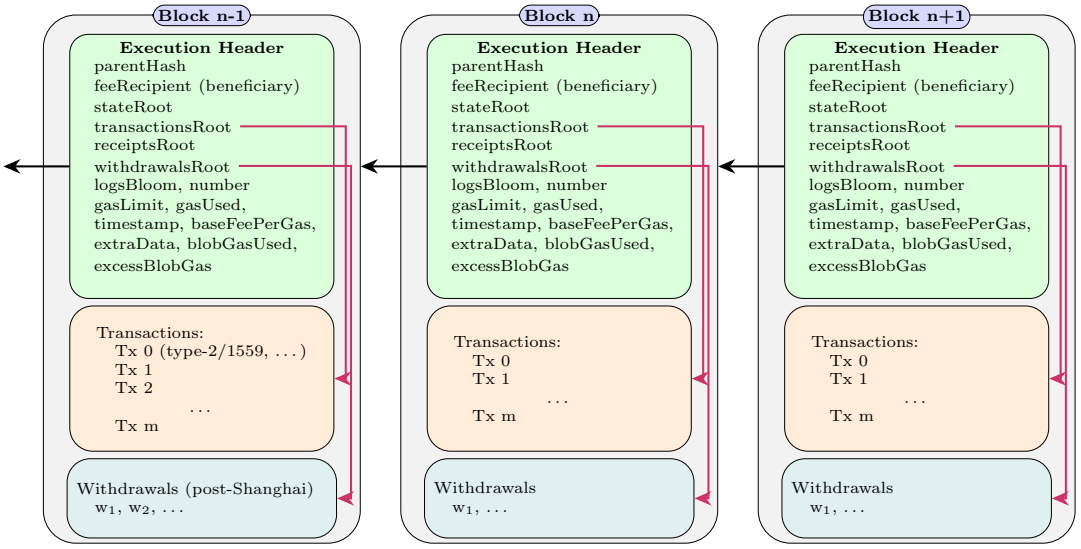


Figure 2.11: Ethereum blocks. The header includes *parentHash* and *beneficiary*, commits to *stateRoot*, *transactionsRoot*, *receiptsRoot*, and after Shanghai to *withdrawalsRoot*. It records *gasUsed*, *gasLimit*, *timestamp*, and *baseFeePerGas*. Recent upgrades add blob related commitments for data availability [35, 124].

Figure 2.12 summarizes the fee fields and the flow of funds during inclusion.

In the proof-of-stake setting, users broadcast signed transactions that enter peer-to-peer gossip. Validators construct execution payloads from their mempool views, the base fee is burned, and the tip steers inclusion and ordering. Attestations drive fork choice, and checkpoints at epoch boundaries acquire economic finality [35]. Figure 2.13 places the execution flow inside consensus.

Ethereum launched on July 30, 2015, with the Frontier genesis block [41], and initially used proof of work where miners expended computational work to propose blocks and the canonical chain followed a longest or heaviest rule [124]. The proof-of-stake research effort known as Casper specified a protocol that

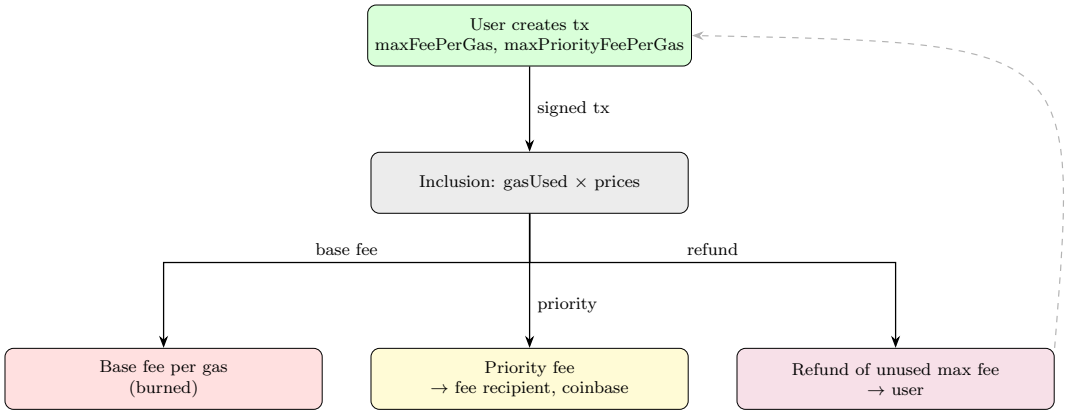


Figure 2.12: Minimal EIP 1559 flow. Inclusion cost equals gas used times the effective price. The base fee per gas is burned. The priority fee goes to the fee recipient. Any unused portion of the maximum fee is refunded to the sender [27, 104].

maintains safety and liveness by placing an economic stake at risk and that enables economic finality with lower energy use [19]. In September 2022, the Merge connected the existing execution layer to the proof-of-stake Beacon Chain. Mining was retired, and a validator set now proposes blocks and attests while the execution semantics of accounts, EVM, and receipts remained unchanged [34, 35].

To give an overview of this handoff, Figure 2.14 depicts the execution layer continuing through the Merge while the consensus role moves from miners to validators. The figure emphasizes continuity at the application layer and the change in who proposes and finalizes blocks [34, 124].

Under proof of stake, time is partitioned into slots of 12 s and epochs of 32 slots. For each slot, one validator is selected using RANDAO-based randomness to propose a block. The other active validators are assigned to committees that attest to the proposed block and to the current head. A validator activates by depositing 32 ETH and participates with an effective balance capped at 32 ETH. Timely proposals and attestations earn rewards. Lateness and inactivity incur penalties. Misbehavior such as proposer equivocation, a double vote, or a surround vote is slashable. If finality stalls, the inactivity leak gradually penalizes offline validators so that an online supermajority can finalize again [19, 35].

Fork choice and finality have separate roles. Latest Message Driven GHOST selects the head by walking from the most recent finalized checkpoint toward the tip, and at each fork, choosing the child with the greatest weight of validators' latest attestations. Figure 2.15 illustrates this greedy walk and the use of the latest votes as weights [35]. Casper the Friendly Finality Gadget places checkpoints at epoch boundaries and finalizes when a checkpoint is justified, and a direct child is justified in the next epoch. Under healthy conditions, finality typically arrives after two epochs, that is $64 \times 12\,\text{s} \approx 12.8$ minutes. Slashing for double votes and surrounding votes protects safety, and inactivity leak restores liveness during extended non-finality. Figure 2.16 shows justification and

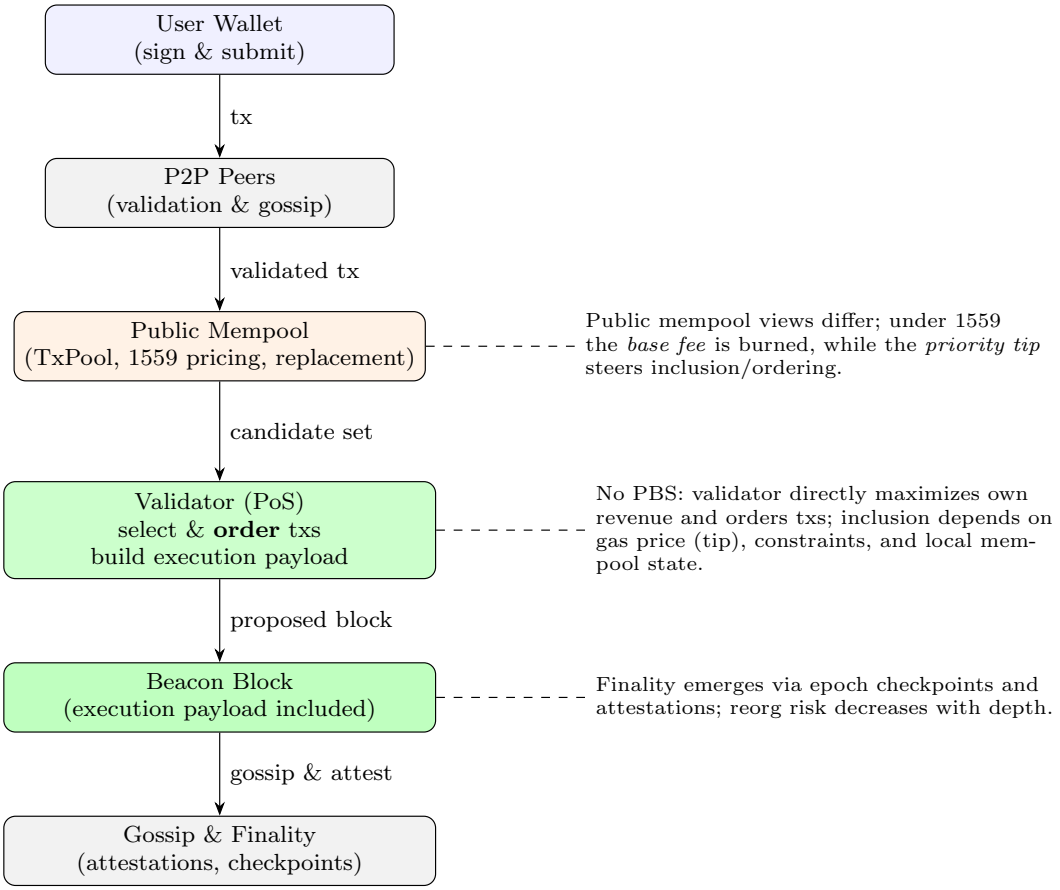


Figure 2.13: Transaction lifecycle under proof of stake and EIP 1559. Users broadcast signed transactions. Validators select and order transactions and propose blocks with execution payloads. Gossip propagates proposals and attestations. Fork choice selects the head and finality confirms checkpoints at the epoch level [35].

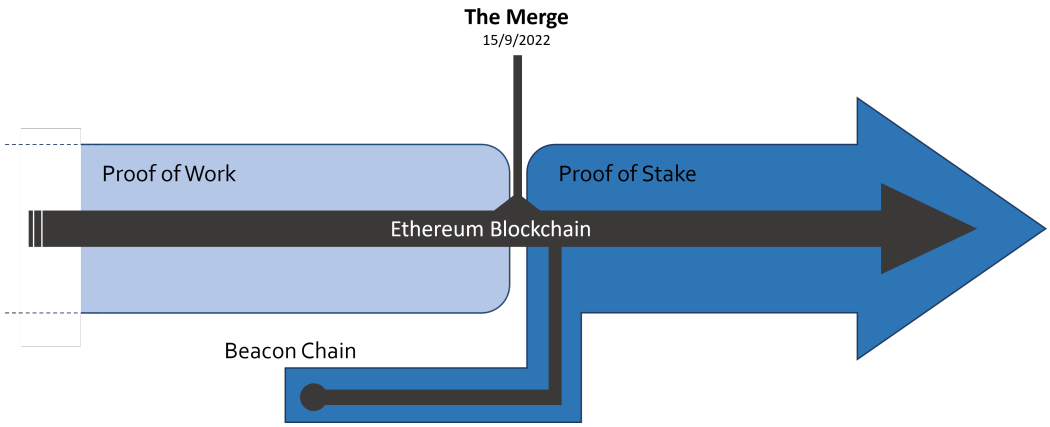


Figure 2.14: Execution layer continuity through the Merge. The EVM and state persist while the proposer role moves from proof of work miners to proof of stake validators on the Beacon Chain [34, 124].

finalization across epochs [19, 35].

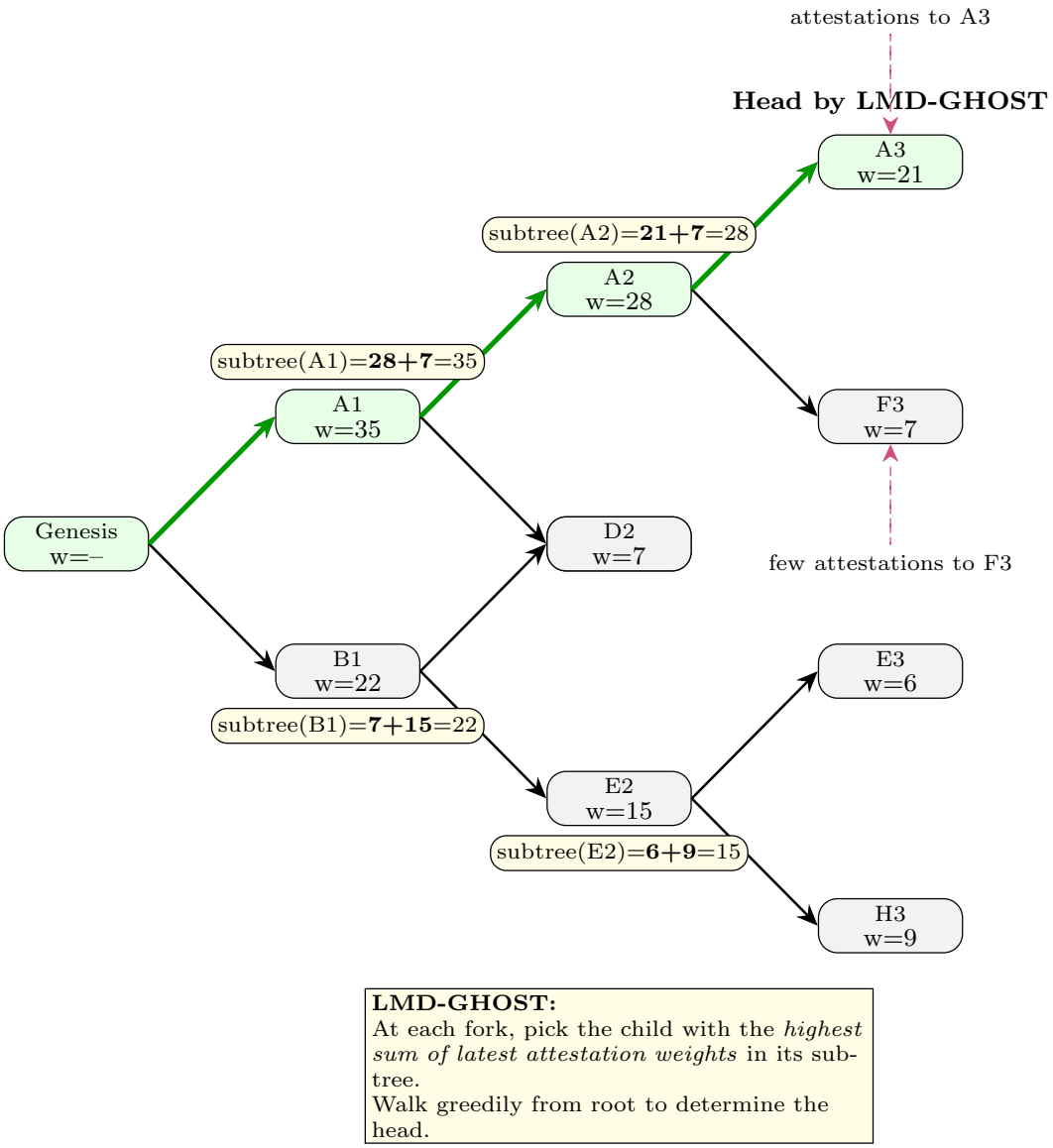


Figure 2.15: LMD GHOST fork choice. Starting from the latest finalized checkpoint, the head is obtained by selecting at each fork the child with the largest subtree weight of validators' latest attestations [35].

These design choices reshape incentives that bear on Maximal Extractable Value. The base fee introduced by EIP 1559 is burned at inclusion time, while priority is paid via tips and via order flow arrangements. After the Merge, many validators outsource block construction to specialized builders through relay-based markets, known as proposer builder separation in practice, so the proposer is compensated for selection and ordering while builders internalize residual value [35, 52]. This subsection focuses on protocol-level mechanics and parameters. Figure 2.17 places the change on a timeline and summarizes the parameters used in this dissertation. Slot duration equals 12 s. Epoch

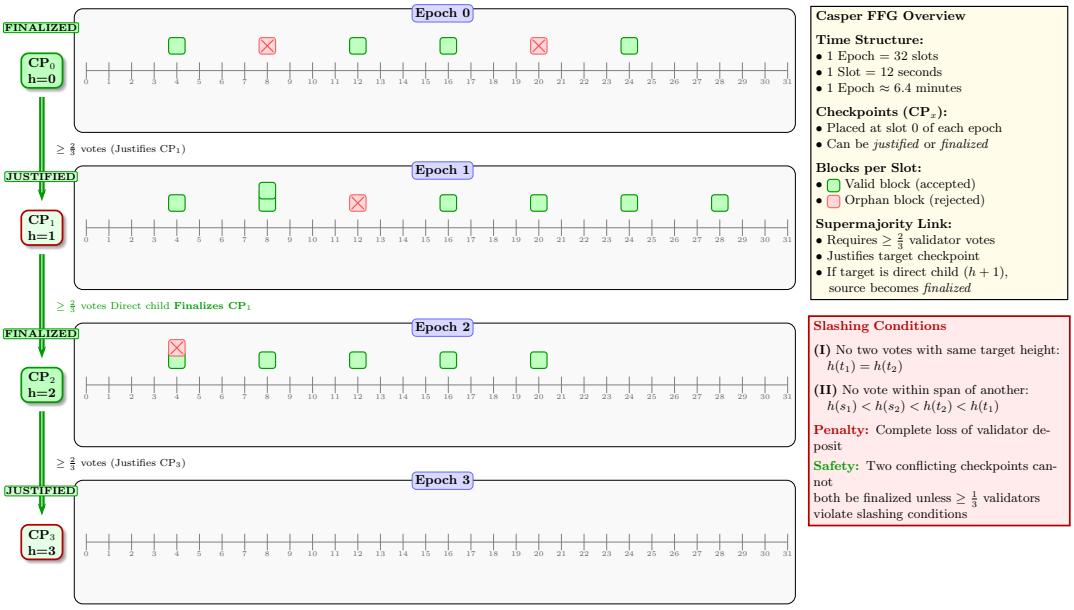


Figure 2.16: Casper FFG finality across epochs. A checkpoint that is justified by a supermajority becomes finalized when a direct child is justified in the next epoch. Slashing rules deter equivocation. Inactivity leak penalizes offline validators and allows finality to resume [19, 35].

length equals 32 slots. Activation deposit equals 32 ETH. Proposer selection uses RANDAO-based randomness. Head selection uses LMD GHOST. Finality uses Casper FFG. Slashing applies to proposer equivocation, double votes, and surround votes. Inactivity leak applies under extended non finality [19, 35].

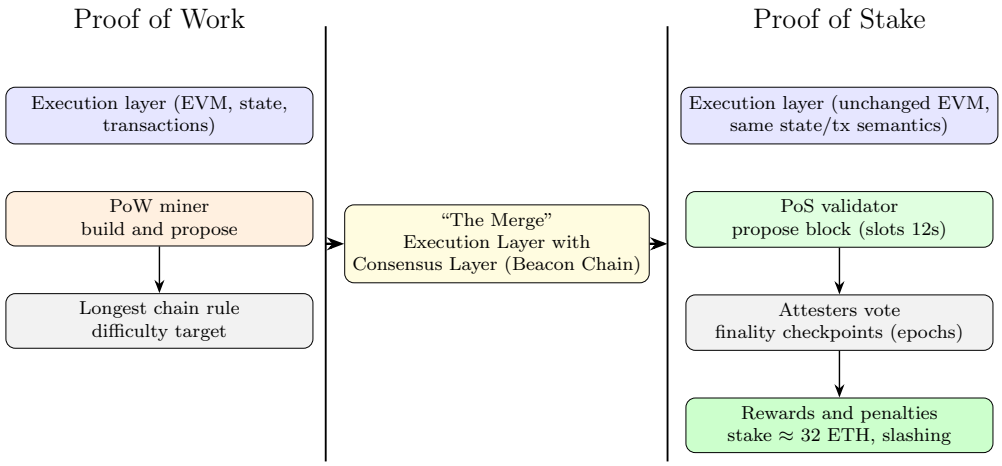


Figure 2.17: Timeline from proof of work to proof of stake [34, 124].

2.4.2 Composability and On Chain Markets

Because accounts can call other accounts directly, on-chain programs compose into higher-level markets. Decentralized exchanges, lending, collateralized stablecoins, liquid staking, and perpetuals emerge as contracts that call other contracts, with deterministic settlement and transparent state [62, 89]. This composability, together with rich price sensitive state, creates a fertile surface for arbitrage, liquidations, and other order sensitive strategies, which later chapters will analyze under the lens of MEV. The programmability that Szabo envisioned in the 1990s, namely "a set of promises specified in digital form", became concrete in Ethereum's EVM and tooling [114, 124].

Historical arc and risks. Ethereum's programmability brought new risks. "The DAO" exploit in 2016 led to a contentious hard fork and a long-running debate about governance, immutability, and the meaning of "code is law" [89]. Scalability constraints and high fees during demand spikes motivated rollups and a roadmap that places much of the execution off-chain while inheriting security from Ethereum [34]. The move to PoS and the introduction of EIP 1559 set the economic foundation for modern block production, which is the backdrop for MEV markets analyzed later [27, 34].

In summary, Ethereum shifted the baseline from currency to a general-purpose ledger plus virtual machine. Its account model and gas metering enable application logic with deterministic global settlement, its shift from PoW to PoS reduced energy and reshaped proposer incentives, and its fee design with EIP 1559 created a predictable but still competitive market for inclusion and position [17, 27, 124].

2.5 Comparative Aside: Ethereum vs. Solana

While Ethereum emphasizes decentralization and security at Layer 1 (and scales via Layer 2), Solana takes a monolithic, high-throughput Layer 1 approach. Solana couples proof of stake with Proof of History, a verifiable delay and clock mechanism, to sequence transactions before consensus and enable very fast confirmation with low fees [125]. Aggressive parallelization and optimized networking underpin Solana's performance envelope [125].

These choices entail trade-offs. Solana's hardware and bandwidth requirements for validators are substantially higher than Ethereum's, contributing to a smaller validator set and raising decentralization concerns relative to Ethereum's broadly accessible staking and validation model [124, 125]. The network has also experienced periods of reduced availability, reflecting the challenges of operating at very high throughput [109]. Conversely, many interactive applications (for example, real-time trading and gaming) benefit from Solana's low latency, low fee environment [125].

Conceptually, Ethereum prioritizes credible neutrality and security at the base layer and leverages rollups for scalability; Solana prioritizes single-chain

performance. Both positions illuminate the design space, and trade-offs for public ledgers [124, 125].

2.6 Maximal Extractable Value: definitions, attacks and strategies

Maximal Extractable Value (MEV) refers to the value that can be captured by actors who can influence transaction inclusion and ordering, or who can otherwise shape the effective execution path of user intent. [29, 37] While MEV is often introduced through simple examples, its practical relevance depends on the gap between *potential* opportunities and *realized* revenues, and on how realized revenues are distributed across the execution pipeline. [47, 96] This section provides the minimal conceptual toolkit used throughout the thesis, without aiming to be exhaustive.

2.6.1 From miner-centric extraction to role-agnostic MEV, from potential to realized

The original "MEV moment" is commonly associated with "Flash Boys 2.0" [29], which framed miner-controlled ordering in proof of work as a consensus-level externality and documented how public mempool competition can induce priority gas auctions and destabilizing incentives. As networks moved toward proof of stake and outsourced block construction, the terminology and the threat model broadened. The term *maximal* emphasizes that extractable value is not specific to miners; it is available to any actor, or coalition of actors, that gains effective control over ordering, inclusion, or execution conditions [37, 60, 83].

In this thesis, a central distinction is between *potential MEV*, the set of opportunities that exist in principle given a state and an ordering mechanism, and *realized MEV*, the revenues that are actually captured in practice. This distinction matters because changes in architecture, and changes in market structure, can raise conversion efficiency, meaning that a larger fraction of potential opportunities becomes executable and monetizable. The empirical chapters study how this conversion efficiency changes under modern execution pipelines, especially under PBS, relays, and private routing [48, 64, 122].

2.6.2 Canonical MEV attack families

Throughout this thesis, we treat MEV not only as a source of revenue, but also as a family of *attacks* enabled by discretionary ordering, state dependence, and pre-inclusion visibility. The term "attack" is appropriate because these behaviors typically impose negative externalities on other participants, for example, worse execution for users, higher fees due to bidding wars, or selective access to opportunities [29, 33]. By contrast, we use "strategies" to refer to *execution and market organization choices*, such as private routing, bundling, order flow

auctions, or PBS mediated block building, which shape how extraction is operationalized. With this distinction, we summarize the canonical MEV attack families used as reference points in the empirical chapters.

Arbitrage. Arbitrage attacks exploit temporary price inconsistencies across venues or paths, and extract value by executing first, before prices converge [96]. A practical example is DEX arbitrage. Suppose token X trades at 100 USDC on DEX_1 and at 101 USDC on DEX_2. An attacker buys X on DEX_1 and immediately sells X on DEX_2, capturing the spread net of fees and gas. The attack dimension arises because the opportunity is time sensitive and competition for priority can impose costs on others, for example, by escalating fees, congesting blocks, and crowding out ordinary users.

Frontrunning. In this thesis, "frontrunning" denotes the attacker observing a pending transaction that creates or reveals an MEV opportunity, and preempting it by securing earlier execution [29]. A practical example is copying a large pending swap. Suppose a user submits a large buy of token X on a DEX, which will move the price upward once executed. An attacker observes the pending transaction and submits its own buy of X with higher priority. In a public mempool setting, this is typically done by outbidding fees, creating a gas war. In a builder-mediated pipeline, priority can also be purchased through explicit builder payments that influence ordering. If the attacker executes first, it buys at the pre-trade price and can then sell after the user's trade moves the price, capturing value that the user effectively created through price impact.

Sandwich attacks. A sandwich attack places two attacker transactions around a victim trade, extracting value from the victim's slippage bound [96, 130]. A practical example is a user swap with tolerance. Suppose a user swaps USDC for token X with a maximum slippage of 1%. The attacker first executes a *frontrun* buy of X, pushing the price up. The victim trade then executes at a worse price but remains within the 1% bound, so it does not revert. The attacker then executes a *backrun* sell of X, unwinding the position at the inflated price. The victim transaction is thus sandwiched between the attacker's legs, and the attacker captures the spread generated by the victim's price impact and fees. This attack is a direct fairness concern because it targets ordinary users and converts their slippage into extractor profit.

Liquidation attacks. Liquidation MEV arises when a lending protocol allows third parties to liquidate under collateralized positions for a reward [96]. A practical example is a collateral ratio threshold. Suppose an account has borrowed stablecoins against ETH collateral. If the ETH price drops, the position becomes eligible for liquidation, and the protocol may allow a liquidator to repay part of the debt and receive collateral at a discount. Attackers monitor the system for such threshold crossings and race to submit the first valid liquidation transaction. The first included liquidation captures the bonus, while losing transactions revert or become unprofitable, and the competition can escalate

fees and amplify volatility periods by tightly coupling liquidations to block-level priority contests.

These canonical attack families highlight common prerequisites. Profitability depends on state-dependent execution, on the ability to observe triggering transactions or conditions before inclusion, and on the ability to secure inclusion and ordering constraints. The rest of the thesis studies how modern execution organization, especially PBS, relay mediation, and private order flow, change these prerequisites, raising conversion efficiency and reshaping both the distribution of revenues and the governance trade-offs that follow.

2.6.3 MEV execution strategies

The canonical MEV attacks described above are enabled or constrained by *how* transactions propagate and *who* can influence inclusion and intra-block ordering. In the public mempool model, transactions are broadcast via gossip before inclusion, which maximizes pre-inclusion visibility and competition, but also enables copying and priority races [29, 47]. Two roles are central in this setting. *Searchers* are specialized actors that monitor pending transactions and on-chain state to detect profitable opportunities, and then submit transactions or bundles intended to realize them. *Validators* are the block proposers in proof of stake (or *miners* in proof of work), and they ultimately decide which transactions become part of the canonical chain by proposing blocks, even when they outsource block construction [37].

In response to public mempool adversariality, the ecosystem increasingly relies on private submission channels. In this thesis, we use "private order flow" to denote transaction submissions routed outside the public gossip mempool and delivered directly to a restricted set of intermediaries, reducing broad pre inclusion visibility and shifting informational advantage toward those intermediaries [53, 61]. This execution strategy changes how searchers compete, because copying and gas wars can be reduced for participants inside private routes, while exclusion from those routes can disadvantage others.

These dynamics become particularly salient under proposer builder separation (PBS). PBS separates the role of assembling a block from the role of proposing it to the network [39]. *Builders* are specialized entities that construct candidate blocks by selecting and ordering transactions, often integrating searcher bundles and private order flow, and they compete by offering bids that share value with the proposer. *Relays* are infrastructure components that facilitate this market by receiving blocks and bids from builders and forwarding them to validators, often providing an interface that reduces trust and limits information leakage, for example by allowing validators to compare bids without directly revealing full block contents until selection [52, 64]. *Validators*, as proposers, select among builder bids and publish the chosen block, capturing revenue via the bid while avoiding the operational complexity of running a competitive block building operation themselves [39].

From a measurement perspective, these execution strategies and roles change what can be observed. If a growing share of flow is routed privately and blocks

are assembled off-chain by builders, mempool-level visibility becomes partial by design, and empirical work must rely on conservative proxies and complementary sources. In this thesis, we use *Mempool Dumpster* [55] as a public propagation baseline for distinguishing public from non-public paths, and we use *ZeroMEV* [128] labels as a disciplined reference when MEV attacks tagging is required, while recognizing that both sources provide partial views of inherently opaque routing decisions [55, 128].

This framing motivates the empirical design of the thesis. The Ethereum chapters study how off-chain block building and relay mediated PBS pipelines affect realized MEV, reward flows, and ordering outcomes [48,122]. The Solana chapter provides a lower bound view of MEV participation through the Jito stack, illustrating how a high throughput architecture supports a distinct MEV market organization [71–73]. The later chapters connect these measurements to fairness, governance, and network structure, and extend the measurement lens to cross-chain settings.

2.7 Related Work

The literature on MEV can be read as a chronicle of successive shifts in who extracts value, what is observable, and which infrastructures arise to operationalize, or mitigate, extraction [29, 60, 83, 86]. This section follows that evolution, starting from early observations of adversarial ordering and frontrunning, then moving through the formalization and generalization of MEV, and finally focusing on realized extraction under proof of stake pipelines and on the cross domain setting that emerges once execution spans multiple ledgers and intermediaries [93, 127].

Concerns about adversarial ordering predate the term MEV, and surfaced early both in practitioner discussions and in academic security work on transaction manipulation. An early practitioner signal is the Reddit discussion on miner frontrunning [102]. A canonical synthesis is the SoK by Eskandari et al., which systematizes front running as a multi-layer phenomenon spanning network visibility, transaction replacement, and block producer discretion [33]. The "MEV moment" is typically associated with Daian et al., who formalize *Miner Extractable Value* and demonstrate how transaction reordering in decentralized exchanges can create consensus-level externalities, including time bandit incentives in Ethereum proof of work [29]. Practitioner work reinforced the intuition that public propagation can become adversarial for ordinary users, motivating private submission and specialized relaying, and popularizing the "dark forest" metaphor [103], which frames the mempool as a hostile environment where publicly visible transactions are continuously monitored and can be opportunistically "preyed upon" through adversarial reordering and insertion. Follow-up measurement work then quantifies the empirical prevalence of frontrunning at scale and connects it to mempool observability and priority races [117].

As DeFi matured, attention shifted from isolated attacks to systematic measurement of distinct MEV families and to the market microstructure that emerges

from priority competition. Qin et al. quantify "blockchain extractable value" in Ethereum proof of work by reconstructing arbitrage, liquidation, and sandwich patterns, emphasizing that realized profits depend on competition costs and on the rules that determine ordering [96]. Related empirical work studies specific arbitrage mechanisms, for example, cyclic arbitrage paths, and connects them to congestion and competition dynamics [123]. In addition, private transaction routing becomes a first-order factor for both security and measurement, because it changes who can observe an opportunity before inclusion. Lyu et al. analyze Ethereum private transaction pools and articulate the security implications of bypassing public propagation, including changes in transparency and adversarial exposure [76].

A second transition is conceptual. The term shifts from miner extractable value to *Maximal Extractable Value*, reflecting that extraction is not tied to proof-of-work miners, but to any actor, or coalition of actors, that can influence inclusion and ordering. Surveys and SoKs consolidate this broader perspective and provide taxonomies by strategy type, layer, and actor [37, 60, 83]. On the mitigation side, Yang et al. systematize countermeasures and evaluate their practical limits, clarifying trade offs among privacy, efficiency, and neutrality goals [126]. Policy-oriented syntheses further highlight fairness and market integrity concerns, and emphasize that measurement remains difficult precisely because an increasing share of activity is intermediated and privately routed [44].

A third transition is empirical and architectural, moving from potential extractable value toward realized extraction under specialized execution pipelines. In Ethereum proof of stake, PBS oriented pipelines, implemented today via relays and out-of-protocol builder markets, reorganize the MEV supply chain and make private order flow operationally central [52, 64, 122]. Measurements of the post-Merge builder market study relay mediation, concentration, reward redistribution via builder payments, and shifts in validator incentives [78, 122]. Additional theoretical and empirical work models how the proof of stake slot structure induces timing games, where proposers may trade off propagation speed against time dependent value, and where network latency becomes economically relevant [105]. At the strategy level, the literature increasingly recognizes that a large share of realized value is non atomic, combining on-chain legs with off-chain venues. Heimbach et al. measure non-atomic arbitrage and connect it to concentration among a small set of sophisticated actors and to the industrialization of extraction in builder-mediated blocks [63]. Other work studies bidding in private MEV markets, including learning based approaches to auction strategies [99]. The same realized lens extends to rollups and layer 2 environments, where sequencer designs and different visibility assumptions change which attacks are feasible and how competition is expressed [5, 118].

A fourth transition is cross-domain, where extractable value depends on coordinating actions across multiple ledgers, execution environments, or venues. This line builds on foundational interoperability research. Herlihy et al. formalize cross-chain deals under adversarial settings, providing protocol abstractions for atomic, or trust-minimized, cross-chain interactions [66]. Zamyatin et al.

systematize communication across distributed ledgers and provide a taxonomy of trust models and connector mechanisms [127]. On the conceptual MEV side, Obadia et al. propose a formalization of cross-domain MEV and highlight how cross-domain coordination creates incentives for collusion among sequencers across domains [93]. McMenamin synthesizes cross-domain MEV and clarifies how shared sequencers, bridges, and interoperability layers alter the unit of analysis from a single chain to an ecosystem of domains [86]. Empirically, early quantification efforts study cross-domain extraction and connect it to decentralization questions [107]. More recent work measures realized cross-chain arbitrage frequency and concentration, and extends to cross-rollup settings and non-atomic routes that span multiple layer 2s, emphasizing that hop length and bridge latency compound feasibility constraints [59, 85].

Positioning of this thesis. Relative to this trajectory, the thesis contributes a measurement-centered view of realized MEV that is explicitly cross-era and cross-architecture. It treats PBS and private order flow as mechanisms that raise conversion efficiency from theoretical opportunity to realized revenue, and then studies how that shift interacts with decentralization and fairness through reward allocation, ordering outcomes, and network structure [77, 78]. Methodologically, it complements existing labeling and dashboard approaches by combining a public propagation baseline with MEV attacks oriented labels, and by constructing graph representations that separate direct ties from PBS mediated ties, enabling community level analyses of organizational structure [55, 128]. Finally, it extends the empirical lens to cross-chain settings through a constrained multi hop detection framework, connecting the cross domain formalizations and early measurements to a high resolution view of realized, but rare, multi hop extraction [79].

3

Empirical Analysis of MEV on Ethereum

This chapter addresses **RQ1** by quantifying the *prevalence* of MEV-linked execution on Ethereum across eras, specifically before and after the Merge. The goal is to measure how often Ethereum block production is mediated by private, off-chain MEV infrastructure, and how this share evolves from proof of work to proof of stake [77, 78].

3.1 Ethereum's Off-Chain MEV Pipeline

Ethereum's modern MEV pipeline is an off-chain market for block construction. Searchers submit bundles or private order flow to builders, builders assemble candidate blocks and compute bids, relays mediate communication and perform basic validity checks, and validators select the highest paying bid for the slot, typically via MEV Boost (proposer builder separation architecture). This chapter uses the high-level pipeline in Figure 3.1 as the reference abstraction for the prevalence measurements that follow.

At the left edge of Figure 3.1 stand *searchers*, automated agents that monitor transaction flows to identify MEV opportunities, for example, arbitrage, liquidations, and sandwich style extraction. In practice, searchers consume signals from the public mempool and from private submission paths, then produce candidate bundles whose internal ordering and inclusion constraints encode the intended strategy. At the right edge lie *validators*, the entities selected by Ethereum's proof of stake protocol to propose blocks for specific slots.

Between these endpoints, two layers operationalize an off-chain market for block construction.

- **Builders.** Builders aggregate transactions and bundles into full candidate blocks, choosing composition and intra-block ordering to maximize expected profit subject to protocol rules and gas constraints. In operational terms, builders combine public mempool traffic with private order flow and with many simultaneous searcher bundles, then solve an optimization problem that must be repeated for each slot.

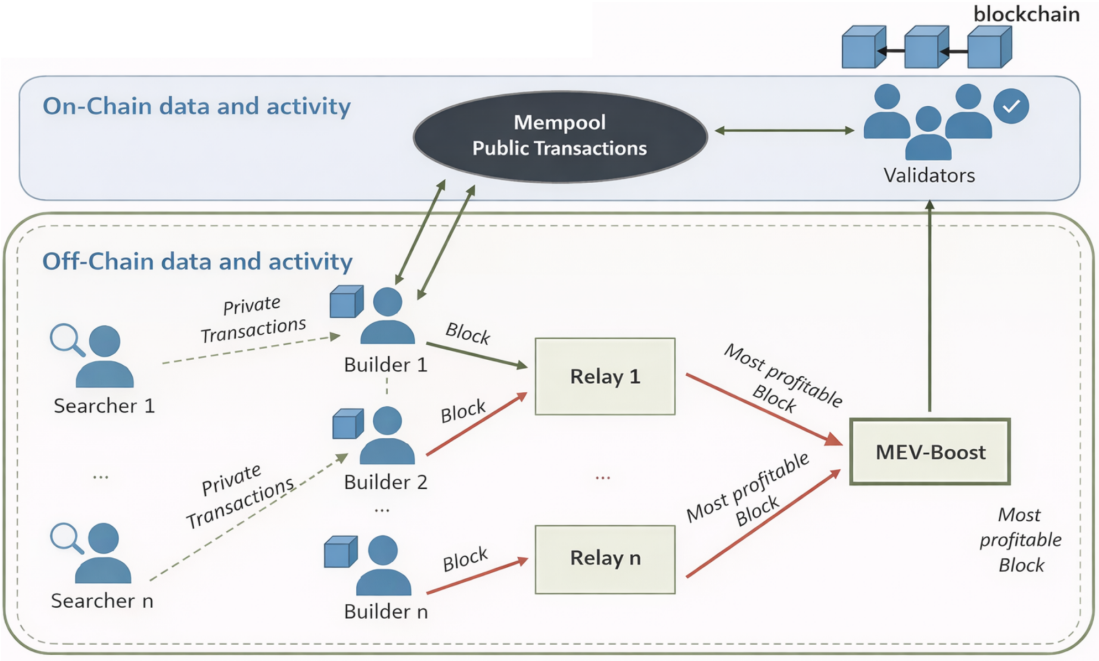


Figure 3.1: Sketch of Ethereum's off-chain MEV ecosystem. Searchers submit bundles to builders, builders assemble candidate blocks and bid via relays, validators select the highest paying block, commonly through MEV Boost.

- **Relays and MEV Boost.** Relays receive candidate blocks from many builders, perform basic validity checks, and forward only the highest paying bid for each slot to the validator. MEV Boost is a client-side middleware that allows validators to connect to a set of relays and select the most profitable bid available for the current slot, without revealing the full block to the proposer before commitment.

A key convention in this ecosystem is the way builders and validators are paid, and this convention is central to the prevalence measurement [77, 78], builders typically:

1. set the *feeRecipient* field of the block header to a builder-controlled address, so priority fees accrue to the builder,

2. insert a final transaction that transfers a negotiated payment from a builder-controlled address to the validator reward address for that slot.

This pattern implies that, in MEV-mediated blocks, the header fee recipient is usually *not* the validator. Figure 3.2 illustrates the canonical structure for a builder assembled block, and it motivates why naive inspection of *feeRecipient* can misattribute builders as if they were proposers, a point emphasized in the post-Merge analysis in [77] and generalized at scale in [78].

From an empirical perspective, the convention in Figure 3.2 provides a robust signal that a block was produced through the off-chain pipeline in Figure 3.1. This chapter exploits that signal, together with relay metadata, to clas-

block number	16474187			
timestamp	1674533363			
fee recipient	0xdaf...98bc5			
num tx	126			
tx index	**from**	**to**	**fee (ETH)**	**value**
0	xxxx	xxxx	0.0000075	xxxx
1	xxxx	xxxx	0.0000000	xxxx
2	xxxx	xxxx	0.0003621	xxxx
...	...	...	...	...
125	**0xdaf...98bc5**	**0xcEC...8867F**	0.0000000	**0.055523**

Figure 3.2: Simplified excerpt of a post-Merge Ethereum block constructed by a builder. The header fee recipient is a builder address, and a final transaction transfers a builder payment to the validator reward address.

sify blocks and to quantify the prevalence of MEV-mediated execution across eras.

3.2 Data and Measurement Setup

The central output for **RQ1** in Ethereum is a time series of MEV-mediated block share, complemented by aggregate prevalence over long windows. To produce this, we combine on-chain block data with off-chain relay metadata.

3.2.1 On-chain blocks and transactions

We ingest block headers and full transaction lists. In the proof-of-stake period, we additionally inspect the last transaction of each block to detect explicit builder payments, as motivated in Section 3.1. The analyzed ranges are:

- **Proof of stake period.** From block 15,537,394 (the Merge, 15 September 2022) to block 18,174,261 (20 September 2023), totaling 2,636,868 consecutive blocks.
- **Proof of work period.** From block 11,834,049 (11 February 2021) to block 15,537,393 (immediately before the Merge).

Overall, the dataset spans 6,340,213 blocks, enabling a direct before and after comparison for **RQ1**.

3.2.2 Off-chain relays and coverage

For the proof-of-stake period, we query public relay endpoints and collect, for each delivered block, the block identifiers and the builder and proposer metadata required to link the off-chain delivery to the on-chain block. Since relay coverage is not uniform and some relays are discontinued, the measurement explicitly tracks which relays are included and their operational status.

Table 3.1 reports the relays considered, together with their claimed censorship and filtering posture and, crucially for interpreting time coverage, their decommission dates when applicable.

Table 3.1: Details and description of relays.

Relay	Uncensored	Unfiltered	Status (as of 11/23/2023)
Aestus [1]	✓	✓	Active
Agnostic Relay [101]	✓	✓	Active
Blocknative [10]	✗	?	Inactive (as of 27/09/2023) [25]
bloXroute Ethical [11]	✓	✗	Inactive (as of 08/11/2023) [12]
bloXroute Max Profit [11]	✓	✓	Active
bloXroute Regulated [11]	✗	✓	Active
Eden Network [91]	✗	?	Active
Flashbots [24]	✗	✓	Active
Manifold [106]	✓	✓	Active
Relayooor [15]	✓	✓	Inactive (as of 03/23/2023) [16]
Ultrasound [121]	✓	✓	Active

3.3 Operational Definition of MEV Prevalence

RQ1 asks about the prevalence of MEV across eras and architectures, and, for Ethereum, this chapter answers it at the level of block production. The empirical question becomes, for each block, whether its construction was mediated by a specialized MEV pipeline, rather than being built solely from the public mempool by the proposer.

We therefore classify blocks into two categories, using signals that are directly tied to the architecture in Figure 3.1 and to the payment convention in Figure 3.2.

3.3.1 Post-Merge classification

In the proof-of-stake period, a block is classified as *MEV-mediated* if it can be linked to at least one relay delivery record, and the on-chain structure is consistent with builder construction. Concretely, the classification leverages:

- **Relay linkage.** A relay reports the delivered block identifiers for the slot, enabling a direct match to an on-chain block (by block hash).
- **Builder recipient patterns.** The block header fee recipient matches a builder-controlled address observed in relay metadata, consistent with builders capturing priority fees before paying the proposer.
- **Builder payment consistency.** When present, the final transaction conveys a payment to the validator reward address, consistent with Figure 3.2.

Blocks that do not match these signals are classified as *no MEV mediated*, meaning that they are not observed to be delivered through the relays in Table 3.1, and thus are treated as blocks built outside the measured off-chain channel set.

3.3.2 Pre-Merge classification

In the proof-of-work period, proposer builder separation and MEV Boost do not exist; the post-Merge relay linkage procedure cannot be applied directly.

However, **RQ1** still requires an operational and comparable notion of *off-chain MEV mediation* before the Merge. So, we treat blocks processed through *Flashbots MEV-Geth* as the primary observable proxy for off-chain MEV execution in pre-Merge Ethereum.

Thus, based on the information available via the MEV-blocks API [54], we discriminate all the blocks that were processed through Flashbots MEV-Geth out of the blocks incorporated in Ethereum since February 11, 2021, and up to the date of the Merge on September 15, 2022. In summary, for the pre-Merge period, we classify 1,755,221 blocks related to off-chain MEV, out of the Ethereum blocks in the range between block number 11,834,049 and block number 15,537,393. By adding the number of post-Merge and pre-Merge blocks collected, we obtain a dataset of 6,340,213 blocks [78].

Concretely, a block in the proof of work range is classified as *MEV mediated* if the MEV-blocks API labels it as processed through Flashbots MEV-Geth [54]. All remaining blocks are classified as *no MEV mediated*. As in the post-Merge case, this definition is conservative and transparent; it only counts pre-Merge MEV mediation when it can be linked to a specific, measured off-chain channel, and it does not attempt to infer MEV solely from behavioral patterns.

3.3.3 Interpretation and lower bound nature

The two era definitions above provide a block-level notion of MEV mediation directly tied to the presence of specialized off-chain mechanisms, Flashbots MEV-Geth in the pre-Merge period, and relay-mediated delivery under MEV Boost in the post-Merge period. The resulting prevalence estimates should be read as *lower bounds* on the true share of MEV-related activity.

3.4 Prevalence of MEV Mediated Blocks Across Eras

We now report the core quantitative output for Ethereum in **RQ1**, namely the prevalence of MEV-mediated blocks over time and in aggregate. Figure 3.3 plots the share of MEV mediated blocks and no MEV mediated blocks, aggregated in buckets of 100,000 blocks, across both proof of work and proof of stake, with the Merge marked explicitly. This figure is the central artifact for comparing prevalence before and after the Merge under a consistent operational definition.

Two broad regimes emerge. In the proof-of-work window, the measured share of MEV-mediated blocks is substantial but not saturating, consistent with private Flashbots mediation coexisting with public mempool-driven block construction [78]. After the Merge, the measured prevalence increases sharply as MEV Boost adoption scales, and the most recent buckets approach near saturation. Over the full proof-of-stake period considered, MEV-mediated blocks account for 85.15% of all blocks [78].

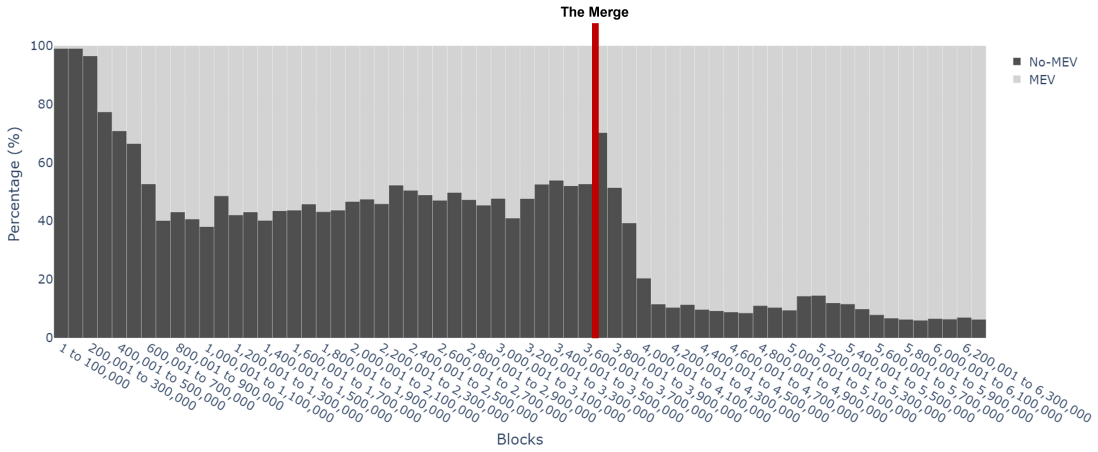


Figure 3.3: Share of MEV mediated blocks and no MEV mediated blocks in groups of 100,000 blocks, spanning proof of work and proof of stake. The vertical marker highlights the Merge.

3.5 Reward Signals That Validate MEV Mediation

While Figure 3.3 answers **RQ1** at the block share level, we also verify that the MEV-mediated label corresponds to a distinct execution and compensation pipeline, rather than to an arbitrary partition. The payment convention described in Section 3.1 and illustrated in Figure 3.2 provides a concrete validation signal, in MEV-mediated blocks, builders commonly set themselves as *feeRecipient* to capture priority fees and then compensate the validator via an explicit builder payment.

Figure 3.4 makes this mechanism observable by reporting the distribution of *validator rewards* in two settings. First, it compares validator rewards in MEV-mediated blocks versus no MEV-mediated blocks. In no MEV-mediated blocks, the validator reward is primarily driven by priority tips accrued through the standard on-chain execution path. In MEV-mediated blocks, the validator reward additionally includes the explicit builder payment negotiated in the relay market, which shifts the distribution upward and induces a heavier tail, consistent with the off-chain pipeline in Figure 3.1.

Crucially, this difference is not only qualitative. The reward distributions in Figure 3.4 show that validator rewards in MEV-mediated blocks are *statistically higher* than those in no MEV-mediated blocks, as confirmed by the Mann-Whitney U test ($p < 0.001$) [78]. This provides an incentive-based explanation for the high prevalence observed in Figure 3.3, validators rationally prefer MEV-mediated blocks because, on average, they yield greater realized compensation than relying on priority tips alone.

Second, Figure 3.4 decomposes the MEV-mediated rewards by *relay* and by *builder identity*. These breakdowns highlight heterogeneity across venues and across the most active builders, but they preserve the same overall pattern: MEV mediation corresponds to a structured and measurable reward process that is

distinct from the direct on-chain path. Together, these distributions validate the MEV-mediated label used for **RQ1**, the category is not only prevalent in terms of block share (Figure 3.3), it also aligns with a statistically superior reward mechanism that helps explain its widespread adoption [77, 78].

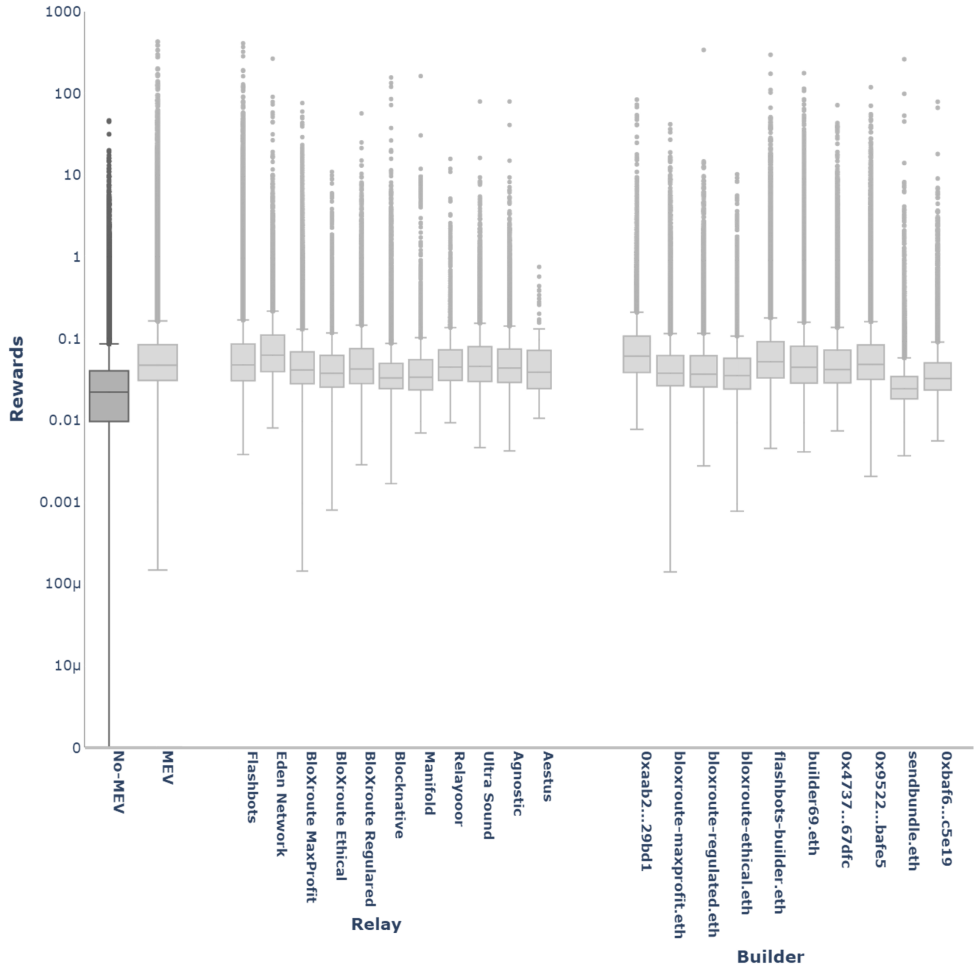


Figure 3.4: Reward distributions for validators in MEV-mediated blocks and no MEV-mediated blocks. For MEV-mediated blocks, the figure further breaks down validator rewards by relay and by the top builders in terms of finalized blocks.

3.6 Attribution Pitfalls Under High MEV Mediation

A second validation and interpretation step is to show how high MEV mediation prevalence changes surface-level on-chain observables, and, in turn, how this can distort any attempt to reason about validator rewards and concentration from finalized blocks alone. Under PBS and widespread relay mediation, the block header *feeRecipient* is frequently a builder-controlled address rather than the validator's reward address, because builders capture priority fees and compensate validators through explicit payments as described in Section 3.1.

As a consequence, aggregations that treat *feeRecipient* as the validator can produce misleading narratives, not only by suggesting that builders "produce" most blocks, but also by hiding which validators are actually proposing blocks and earning slot-level compensation.

We illustrate this attribution problem with three complementary pie charts that progressively refine what is being measured. Figure 3.5 reports the fee recipient distribution in the first 1,000 blocks after the Merge. In this early window, the distribution exhibits a strong concentration among a small number of recipient entities, and, crucially, these recipients largely correspond to validators, thereby reproducing the common observation that block validation is dominated by a limited set of large players.

Figure 3.6 repeats the same aggregation over the full post-Merge dataset. The resulting picture changes qualitatively; the dominant fee recipients are no longer validator entities, but builder identities. This does not mean that the validators disappeared or stopped proposing blocks. Rather, it is an expected footprint of high MEV mediation prevalence, where builders build most blocks; the header fee recipient shifts to the builder's side and therefore masks the validator's side reward recipients if one only inspects finalized blocks. In other words, after PBS, fee recipient-based aggregation primarily reveals who captures priority fees, not who actually proposes blocks and receives slot compensation.

Finally, Figure 3.7 restores the intended interpretation by correcting the attribution. For each MEV-mediated block, we replace the builder fee recipient with the corresponding validator identified through relay linkage and the builder payment signal motivated by Figure 3.2. For no MEV-mediated blocks, we keep the validator identity inferred directly from the on-chain *feeRecipient*. The corrected distribution in Figure 3.7 therefore represents the same dataset as Figure 3.6, but with builder addresses substituted by the specific validators responsible for proposing each block. The key takeaway is that, after PBS, tracking validator rewards and concentration becomes inherently more complex because the relevant reward flows are split across on-chain header fields and off-chain mediation metadata. Accurate measurement, therefore, requires additional steps beyond finalized blocks alone, namely distinguishing MEV-mediated from non-MEV-mediated blocks and attributing builder-constructed blocks back to their proposers using off-chain information [77, 78].

3.7 Builder Competition as an Extent Signal of MEV Mediation

Finally, **RQ1** asks not only for the share of blocks mediated by MEV infrastructure, but also for the "extent" of MEV-linked mechanisms in the execution architecture. In post-Merge Ethereum, an operational way to quantify this extent is to measure the intensity of relay side competition, namely, how many candidate blocks builders submit per slot. This quantity is directly connected to the off-chain market in Figure 3.1, and it becomes observable precisely because MEV mediation is prevalent.

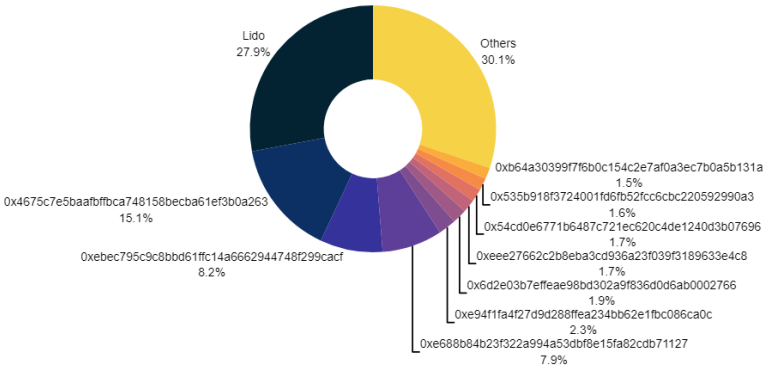


Figure 3.5: Fee recipient distribution in the first 1,000 blocks after the Merge. This early window highlights concentration among a small number of validator recipients.

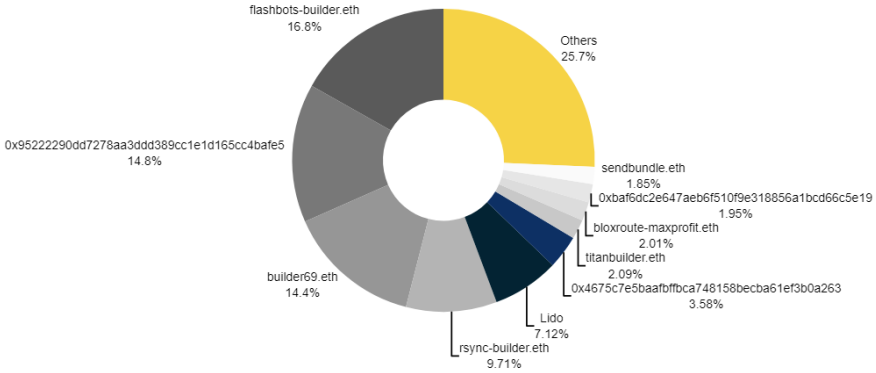


Figure 3.6: Fee recipient distribution over the full proof of stake window. Under high MEV mediation, fee recipients are dominated by builders, masking the validator side recipients if one relies on finalized blocks alone.

Following [78], we focus on the latest 1,000 MEV mediated blocks finalized in Ethereum up to September 20, 2023, and we extract from relay logs all bid submissions that builders sent when competing for those slots. We collect submissions from the subset of relays exposing the *builder_blocks_received* endpoint [51], namely *Aestus*, *Agnostic Relay*, *Manifold*, *Flashbots*, and *Ultrasound*. For readability, we sort slots by the total number of submissions they received.

We report two complementary views. Figure 3.8 counts all builder submissions per slot across the five relays, including duplicates when identical candidate blocks are sent to multiple relays. In this view, the median competition intensity is 720 builder bids per slot, and the maximum reaches 2,230 submissions for a single slot, quantifying the scale of off-chain redundancy induced by the market.

We then refine the analysis to account for the fact that a builder may multicast the same candidate block to multiple relays. We therefore discard duplicate

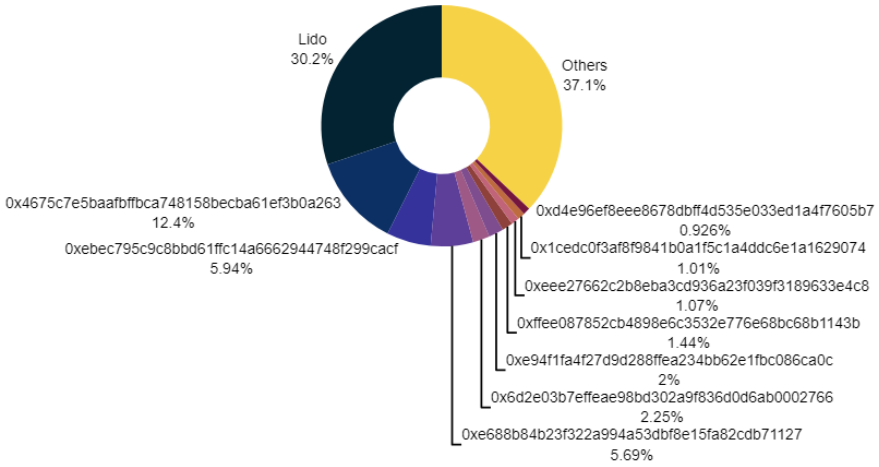


Figure 3.7: Validator distribution after correcting fee recipient based attribution using relay linkage and builder payment signals. Builder addresses in Figure 3.6 are replaced with the specific validators associated with each block.

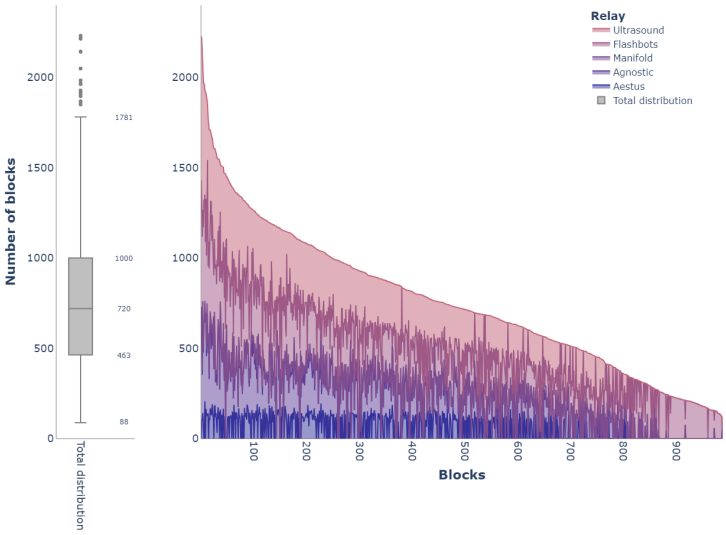


Figure 3.8: Number of blocks sent by builders to each relay to win a slot.

submissions of blocks that share the same hash (that is, the same block submitted by the same builder to another relay), and we count unique block hashes per slot. Figure 3.9 reports the results after this refinement. Even after deduplication, the median remains high, 471 unique candidate blocks per slot, and the maximum reaches 1,295 unique blocks for a single slot. This indicates that the off-chain market is not only redundant because of relay multicast, but also genuinely diverse in candidate block production, consistent with a highly competitive builder landscape.

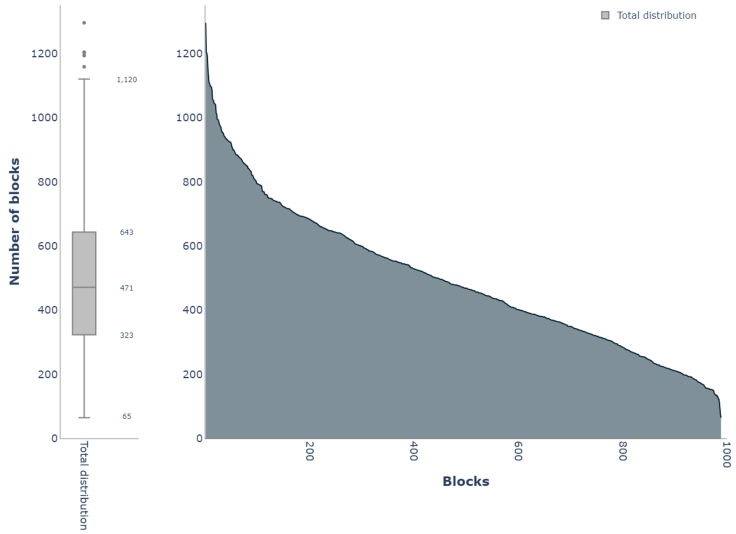


Figure 3.9: Number of unique blocks sent by builders per slot.

3.8 Summary

This chapter addressed the Ethereum component of **RQ1** by quantifying the prevalence of MEV-mediated execution before and after the Merge, using an operational definition grounded in measured private channels, namely Flashbots MEV-Geth mediation in the proof of work era via the MEV-blocks API [54], and relay linkage in the proof of stake era through the post-Merge PBS pipeline [77, 78]. The main quantitative outcome is the time series in Figure 3.3, which shows a transition from a mixed pre-Merge regime, where a substantial fraction of blocks were processed through Flashbots MEV-Geth, to a post-Merge regime where relayed, builder-constructed blocks become the dominant execution mode, exceeding 90% in the most recent windows [78].

The additional analyses strengthen the interpretation and robustness of the prevalence measurement. Figure 3.4 validates that the MEV-mediated label corresponds to a distinct compensation pipeline, and that validator rewards in MEV-mediated blocks are statistically higher than those in no-MEV blocks, consistent with the explicit builder payment convention illustrated in Figure 3.2 and with the incentive-based explanation for the widespread adoption of MEV-Boost [78]. Figures 3.5, 3.6, and 3.7 show that, under PBS, naive reliance on *feeRecipient* obscures validator identities by surfacing builders instead, and that accurate reasoning about validator rewards and concentration requires an additional attribution step using off-chain relay metadata [77, 78]. Finally, Figures 3.8 and 3.9 quantify the intensity of the off-chain market per slot, with medians of 720 total submissions and 471 unique candidate blocks per slot in our 1,000 slot sample, providing an extent signal of MEV-linked activity that complements block-share prevalence within the scope of **RQ1** [78].

Having established a lower bound for MEV prevalence in Ethereum across eras, together with the observable footprints induced by PBS, the next chapter

extends **RQ1** to a contrasting execution architecture. We move from Ethereum's PBS-mediated pipeline to Solana's high-throughput design, where MEV prevalence and visibility are shaped by different runtime constraints and by a different off-chain infrastructure, most notably Jito Labs.

4

MEV on Solana

This chapter extends the empirical analysis from Ethereum to a contrasting high-throughput architecture, Solana. The aim is not to exhaustively characterise all forms of MEV on Solana, but to quantify the share of MEV-linked activity that is directly observable through Solana's off-chain infrastructure, in particular Jito Labs, and to understand how Solana's design mediates extraction.

Solana is selected as the comparative case for three reasons that are specific to its MEV market structure, beyond the widely cited throughput advantages. First, Solana operates a *domain* [86] with a distinct sequencer architecture: the predictable leader schedule (analogous to Ethereum's proposer selection) creates well-defined ordering windows that can be targeted by off-chain infrastructure, even in the absence of a traditional public mempool. This makes Solana a natural laboratory for studying how MEV markets form under different sequencer designs. Second, Solana hosts a large and liquid DeFi ecosystem, decentralized exchanges, lending protocols, and stablecoin markets, that provides the same price-sensitive on-chain state that drives MEV on Ethereum [60]. Third, Jito Labs provides the first publicly observable, systematically documented MEV infrastructure on Solana, enabling a measurement methodology that mirrors the relay-based approach used in Chapter 3, making cross-architecture comparison methodologically consistent [65]. Together, these features make Solana the most appropriate high-throughput case study for extending **RQ1** beyond Ethereum.

Our measurement relies on a limited 500 block snapshot and on publicly observable Jito participation signals. As a result, the reported prevalence reflects only the portion of MEV that leaves an explicit, traceable footprint under this observation model. Nevertheless, this constrained window is sufficient to confirm the same qualitative pattern observed on Ethereum, namely that off-chain MEV infrastructure leaves systematic, measurable footprints in block production.

This chapter supports **RQ1** by enabling a cross architecture comparison with the Ethereum measurements in Chapter 3, and it proceeds in four steps. Section 4.1 briefly recalls the aspects of Solana's design that are most relevant for MEV. Section 4.2 introduces Jito Labs and the Solana MEV stack, with a focus

on the Jito MEV dashboard and the Jito Block Engine. Section 4.3 describes the data collection and matching methodology. Section 4.4 presents a lower-bound estimate of MEV prevalence and a simple centralisation analysis for Jito validators, and Section 4.6 discusses implications and contrasts with Ethereum.

4.1 Architectural Features Relevant to MEV

Solana is a public, proof-of-stake blockchain launched in March 2020 and designed around high throughput and low latency. Its architecture couples a classical PoS validator set with a cryptographic time-stamping mechanism known as *Proof of History* (PoH) [108, 125]. In PoH, a verifiable delay function produces a sequence of hashes that serve as a global clock: each event is recorded together with a hash value in this sequence, which allows nodes to verify the relative ordering of events without repeated coordination.

At a high level, Solana's block production proceeds through a leader schedule. Validators stake SOL and are selected as leaders in proportion to their stake; each leader is responsible for ordering and producing blocks for a sequence of time slots. PoH hashes, together with the leader schedule, allow transactions to be batched and disseminated quickly, reducing the need for repeated global agreement on every intermediate ordering step [108, 125].

These design choices have two important consequences for MEV:

- **High throughput and short timescales.** Solana routinely reaches thousands of transactions per second. MEV opportunities, such as arbitrage or liquidations, arise and disappear quickly against this background. The speed and volume complicate direct, on-chain detection of MEV patterns but also create a rich space for searchers.

- **Predictable leadership.** As in Ethereum proof of stake, the leader schedule is known ahead of time, which allows validators and specialised off-chain infrastructure to plan for upcoming slots. This predictability can be used to organise MEV markets.

While the details differ, Solana shares with Ethereum the basic ingredients that make MEV possible: a public mempool or order flow, price-sensitive on-chain state, and actors with influence over ordering and inclusion. The key difference lies in how MEV is mediated and internalised. On Solana, a central role is played by Jito Labs.

4.2 Jito Labs and the Solana MEV Stack

Jito Labs is a MEV infrastructure provider founded in 2021 that focuses on Solana. Its documentation frames the stack as a way to make MEV extraction more systematic and to redistribute part of the resulting value to validators and their stakers. In the architecture described by Helius, this is achieved by running the bundle auction and selection logic off-chain, while a dedicated relayer

introduces a short delay for ordinary transaction ingress, creating a window for off-chain order flow auctions via the Block Engine [65, 73]. To this end, Jito maintains four main components:

1. A modified Solana validator client (often referred to as *Jito-Solana*) that integrates MEV-aware features and tip handling.
2. A public MEV dashboard that reports statistics about extracted value and participating validators.
3. The Jito Block Engine, an off-chain service that coordinates searchers and validators through an auction for blockspace [72].
4. The Jito Relayer, which acts as a transaction proxy for leaders running Jito. In the architecture described by Helius, the relayer holds incoming transactions for about 200 milliseconds before forwarding them to the leader, providing a short "speed bump" that enables off-chain auctions to decide which bundles are forwarded for inclusion [65].

4.2.1 Jito MEV Dashboard

Jito's MEV dashboard is a public website that aggregates statistics about MEV opportunities and realised MEV on Solana [71]. It reports, among other quantities, estimates of:

- Total MEV extracted over selected time windows.
- Rewards paid to validators participating in Jito-mediated MEV.
- The share of blocks in which Jito validators have been involved in MEV strategies.

From the perspective of this thesis, the dashboard plays two roles. First, it provides an external, independently maintained measurement that confirms the presence and economic significance of MEV on Solana. Second, it exposes downloadable CSV data on validators that have participated in Jito-mediated MEV auctions. These data can be matched against the set of validators that appear in a sample of on-chain blocks, yielding a lower bound on the fraction of blocks with Jito-linked MEV activity.

Figure 4.1 shows a screenshot of the Jito MEV dashboard around October 2023, during the period considered in this analysis.

4.2.2 Jito Block Engine and Order-Flow Auctions

The Jito Block Engine is the counterpart, in the Solana ecosystem, of the builder infrastructure studied in Chapter 3. Rather than relying on independent block builders, Jito provides an engine that coordinates searchers and validators through a dedicated off-chain pipeline. In this design, inclusion and state transition still occur on-chain, but bundle ingestion, simulation, and bid-based selection are

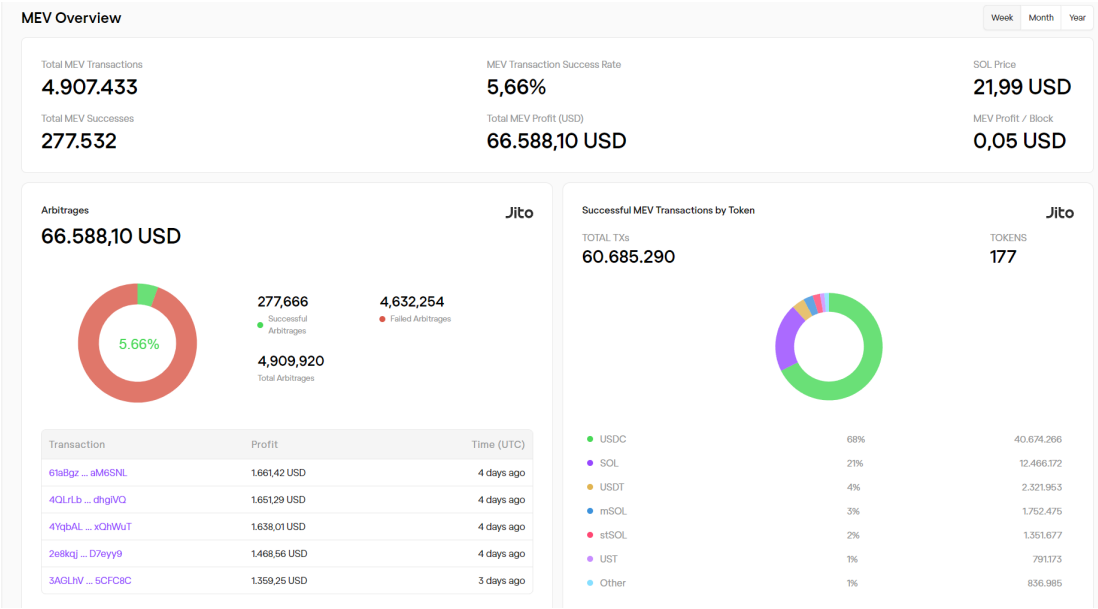


Figure 4.1: Jito Labs MEV dashboard (October 2023). The dashboard reports aggregate statistics about MEV extraction and validator participation on Solana [71].

mediated off-chain by the Block Engine, and, as described by Helius, leaders running Jito route ordinary transaction ingress through the Jito Relayer, which introduces an approximately 200 millisecond delay. This short delay creates a practical window for the Block Engine to run its order flow auction and to forward the selected bundle set to the leader for the upcoming slot [65]. The same off-chain mediation introduces additional trust, availability, and observability dependencies relative to purely protocol native propagation, and it can influence on-chain outcomes by shaping which transactions reach the leader, their ordering within the slot, and the execution guarantees associated with accepted bundles, such as atomic all or nothing execution [73]. Conceptually, the process is as follows [72]:

1. **Searchers** scan Solana's transaction flow for profitable opportunities (for example, arbitrages, liquidations, or other order-dependent strategies). When they identify a candidate strategy, they construct one or more bundles of transactions that implement it.

2. For each bundle, searchers attach a bid: an amount they are willing to pay for the inclusion of the bundle in the next blocks. This bid is analogous to a priority fee or tip, but structured through the Block Engine rather than only via transaction fees.

3. The **Block Engine** receives bundles and bids from many searchers, simulates different combinations and orderings of bundles, and evaluates which combination yields the highest net payment to the validator, subject to correctness and protocol constraints.

4. The Block Engine then forwards to a **validator** running the Jito-Solana client the best-paying bundle set for the slot. The validator includes the corresponding transactions and records the associated payment.

Figure 4.2 summarises these interactions.

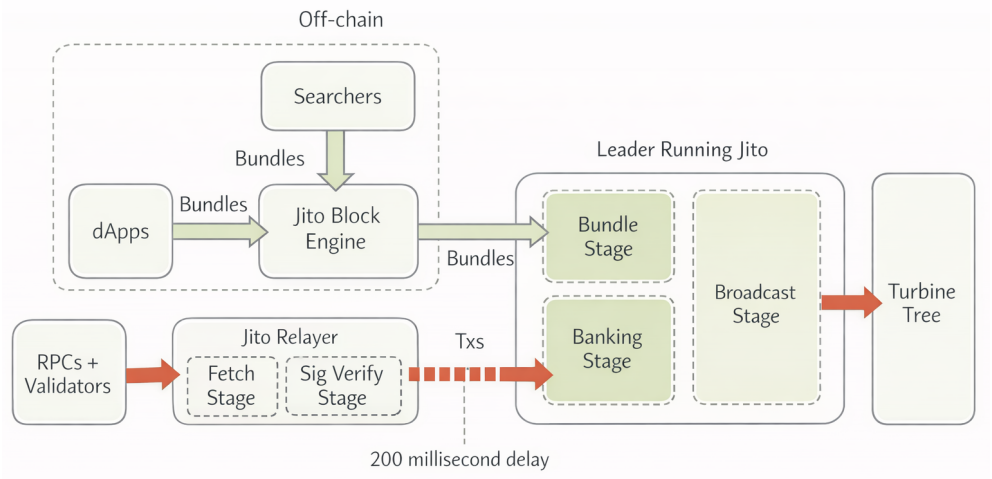


Figure 4.2: Jito mediated MEV pipeline on Solana, adapted from the Helius "Solana MEV Report". Searchers and dApps submit bundles off-chain to the Jito Block Engine, while ordinary transactions are routed through the Jito Relayer. The relayer introduces an approximately 200 millisecond delay before transactions reach a leader running Jito, creating a short window for off-chain bundle selection before broadcast to the Turbine tree [65, 72].

Relative to the Ethereum architecture in Chapter 3, two differences are noteworthy:

- Jito Block Engine collapses the distinction between builders and relays: a single service aggregates bundles, simulates blocks, and forwards the best candidate to validators.
- The interface to validators is provided by a customised validator client rather than a client-side middleware like MEV-Boost.

Despite these differences, the economic logic is similar: searchers pay for ordering, validators sell their slots, and a piece of the extracted value is shared with stakers.

4.3 Data Sources and Methodology

The empirical goal in this chapter is: to obtain a lower bound estimate of the prevalence of Jito-mediated MEV in recent Solana blocks and to examine whether MEV participation appears concentrated among a small subset of Jito validators.

To that end, we combine two data sources:

1. **On chain block data.** For a contiguous window of 500 blocks, we retrieve metadata for each block, including the block number and the identity of the validator that produced it.

2. **Jito validator data.** From the Jito MEV dashboard, we download a CSV file containing the set of validators that have participated in Jito-mediated MEV auctions during the same period.

The on-chain data are collected through the public API of Solscan, a Solana block explorer [110, 111]. At the time of data collection, the Solscan *last* endpoint reported a latest block number of $h_{\max} = 189{,}327{,}223$, corresponding to April 19, 2023. Consequently, we adopt an iterative approach that walks backwards from $h_{\max}$, using a *for* loop to query blocks in reverse order and collect a contiguous window of 500 blocks. Solscan exposes an endpoint that returns the latest block number, and a second endpoint that returns detailed information about an individual block.

The data collection procedure is as follows:

1. Call the Solscan *last* endpoint to obtain the latest block height $h_{\max}$, in our snapshot $h_{\max} = 189{,}327{,}223$.

2. Iterate in reverse order over block numbers $h = h_{\max}, h_{\max} - 1, \ldots, h_{\max} - 499$, and for each h call the Solscan *block* endpoint to retrieve the block's JSON representation.

3. From each JSON object, extract the identity of the validator that produced the block, and store the resulting records in a list.

4. Normalise the list into a tabular format and save the resulting table as a CSV file for further analysis.

The Jito validator data are obtained by exporting from the Jito MEV dashboard the table of validators that have received MEV rewards in a recent period. The exported CSV includes, for each validator, an identity account and associated statistics about MEV rewards. We denote this table by V_{Jito}.

The core matching step then proceeds as follows. Let B be the table of the last 500 blocks, and let V_{Jito} be the set of Jito validators. For each block $b \in B$, we check whether the validator identity listed in b appears in V_{Jito}. If so, we label the block as *Jito-MEV*; otherwise, we label it as *No-MEV* for the purposes of this lower-bound analysis. This procedure can be expressed as:

$$\text{label}(b) = \begin{cases} \text{MEV} & \text{if validator}(b) \in V_{\text{Jito}}, \\ \text{No-MEV} & \text{otherwise.} \end{cases}$$

The resulting labels allow us to compute:

- The share of recent blocks in which a Jito validator is demonstrably involved in MEV.

- The distribution of MEV participation across Jito validators in the sample.

This measurement captures only the MEV that is observable through the chosen public API and Jito participation signals within a 500 block window, and therefore it cannot represent the full MEV surface on Solana, the resulting estimates should be read as partial prevalence figures and as preliminary evidence about centralisation.

4.4 Empirical Results: Lower-Bound MEV Prevalence

Applying the labelling procedure described above to the 500 block window ending at block number 189,327,223 (April 19, 2023) yields the following result:

- Out of the 500 blocks sampled, 85 blocks were produced by validators that appear in the Jito MEV dashboard export.
- The remaining 415 blocks were produced by validators that do not appear in the Jito export for the period.

This implies that at least

$$\frac{85}{500} = 0.17$$

of the blocks, that is **17%**, are certainly associated with Jito-mediated MEV activity. The rest of the blocks may still contain MEV, but through another channel.

Figure 4.3 summarises these proportions as a donut chart.

Several caveats are in order:

- The window of 500 blocks is relatively short. At Solana's block times, this corresponds to a snapshot over a limited horizon rather than a long-term average.
- Other MEV channels, if present, are not captured by this method.

Nevertheless, the 17% figure provides an informative floor: a non-trivial fraction of recent Solana blocks involve explicit MEV coordination through Jito.

4.5 Validator Centralisation in Jito-Mediated Blocks

A second question concerns the distribution of MEV participation across validators using Jito. Given the 500-block sample and the labels described above, we can count, for each Jito validator, how many of the 500 blocks it validated.

Let $C(v)$ denote the number of sampled blocks in which validator v both appears in the Jito export and is the block producer. Plotting the distribution of

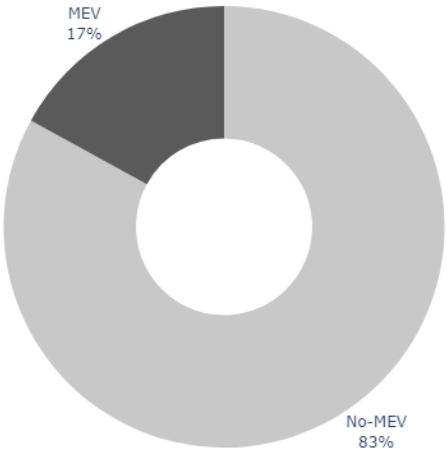


Figure 4.3: Share of MEV and No-MEV blocks among the last 500 Solana blocks, using Jito validator presence as a lower-bound indicator. About 17% of blocks are certainly associated with Jito-mediated MEV.

$C(v)$ across all such validators yields a simple proxy for MEV centralisation in the Jito ecosystem over this window.

Figure 4.4 shows the resulting distribution as a pie chart, where each slice represents the share of MEV-labelled blocks attributed to a given Jito validator.

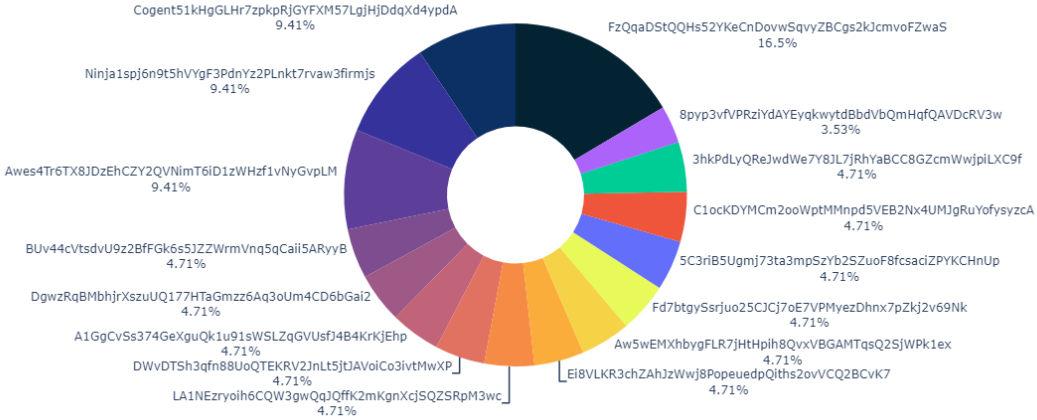


Figure 4.4: Share of MEV-labelled blocks attributed to each Jito validator in the 500-block sample. No single validator dominates MEV participation in this window.

Within the limitations of the sample, two observations emerge:

- The distribution of MEV blocks across Jito validators appears relatively flat: several validators validate comparable numbers of MEV-labelled blocks, and no single entity accounts for a majority.

- Given the small sample size and the short time horizon, we cannot draw strong conclusions about long-run centralisation. The absence of a dominant validator in this snapshot is compatible with a genuinely diffuse MEV ecosystem and with slower dynamics in which leadership would emerge only over longer windows.

4.6 Discussion and Contrast with Ethereum

Before comparing prevalence figures across architectures, a methodological caveat is necessary. The Ethereum post-Merge dataset in Chapter 3 covers September 2022 through September 2023, while the Solana window analyzed here is a 500-block snapshot from April 2023. The two datasets are therefore not fully contemporaneous. In particular, the Ethereum window includes the early post-Merge period when MEV-Boost adoption was still ramping up [64], which mechanically lowers the aggregate Ethereum prevalence figure relative to a window restricted to mature PBS operation. Direct percentage comparisons between the 85.15% Ethereum figure and the 17% Solana lower bound should therefore be treated as indicative cross-architecture patterns rather than as precise simultaneous benchmarks.

This chapter answers **RQ1** in a cross-architecture sense by extending the prevalence analysis beyond Ethereum, using Solana as a contrasting high--throughput case. Using a Jito-based label on a 500 block window, we obtain a lower bound estimate of 17% Jito-mediated MEV blocks, and we observe no single Jito validator dominating within that snapshot. Compared to Chapter 3, where relay-mediated, MEV Boost style infrastructure dominates post-Merge block construction over multi-million block windows, the Solana estimate is far from saturation and is tied to a single provider, highlighting both architectural differences and measurement conservatism in the Solana setting.

Relative to Chapter 3, the contrast is twofold. First, Ethereum exhibits near saturation of off-chain block building in the post-Merge period, with MEV Boost-mediated construction dominating multi-million block windows, whereas the Solana estimate here captures only the share of blocks linked to one provider and remains far from saturation. Second, the architectural locus of coordination differs. Ethereum relies on a modular PBS pipeline where independent builders compete behind relays and MEV Boost, while Solana, in the configuration studied here, concentrates order flow coordination in the Jito Block Engine and a modified validator client. Despite these differences, the qualitative logic aligns across systems: extracting MEV at scale requires off-chain coordination and computation that shapes on-chain ordering outcomes, even though it is not directly visible in base layer consensus metrics.

Relying on Jito as the sole observation channel introduces measurement constraints that should be interpreted carefully when comparing results to Ethereum.

Jito is, as of the period studied, the dominant MEV infrastructure provider on Solana, but it does not capture all MEV activity on the chain. Validators not running the Jito-Solana client, or searchers routing bundles through alternative channels, are invisible under this approach. In principle, alternative detection methods are available: direct on-chain analysis of swap and liquidation patterns could label MEV strategies without relying on Jito's off-chain infrastructure [60]. However, Solana's lack of a persistent public mempool means that front-running and sandwich attacks cannot be detected by mempool comparison, the primary approach used for Ethereum in Chapter 3, and post-execution heuristics on Solana are less established in the literature. Under these constraints, the 17% Jito-mediated lower bound should be interpreted as a realized extraction floor for this specific provider and window, analogous to measuring only Flashbots MEV-Geth blocks in pre-Merge Ethereum [78]. The true share of MEV-influenced blocks on Solana is likely higher, but quantifying the gap would require either a Solana-specific behavioral classifier or broader access to off-chain bundle metadata beyond what is publicly available.

5

Ordering and Democratic Principles

Chapter 5 addresses **RQ2 (Ordering and favoritism)**, namely whether realized intra block ordering systematically favors MEV related activity, and whether private channel submissions and direct payment to the block builder flows[1] obtain priority beyond what fee incentives would imply. The chapter is empirical and normative at once. Empirically, it operationalizes a transparent baseline that we term *fee-based meritocracy*, then tests it against observed ordering at scale in a post-Merge dataset. Normatively, it evaluates systematic deviations from this baseline through the lens of a three principle framework, the *democratic triad*, no collusion, meritocracy, legitimate MEV, and asks whether PBS and private order flow weaken credible neutrality for ordinary users.

5.1 The democratic triad

We evaluate ordering practices through a three principle framework that we refer to as the *democratic triad*, composed of *no collusion*, *meritocracy*, and *legitimate MEV*.

Principle 1, No collusion. Blockchain actors may pursue MEV strategies, provided that they operate independently from each other. The principle is violated when block producers and third parties engage in exclusive arrangements that grant privileged inclusion or ordering to specific transactions.

Principle 2, Meritocracy. Block producers should behave as economically rational actors, selecting transactions based on transparent incentives such as fees, rather than on hidden side deals. A prototypical violation is favoritism granted to traders who provide direct, off chain, or otherwise concealed payments in exchange for inclusion and ordering guarantees.

[1] In this thesis, "direct payment to the block builder flows" denotes direct payments from a searcher to a builder in exchange for inclusion and ordering guarantees, implemented either as a dedicated payment transaction or as a cryptocurrency transfer executed within a smart contract.

Principle 3, Legitimate MEV. Actors shall pursue only legitimate, that is non malicious, MEV strategies. Under this principle, practices that systematically extract value by harming other users, such as sandwiching victims, are considered illegitimate.

Each dimension of the democratic triad is operationalized through a specific metric, and each metric carries inherent false positive (FP) and false negative (FN) risks that bound what the empirical results can and cannot claim.

Meritocracy is measured via *Tx_Shift*, the distance between realized transaction position and the fee-based ideal ranking. FP: transactions legitimately reordered for technical reasons may exhibit negative Tx_Shift that resembles favoritism without reflecting discretionary off-chain deals. FN: favoritism exercised at the bundle level, where builders optimize joint bundle profitability rather than individual transaction position, is not captured by per-transaction Tx_Shift [64].

Legitimate MEV is measured as the frequency of MEV transactions not labeled as front-running or sandwich attacks by Zeromev [128]. FP: benign strategies may be incorrectly labeled as harmful by rule-based heuristics, leading to an underestimation of legitimate MEV. FN: harmful activity routed exclusively through private bundles may evade detection if Zeromev does not flag its internal execution pattern, causing malicious MEV to be miscounted as legitimate in the private channel.

No Collusion is measured via graph modularity and cluster stability (Chapter 6). FP: high modularity indicates organized structure, but it does not prove explicit collusion; structural cohesion can arise from shared infrastructure or rational routing toward high-quality builders [119]. FN: implicit collusion, for example coordinated fee schedules not reflected in on-chain payments, is not detectable through graph structure alone.

Table 5.1 summarizes how each principle of the democratic triad is operationalized in this chapter, together with the primary measurement limitations.

In this chapter, we label an ordering behaviour as *undemocratic* when there is empirical evidence of persistent, mechanism driven violations of at least one principle of the democratic triad. In operational terms, we treat systematic priority beyond fee incentives, consistent priority for privately routed order flow, or consistent priority associated with direct builder payment flows, as evidence against meritocracy and, depending on the mechanism, as suggestive of collusion.

5.2 Ethereum gas, fees, and why they matter for intra-block ordering

Any empirical notion of "fee-based meritocracy" requires a precise account of what it means, mechanically and economically, for a transaction to "pay more" in Ethereum. Transaction ordering is the interface between users' fee signals

Table 5.1: Operationalization of the democratic triad: metrics, data sources, and limitations.

Principle	Metric	Data source	Main limitations
Meritocracy	Tx_Shift: distance from fee-based ideal rank	XBlock ETH [129], Etherscan [42]	Bundle-level optimization not captured; RBF and gas-packing may produce spurious shifts
Legitimate MEV	Frequency of malicious attack labels (frontrun, backrun, sandwich)	Zeromev API [128]	Private-bundle attacks may evade detection (FN); aggressive-but-benign strategies may be over-labeled (FP) [22]
No Collusion	Graph modularity (Leiden) and temporal cluster stability	MEV interaction graph (Chapter 6)	High modularity ≠ collusion; implicit coordination not detectable on-chain [119]

and producers' incentives, and it is therefore the natural place where preferential treatment, if present, becomes measurable.

Ethereum charges *gas* as an abstract unit for computation and state access. Each transaction specifies a *gas limit*, namely the maximum amount of gas units the sender is willing to purchase, and execution consumes a realized amount of gas that depends on the executed operations, state reads, and writes. Users pay fees in ETH, typically quoted in gwei (where $1,\text{gwei} = 10^{-9},\text{ETH}$), and the total amount paid by a transaction is proportional to gas used times an effective price per unit of gas. Gas also serves a security role, because requiring payment for computation limits spam and makes denial of service attempts costly. [38]

Since the London upgrade, Ethereum uses the EIP 1559 fee market. Each block includes a protocol-determined *base fee* per gas, denoted *baseFeePerGas*, which adjusts with congestion, and is *burned* rather than paid to the block producer. In addition, users can specify a *priority fee* per gas, expressed through *maxPriorityFeePerGas*, which acts as a tip to the producer, and a total cap *maxFeePerGas* that bounds what they are willing to pay per gas overall [36, 40].

The split between base fee and tip is critical for ordering incentives. The user pays, per unit of gas, the sum of the base fee and an effective priority fee, while only the priority component accrues to the producer. In the EIP 1559 transaction format, the effective priority fee per gas is bounded by the gap between the user cap and the block base fee, namely

$$priorityFeePerGas = \min\Big(maxPriorityFeePerGas,\, maxFeePerGas - baseFeePerGas\Big)$$

As a result, the total fee paid by the sender can be written as

$$feePaid \;=\; gasUsed \cdot \big(baseFeePerGas + priorityFeePerGas\big),$$

while the component that constitutes protocol native producer revenue is

$$tipRevenue \;=\; gasUsed \cdot priorityFeePerGas$$

This distinction motivates why a "fee-based" ordering rule, in a meritocratic sense, is naturally expressed in terms of a gas price signal tied to the tip, rather than in terms of the base fee, which is constant within the block and not received by the producer [36, 38].

From the producer's perspective, ordering is a constrained optimization under the block gas limit. When choosing among candidate transactions, a rational producer compares their marginal revenue per gas, so ranking by a per-gas tip signal is coherent with revenue maximization, whereas comparing unnormalized totals would mechanically favor gas-intensive transactions even when they deliver lower revenue per unit of gas. This is the operative sense in which "fee-based meritocracy" is defined in this chapter, it is a transparent counterfactual in which higher effective tip signals, per unit of gas, receive earlier placement.

Post-Merge PBS further expands the space of ordering incentives. The validator proposes the block, but a builder may construct it, select and order transactions, and specify the execution payload, including the *feeRecipient* that captures tip revenue. Within this architecture, there are two conceptually distinct channels that can buy priority. The first is protocol native, users compete by increasing their effective priority fee. The second is a side payment channel, searchers can directly remunerate builders in exchange for inclusion or ordering guarantees, for example, via explicit payment transactions, or via transfers embedded inside contract calls. These direct builder payment flows are central to RQ2, because they can rationally generate ordering advantages that are not explained by gas price signals alone. [36, 40]

A fee-based baseline remains informative even though builders often optimize at the bundle level. If a builder includes bundles whose profitability comes primarily from MEV extraction rather than from tips, then realized ordering will not, in general, coincide with a pure tip maximization rule. The role of the baseline in this chapter is therefore not to claim a unique ground truth ordering, but to provide a transparent neutrality benchmark. Systematic deviations that repeatedly advantage specific categories, in particular private submissions and direct builder payment flows, become evidence that off-chain MEV mechanisms reshape on chain ordering outcomes in ways that ordinary fee competition does not fully explain [38, 80].

5.3 Empirical Setting and Data Construction

We analyze all Ethereum blocks produced in January 2024, a post-Merge window comprising 220,993 blocks, from block 18,908,895 to block 19,129,888, and 36,015,340 transactions therein. We obtain block and transaction level data from the *XBlock ETH* datasets [129], and we enrich them with block header and receipt level information via the *Etherscan* API [42] and auxiliary metadata queries via the *MetaMask* developer API [87].

The choice of January 2024 as the analysis window is deliberate. By this point, MEV-Boost adoption had reached near-saturation [78], the PBS supply chain was structurally stable, and no major protocol upgrade had yet altered the

fee and ordering mechanics: the Dencun upgrade (EIP-4844) landed in March 2024, after the analyzed window. January 2024 therefore represents a mature, steady-state configuration of post-Merge Ethereum, free from the transitional dynamics of the weeks immediately following the Merge and from the structural discontinuities that Dencun later introduced for blob-carrying transactions [34]. Restricting the ordering analysis to this single month also allows fine-grained, block-by-block matching with the Mempool Dumpster dataset [55] and with relay delivery records, both of which are queried at daily granularity and whose coverage can degrade for longer retrospective windows.

The core dataset links each transaction to a set of attributes that allow us to test fee-based meritocracy and to isolate the potential mechanisms behind deviations.

Public versus private transaction visibility. We classify a transaction as *public* if it is observable in public mempool traces before inclusion, and as *private* otherwise. To implement this, we use *Mempool Dumpster*, a Flashbots research dataset that collects transaction hashes observed in the public mempool [55]. Transactions in our January 2024 block sample whose hashes match the mempool traces are labeled as *public*. All remaining transactions are labeled as *private*, under the interpretation that they likely reached the block producer through an off-chain channel, for example private RPC endpoints, relay-mediated order flow, or direct bundle submission in the MEV ecosystem.

MEV blocks versus NoMEV blocks. To connect ordering outcomes to PBS-mediated construction, we partition blocks into *MEV blocks* and *NoMEV blocks*. A block is labeled as *MEV* if its block hash matches a block hash reported by at least one relay API in the MEV Boost ecosystem. We collect relay reported blocks from the following relays, Aestus [2], Agnostic Gnosis [100], BloXroute MaxProfit [13], BloXroute Regulated [14], Eden Network [90], Flashbots [50], Manifold [81], Titan [116], Ultrasound [121]. In this January 2024 dataset, 90.6% of blocks can be traced to these relays, which provides a large, real-world sample of PBS-mediated construction.

Direct payment to the Builder detection and economic value. A central mechanism in RQ2 is explicit side payments that reward the builder for preferential inclusion or ordering. We therefore label as *builder payment* those transactions that include a payment to the builder of the block that contains them. Detecting these payments requires receipt-level inspection because builder compensation can be implemented either as a simple transfer, or as a smart contract call that triggers one or more internal transfers. We identify smart contract calls using Etherscan metadata, and, for those transactions, we inspect the transaction receipts to recover the set of payments emitted by the execution trace. If a paid recipient matches the block builder, we classify the transaction as *builder payment*. For the economic magnitude of each transaction, we compute a USD value by gathering token prices at inclusion time, then multiplying transferred token amounts by their corresponding USD price. For smart contract calls, we

conservatively assign the transaction's value as the maximum payment observed in the receipt trace, which serves as a proxy for the monetary importance of the action and avoids undercounting multi-transfer executions [42].

MEV attack labels. To relate preferential ordering to profitable and malicious behavior, we attach per-transaction MEV attack labels using the Zeromev API [128]. The labels include *arbitrage*, *liquidation*, *swap*, *frontrun*, *backrun*, *sandwich*, and *none*. In this labeling scheme, *sandwich* denotes victim transactions of sandwich attacks, while the attacker legs are typically labeled *frontrun* and *backrun*. These labels allow us to test whether the transaction groups that are prioritized under PBS also concentrate malicious activity.

Labeling tool: rationale and alternatives. The choice of Zeromev [128] as the labeling source is motivated by its public availability, its transaction-level granularity, and its direct support for the post-Merge Ethereum period covered by this chapter. Two main alternatives exist in the literature. The first is *MEV-Inspect* [56], a Flashbots open-source pipeline that classifies MEV strategies by replaying transaction traces. Zeromev builds on this line of work and incorporates MEV-Inspect within a broader labeling framework, extending it with additional detection functionality and wider transaction-level coverage beyond Flashbots-mediated bundles. The second is *EigenPhi* [32], a commercial labeling service with broader heuristic coverage, but whose detection methodology is partially proprietary and therefore harder to audit. Zeromev applies a rule-based classifier that covers sandwich attacks, arbitrage, and liquidations for any Ethereum block, independent of the submission channel. Its use is therefore particularly appropriate in this chapter, since it combines the open and reproducible foundations of MEV-Inspect with a broader labeling scope more suited to post-Merge Ethereum. Its main known limitation is that transactions routed exclusively through private bundles may be underdetected if their economic footprint is not visible in finalized state differences [22]. This implies that the malicious MEV concentration reported in Section 5.6 for MEV private transactions is likely a lower bound: the true rate of malicious activity in private flows could be higher if private-bundle attacks systematically evade Zeromev's classifier.

5.4 Fee-Based Meritocracy as a Baseline

A simple, transparent baseline that approximates "credible neutrality" is *fee-based meritocracy*, namely the idea that, given a set of candidate transactions, inclusion and ordering should be broadly aligned with the fee signals that users submit. This baseline is not meant to claim that fee maximization is socially optimal, especially under EIP 1559 dynamics. Instead, it functions as an interpretable counterfactual rule. If realized that ordering repeatedly deviates from this rule in a structured way that advantages specific off-chain flows, then

the deviation becomes evidence of discretionary ordering power and potential collusion.

Operationally, we compute, for each transaction, two indices within its block.

- *Actual_Tx_Index*, the realized position of the transaction in the block, starting at one for the first transaction and incrementing sequentially.
- *Ideal_Tx_Index*, the counterfactual position the transaction would have if all transactions in the block were sorted in descending order of their gas price signal. In this construction, the highest paying transactions appear first, and ties follow the induced stable ordering from the sorting procedure.

We then define the ordering deviation metric

$$Tx\_Shift = Actual\_Tx\_Index - Ideal\_Tx\_Index.$$

A negative *Tx_Shift* implies that the transaction is placed *earlier* than the fee-based baseline would predict, which we interpret as *prioritization*. A positive *Tx_Shift* implies that the transaction is placed *later* than the baseline, which we interpret as *delay*. Because *Tx_Shift* is defined transaction by transaction within each block, it supports comparisons across blocks of different sizes and allows us to aggregate ordering outcomes over large windows.

To connect *Tx_Shift* to the off-chain MEV ecosystem, we partition transactions into four mutually exclusive groups:

1. **NoMEV public**, the mempool transactions included into blocks that did not originate from MEV relays: 2,872,835 transactions (8% of the dataset).
2. **MEV public**, the mempool transactions included into blocks from MEV relays: 29,406,322 transactions (81.6% of the dataset).
3. **MEV private**, transactions. Private transactions that do not involve payment to the block builder: 3,460,670 transactions (9.6% of the dataset).
4. **Builder payment**. Private transactions that involve payment to the builder: *275,513 transactions* (0.8% of the dataset).

This grouping isolates two mechanisms emphasized by RQ2, private channel submission and direct side payments, and it allows us to compare them against a natural "control" group, NoMEV public blocks, where off-chain mediation is not present.

5.5 Preferential Ordering Beyond Fees

This section provides the core empirical answer to **RQ2**. Using the fee-based meritocracy baseline defined in Section 5.2, we quantify how far realized intra-block ordering departs from what fee signals alone would predict. We first compare the ordering deviation distributions in MEV versus NoMEV blocks, then we isolate which transaction categories drive the deviations, with a specific focus on private submissions and direct payments to the builder.

5.5.1 Deviation patterns

Figure 5.1 reports the distribution of *Tx_Shift* for the four transaction groups. The most striking reference point is the NoMEV public group. Across the transactions in NoMEV blocks, the distribution is essentially degenerate at zero, excluding outliers, meaning that these blocks closely follow the fee-based ordering rule. In other words, when a block is not traced to the relay-based PBS ecosystem, the realized ordering is well approximated by fee-based meritocracy.

The picture changes sharply inside MEV blocks. For MEV private transactions, the distribution of *Tx_Shift* is skewed toward negative values. In the January 2024 dataset, MEV private comprises about 3.4 million transactions, and its first quartile is −28. This indicates that a large fraction of private transactions are systematically advanced relative to their fee-based position.

By contrast, MEV public transactions, which comprise about 29.4 million transactions in the same window, show the opposite tendency. Their median *Tx_Shift* is positive, equal to 3, indicating that many public mempool transactions are delayed in MEV blocks. The joint pattern, private transactions advanced and public transactions delayed, is consistent with a reordering mechanism that prioritizes private order flow at the expense of public order flow within PBS-mediated construction.

The strongest deviation appears for builder payment transactions. Their *Tx_Shift* values are strongly negative, with a median centered at −139. Under the fee-based meritocracy baseline, such a shift is difficult to attribute to ordinary fee competition. Instead, it aligns with the interpretation that explicit side payments purchase preferential ordering.

5.5.2 Actual transaction index confirms systematic prioritization

A complementary perspective is obtained by looking at realized transaction positions directly, rather than relative shifts. Figure 5.2 plots the distributions of *Actual_Tx_Index* for the same groups, together with the population-wide distribution. The results reinforce the interpretation of preferential treatment.

Builder payment transactions and MEV private transactions are consistently placed close to the beginning of blocks. This is coherent with the negative *Tx_Shift* results, and it clarifies that prioritization is not merely a statistical artifact of the baseline. It manifests as early block placement.

Taken together, the *Tx_Shift* and *Actual_Tx_Index* analyses support a clear empirical conclusion for RQ2, observed intra-block ordering in PBS traced blocks systematically favors private order flow and builder payment flows beyond what fee-based meritocracy would imply, while non-relay traced blocks behave as an approximate neutrality control group.

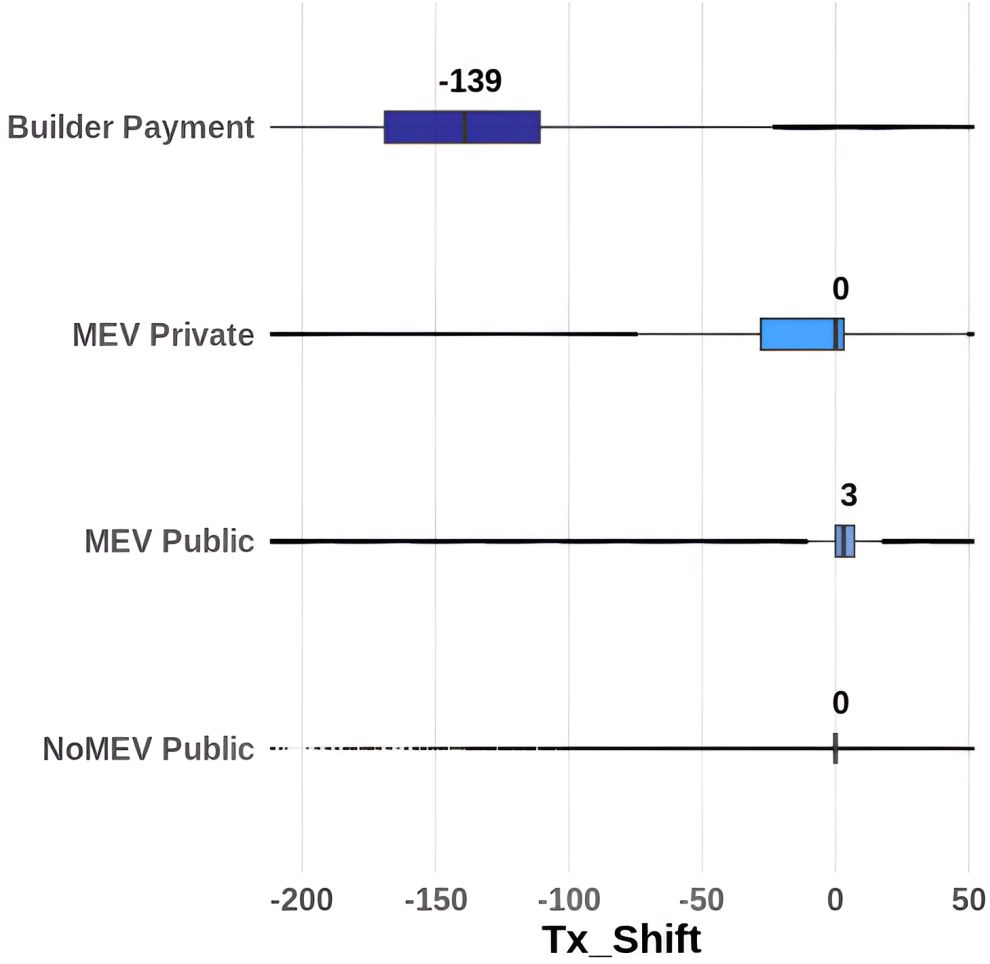


Figure 5.1: Ordering deviation *Tx_Shift* by transaction group, computed as *Actual_Tx_Index* minus *Ideal_Tx_Index* under a fee-based ordering baseline. Negative values imply earlier than baseline placement, positive values imply delay. NoMEV blocks remain close to the neutral baseline, whereas MEV blocks exhibit systematic reordering that prioritizes MEV private and builder payment transactions.

5.5.3 Signals consistent with favoritism, builder payments as bought priority

The most direct ordering signal consistent with collusion is the behavior of builder payment transactions. A large negative *Tx_Shift* means that transactions tied to explicit transfers to the builder appear much earlier than their fee-based position predicts. This conflicts with the meritocracy component of the democratic triad. Ordering is not simply a transparent function of fee signals, rather it is consistent with an additional compensation channel that buys priority, and, in practice, can buy ordering guarantees for time-sensitive strategies.

Importantly, this interpretation does not require claims about intent that cannot be directly verified on-chain. The claim is empirical and statistical: the

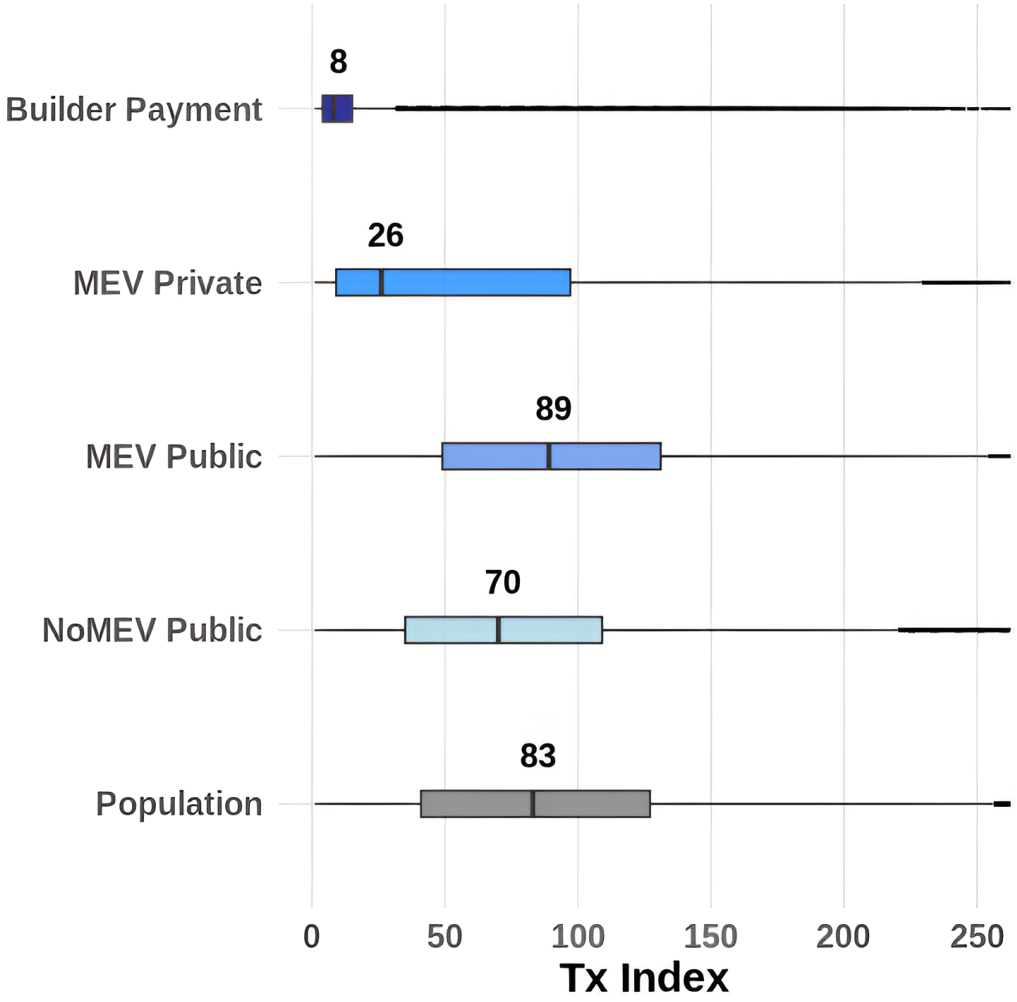


Figure 5.2: Distributions of actual transaction index for builder payment, MEV private, MEV public, and NoMEV public transactions. Preferential groups, builder payment, and MEV private, are disproportionately concentrated near the beginning of blocks.

repeated magnitude and directionality of ordering advantages for builder payment flows are difficult to reconcile with neutral, fee-ordered construction. In the language of **RQ2**, these patterns are best read as signals of *favoritism* toward searchers who transact through direct builder payments and private channels, and, via PBS, toward the broader off-chain pipeline that validators accept when they propose builder-constructed payloads.

5.6 Favoritism and Malicious MEV Roles

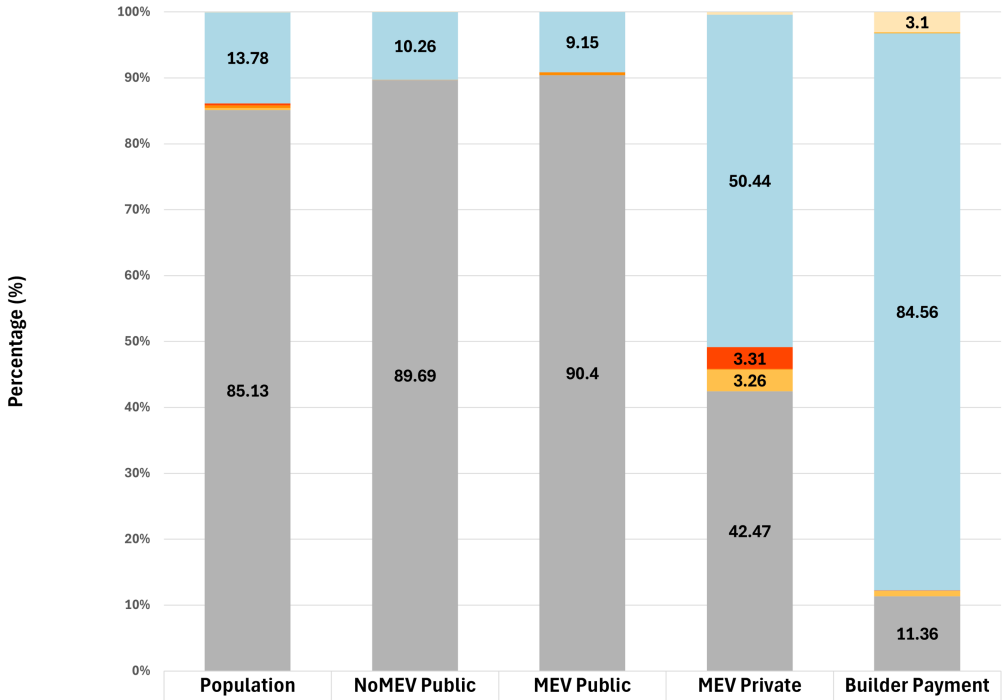
RQ2 also asks whether prioritized categories are linked to speculative or malicious behavior, and, more broadly, whether favoritism undermines the "legitimate MEV" component of the democratic triad. We address this by combining the ordering groups with Zeromev role labels [128].

Figure 5.3 reports the composition of MEV roles across the full population

and across transaction groups, and Figure 5.4 provides the orthogonal view, for each MEV role, the distribution of transaction groups that realize it.

Two patterns are particularly relevant. First, malicious MEV activity, proxied by *frontrun* and *backrun* transactions, concentrates overwhelmingly in MEV private transactions. This is consistent with strategic incentives, malicious actors benefit from hiding order flow from the public mempool to reduce detection and competition. Second, highly remunerative strategies, notably *arbitrage* and *liquidation*, have substantial mass in the builder payment category. This is consistent with a rational demand for ordering guarantees, profitable strategies may choose to pay side transfers to ensure timely inclusion and preferred ordering.

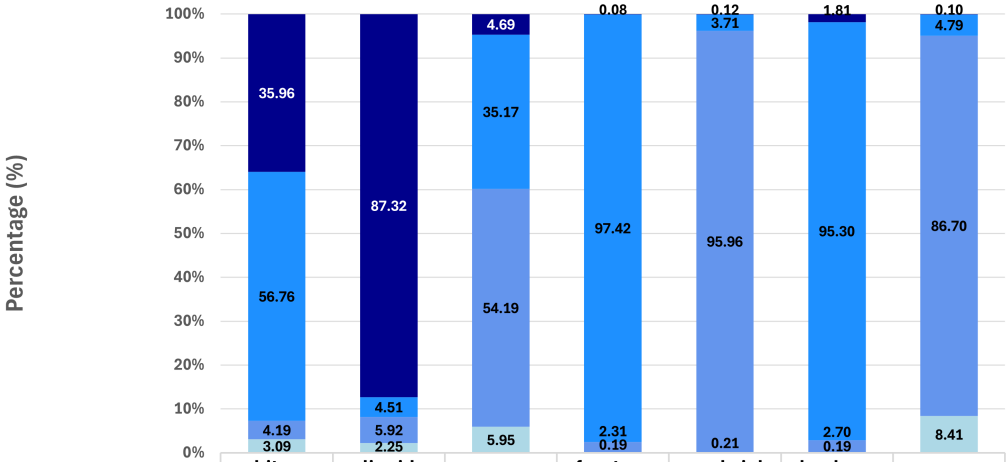
Finally, the labeling also clarifies the victim's side of the harmful extraction. In the dataset, victims labeled *sandwich* are public transactions in 96% of cases, which is consistent with the view that public mempool observability exposes ordinary users to adverse ordering.



	Population	NoMEV Public	MEV Public	MEV Private	Builder Payment
arbitrage (%)	0.06	0.03	<0.01	0.39	3.1
liquidation (%)	<0.01	<0.01	<0.01	<0.01	0.11
swap (%)	13.78	10.26	9.15	50.44	84.56
frontrun (%)	0.33	0.01	0.01	3.31	0.03
sandwich (%)	0.37	0.01	0.43	0.14	0.06
backrun (%)	0.33	0.01	0.01	3.26	0.78
none (%)	85.13	89.68	90.4	42.46	11.36

Figure 5.3: MEV role composition across the full dataset and across transaction groups. Malicious roles, frontrun and backrun, concentrate on MEV private flows, while a substantial portion of profitable roles appear in builder payment transactions.

These results tighten the normative interpretation. Preferential ordering is



	arbitrage	liquid	swap	frontrun	sandwich	backrun	none
Builder Payment (%)	35.96	87.32	4.69	0.08	0.12	1.81	0.10
MEV Private (%)	56.76	4.51	35.17	97.42	3.71	95.30	4.79
MEV Public (%)	4.19	5.92	54.19	2.31	95.96	2.70	86.70
NoMEV Public (%)	3.09	2.25	5.95	0.19	0.21	0.19	8.41

Figure 5.4: For each MEV role, the share of transactions belonging to builder payment, MEV private, MEV public, and NoMEV public groups. Sandwich victims are overwhelmingly public, while frontrun and backrun legs are predominantly private.

not isolated to benign flows, rather, the transaction categories that are prioritized under PBS, especially private order flow, are also the categories where malicious behavior concentrates. This joint pattern directly supports a negative answer to the "legitimate MEV" ideal, the same off-chain mediation that enables efficient extraction also creates a protected channel for harmful strategies.

5.7 Economic Magnitude and Recipient Structure

Preferential treatment is most concerning when it applies to economically meaningful flows. We therefore analyze the distribution of transaction values in USD across transaction groups. Figure 5.5 shows the results as boxplots on a logarithmic scale, after filtering out negligible values, specifically transactions with value less than 0.1 USD.

The builder payment group exhibits the highest economic values, with the third quartile exceeding 10,000 USD in the January 2024 sample. More broadly, MEV private transactions have significantly higher values than public transactions, both MEV public and NoMEV public. This supports the interpretation that private order flow and bought priority are central mechanisms in contemporary PBS-mediated block building, not marginal edge cases.

We also investigate the type of transaction recipients, to better understand the infrastructure that these prioritized flows interact with. Using Etherscan labels, we classify recipients as externally owned accounts, generic smart contracts, and, when available, specialized categories such as MEV Bot contracts and DEX contracts [42]. Figure 5.6 summarizes the distribution of recipient

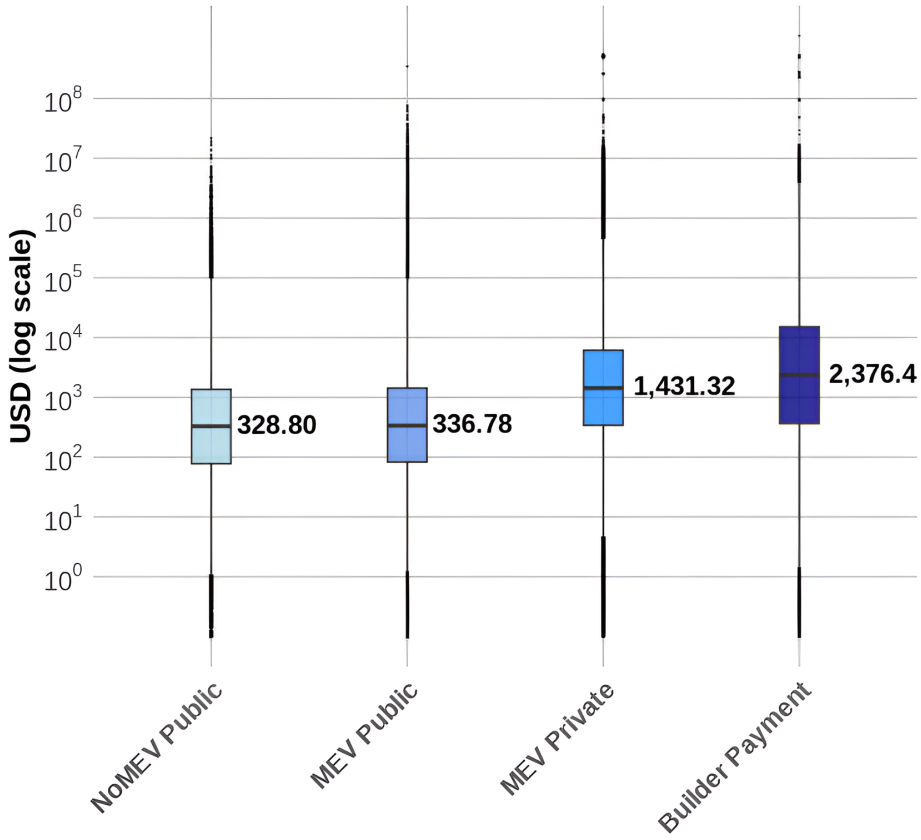


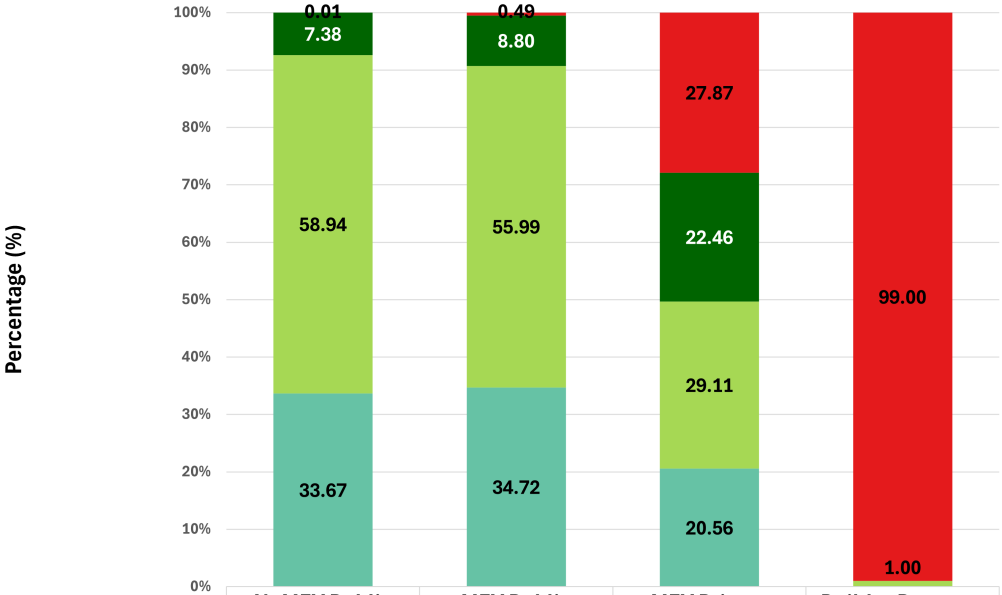
Figure 5.5: Distribution of transaction values in USD, log scale, across transaction groups, after filtering out transactions with a value below 0.1 USD. Builder payment transactions dominate the upper tail, and MEV private exceeds public groups in magnitude.

types across transaction groups.

Two observations stand out. First, the vast majority of builder payment transactions interact with MEV Bot, and MEV Bots also represent a significant share of recipients for MEV private transactions. This is consistent with the view that many of the prioritized and favoritism indicating flows are associated with automated MEV execution. Second, DEX contracts appear in both public and private transactions, but they are disproportionately called by private transactions. This is consistent with the strategic logic of private order flow; traders may route transactions privately to protect profitable strategies from public mempool competition and from becoming sandwich victims.

5.8 Discussion and Limitations

The empirical results provide a concrete answer to **RQ2**. Observed intra-block ordering in relay traced MEV blocks systematically favors MEV private and



	NoMEV Public	MEV Public	MEV Private	Builder Payment
MEV Bot (%)	0.01	0.49	27.87	99.00
DEX (%)	7.38	8.80	22.46	0.00
Smart Contract (%)	58.94	55.99	29.11	1.00
EOA (%)	33.67	34.72	20.56	0.00

Figure 5.6: Distribution of transaction recipient types across transaction groups, using Etherscan recipient labels. Builder payment transactions overwhelmingly interact with MEV bot contracts, and DEX contract interaction is substantially higher in MEV private flows than in public groups.

builder payment flows beyond what fee-based meritocracy would imply. NoMEV blocks behave as an informative contrast class, because their realized ordering is closely aligned with the fee-based baseline, suggesting that the strongest deviations are tied to PBS-mediated construction rather than being a universal property of block building.

From the perspective of the democratic triad, the findings stress meritocracy and legitimate MEV. They conflict with meritocracy because ordering is not explainable as a transparent function of fee signals when private order flow and builder payment flows are repeatedly advanced. They also raise concerns about legitimate MEV because the flows that are prioritized, especially MEV private, are the locus where malicious roles, frontrun and backrun, concentrate, while public transactions disproportionately constitute sandwich victims. In this sense, the evidence suggests not only that ordering is reshaped by off-chain mechanisms, but also that the reshaping interacts with harmful extraction rather than being confined to benign efficiency-improving strategies.

At a higher level, the chapter highlights a post-Merge tension. PBS and relays can be viewed as an efficiency innovation that makes block markets explicit and specialized. Yet the same infrastructure creates a parallel economy of private order flow and direct builder payments that weakens the credible neutrality narrative for ordinary users. Under this view, the ordering layer, not only

the payment layer, becomes a key site where off-chain power is exercised and where favoritism can be operationally detected.

Several limitations qualify the interpretation and clarify what the results do and do not claim.

Mempool dataset completeness. The public versus private classification depends on observability in public mempool traces, and any mempool dataset can be incomplete, which can inflate the private category. Mempool Dumpster [55] is the primary source for public-mempool visibility in this chapter. It is a continuously maintained Flashbots research dataset that monitors multiple Ethereum nodes, but it cannot guarantee exhaustive coverage of all transactions that briefly enter and leave the public gossip layer. Alternative datasets exist, for example Blocknative's historical mempool archive [10] and EigenPhi's labeled transaction stream [32], but they cover different node vantage points and time windows, and a full cross-source reconciliation is outside the scope of this chapter. Critically, the *builder payment* group is identified via receipt-level inspection of on-chain payment flows rather than via mempool absence, and is therefore unaffected by this limitation. Since the strongest ordering deviations are documented for builder payment transactions (median Tx_Shift = −139), the core conclusions remain robust even under conservative re-interpretation of the private/public boundary [69].

Finally, attribution and intent remain inherently limited in on-chain inference. Preferential ordering patterns, including strong advancement of builder payment flows, are consistent with discretionary favoritism and with side payment mediated arrangements, but they do not prove explicit agreements among searchers and builders. The chapter, therefore, treats favoritism as an empirical ordering signature, supported by repeated statistical patterns, rather than as a claim about intent. Extending the analysis across additional months and stress periods would further test stability and support time-varying auditing of neutrality.

Overall, the evidence supports the thesis' central normative claim for **RQ2**. PBS and private order flow reshape not only who is paid, but also how ordering power is exercised, with systematic and economically meaningful prioritization for private and builder payment transactions in MEV blocks. This sets up the network-based analysis in the next chapter, where preferential relationships are represented as edges and studied at the level of ecosystem structure.

6

Network Structure and Communities in the MEV Ecosystem

This chapter answers **RQ3 (Social structure and centralization)**: *which stable groups and interest clusters form within the MEV ecosystem, how PBS mediated relationships differ from direct transfers, and what degree of centralization emerges among extracting actors*. The previous chapters showed, first, that post-Merge Ethereum largely relies on off-chain block construction under PBS, second, that realized ordering often deviates from a fee-based baseline when private order flow and builder payment flows are involved. Here, we adopt a complementary lens, we represent realized MEV as a graph of repeated economic interactions among labeled actors, and we test whether these interactions resemble random mixing, or instead concentrate into persistent communities and structurally central hubs.

Within the framework of the democratic triad introduced in Chapter 5, this chapter operationalizes the *No Collusion* principle. High modularity and persistent clusters in the MEV interaction graph are the empirical signature against which we evaluate whether the ordering relationships documented in Chapter 5 reflect isolated, opportunistic extraction or organized, durable coalitions.

A key design choice is to separate two meanings of "relationship". Some edges are *procedural*, they appear because PBS auctions and the MEV-Boost payment convention require builders to transfer value to validators, so the tie is mechanically induced by the architecture. Other edges are *discretionary*, they reflect direct searcher routing, repeated bilateral exchanges, or payment patterns that are not mandated by PBS and therefore can encode stable partnerships, preferred pathways, or concentrated dependencies. The chapter, therefore, builds a family of MEV graphs, compares their community structure, and uses Sankey views to make centralization visible, especially the concentration of searcher outflows toward a restricted set of builders in blocks that carry MEV.

6.1 Role assignment

The network analysis in this chapter depends on assigning an interpretable role to each on-chain address, and on defining edges whose semantics follow the empirical MEV supply chain rather than raw contract-level execution. We start from the transaction-level dataset assembled in the earlier chapters for January 2024, where each transaction is already enriched with (i) visibility information (public versus private), (ii) MEV-related annotations when available, and (iii) traces that identify direct payments associated with block building. On top of this dataset, we introduce a conservative role taxonomy and a preprocessing step that prevents the graph from collapsing into a smart-contract-centric representation.

Role taxonomy. Each node corresponds to an address and is assigned a role label. We distinguish a specialized layer that captures the MEV and consensus pipeline, and a user layer that captures ordinary participation. Concretely, we use the labels *Searcher*, *Builder*, *Validator*, and *Builder–Validator* for specialized actors, and *Wallet*, *Smart Contract*, and *Sender* for the substrate. The *Smart Contract* label is tracked explicitly because a large fraction of economically meaningful interactions on Ethereum are routed through contract-mediated execution, and without explicit contract handling, the resulting network would primarily represent wallet-to-contract routing rather than relationships among economic agents.

How labels are assigned, native classes. Role assignment begins from a native separation between *Wallet* (externally owned accounts, EOAs) and *Smart Contract* addresses (accounts with deployed bytecode). In preprocessing, we label an address as *Smart Contract* when contract metadata or bytecode presence indicates deployed code, and as *Wallet* otherwise. This distinction is not merely descriptive. It is a prerequisite for two steps that recur throughout the chapter, first, the construction of a sender-centric graph, and second, the interpretation of communities as groups of economic agents rather than clusters of widely used execution routers.

How labels are assigned, discriminating actions. On top of the native classes, we assign operational MEV roles using on-chain signals that correspond to distinct capabilities in the supply chain.

Sender. A *Sender* is a wallet address that appears at least once as the *from* field of a transaction classified as private under our mempool observability definition, namely a transaction not seen in public-mempool traces prior to inclusion. This label is used to separate ordinary wallets that never originate private submissions from those that demonstrably use private channels.

Searcher. A *Searcher* is an address that initiates at least one transaction labeled as MEV-related by the transaction-level labeling pipeline used in the earlier chapters. In particular, this includes addresses associated with strategies or roles such as *arbitrage*, *liquidation*, *frontrun*, and *backrun*. Operationally, this criterion captures the discovery and strategy layer, the actors that repeatedly originate time-sensitive transactions whose profitability depends on ordering.

Builder. A *Builder* is identified via post-Merge builder attribution signals in

MEV-Boost-style blocks. Concretely, we assign the builder label to entities that are observed as *fee_recipient* in relayed blocks for which our attribution pipeline indicates external block construction under PBS conventions. This criterion is intentionally aligned with the separation between construction and proposal in the post-Merge regime, and it is designed to capture the entities that repeatedly construct blocks that include MEV-related transaction sequences.

Validator. A *Validator* is identified from proposer-side metadata and block-level revenue signals. We label an entity as a validator when it is observed as the *proposer_fee_recipient* in MEV-Boost blocks, and we additionally treat the *fee_recipient* in non-MEV-Boost blocks as a validator-side revenue address, because in that regime the block producer directly collects protocol-native revenues.

Builder–Validator and hybrid roles. A *Builder–Validator* is an address that satisfies both builder-side and validator-side signals over the observation window. This can occur for vertically integrated entities, for entities that appear on both sides in different blocks, or for intermediaries that mediate PBS-like flows across the two layers. The role system, therefore, explicitly allows hybrid categories to arise from co-occurring discriminating actions rather than assuming that actors are single-function by design.

For visualization and single-label summaries, we use a priority rule that assigns a unique role when an address satisfies multiple criteria. The guiding principle is to prioritize vertically integrated and infrastructure-heavy roles over substrate roles. In particular, the priority order is:

$$\text{Builder–Validator} \;>\; \{\text{Builder}, \text{Validator}, \text{Searcher}\} \;>\; \text{Sender} \;>\; \text{Wallet},$$

This priority order reflects a principle of *infrastructure responsibility*: we assign precedence to roles that indicate greater influence over the ordering and construction pipeline, since misclassifying a builder as a mere wallet would cause the graph to lose its most structurally significant nodes. The Builder-Validator compound role takes highest precedence because an entity that satisfies both signals operates across both the block construction and proposal layers, representing maximal vertical integration and ordering discretion. Searchers are ranked above passive substrate roles because their on-chain footprint is defined by active MEV strategy execution. Wallet and Sender are substrate roles that carry no infrastructure signal and are therefore deprioritized in ambiguous cases. with *Smart Contract* treated as a separate native category that is not interpreted as an economic agent and is handled explicitly in the preprocessing transformation described next.

Figure 6.1 summarizes the adopted taxonomy and the way hybrid roles arise from co-occurring signals. A structural fact that will recur throughout the chapter is that specialized roles are a minority by count while the user layer dominates numerically, yet specialized roles concentrate structural influence because they mediate the Searcher→Builder→Validator pipeline that determines ordering and block construction outcomes.

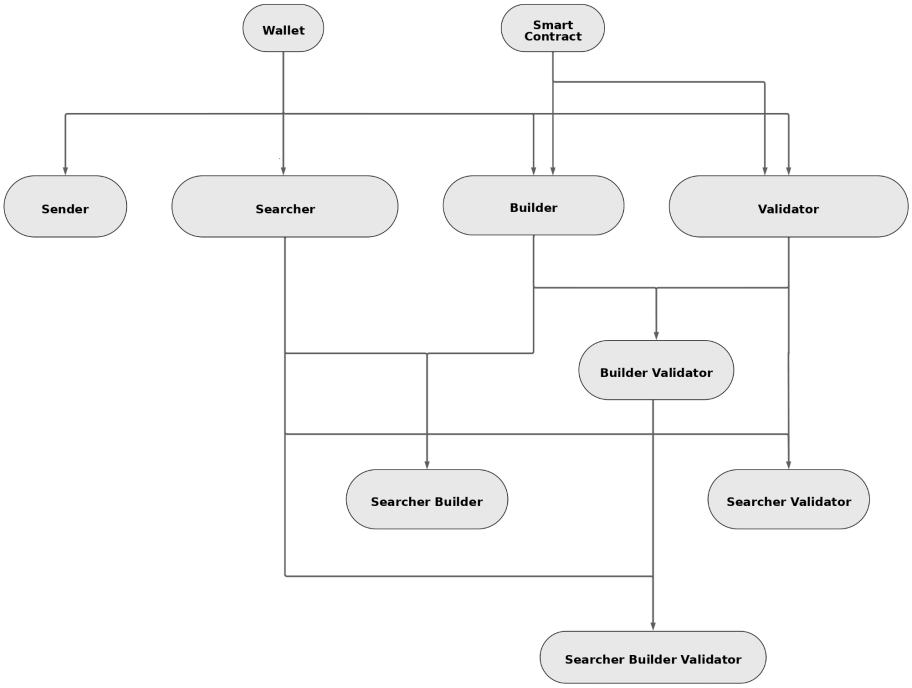


Figure 6.1: Role taxonomy adopted for the MEV graph. Nodes are addresses classified into specialized MEV roles (Searcher, Builder, Validator, Builder–Validator) and user-layer roles (Wallet, Sender), with smart contracts tracked explicitly as execution intermediaries. Hybrid roles arise when the same address satisfies multiple role-discriminating signals over the observation window.

6.1.1 From smart-contract centric to sender-centric

To surface real economic relationships, we transform the *smart-contract centric* view into a *wallet-centric* view. For each call chain, we replace intermediary smart contract addresses with the relevant wallet, namely the *original sender* that initiated the transaction, so that edges connect effective senders and recipients (Wallet, Sender, MEV actors). This transformation exposes relationships that would otherwise remain "hidden" behind proxy contracts, routers, and execution intermediaries.

Operationally, the transformation follows the steps in Fig. 6.2:

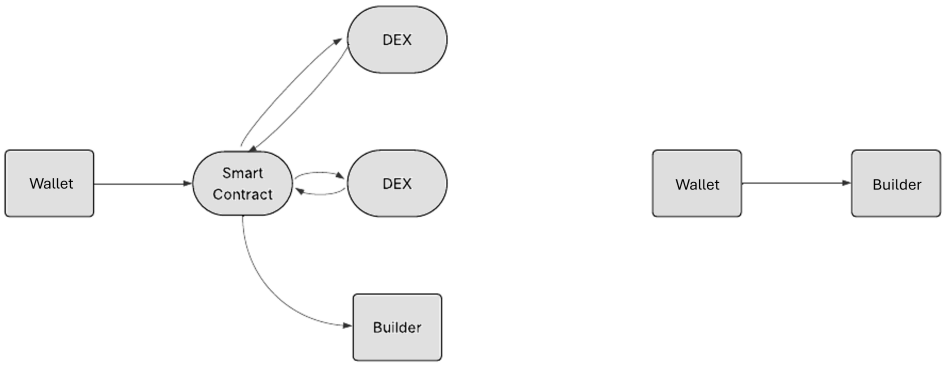
1. we start from a transaction of the form *Wallet* $\rightarrow$ *Smart Contract* and reconstruct the full *call trace*,

2. whenever the target of a call is a smart contract, we substitute it with the *original sender* address, so that the relationship becomes *Wallet* $\rightarrow X$, where X is the effective actor involved (native transfer, ERC-20 transfer, payment, forwarding),

3. redundant edges of the form *Wallet* $\rightarrow$ *Wallet*, which do not convey additional economic information, are removed from the graph,

4. the result is a *wallet-centric* view that removes intermediary contracts (proxy, router, pool, ...), and makes visible the effective economic ties that are relevant for interest-group analysis.

To resolve the ambiguity in multi-level call chains, the *original sender* of the outermost transaction is propagated as the effective actor for all internal calls that pass through intermediate contracts. Concretely, consider the call chain $W \xrightarrow{\texttt{call}} C_1 \xrightarrow{\texttt{call}} C_2 \xrightarrow{\texttt{transfer}} V$, where W is a wallet, C_1 and C_2 are smart contracts, and V is the final value recipient. Under the sender-centric transformation, both C_1 and C_2 are replaced by W, producing the edge $W \rightarrow V$ in the sender-centric graph. If V is itself a smart contract, we additionally inspect whether V's outbound transfers within the same transaction produce a final wallet recipient; if so, the edge is extended to $W \rightarrow V'$ where V' is the terminal wallet.



(a) Smart-contract centric view. (b) Transformed sender-centric view.

Figure 6.2: Comparison between a smart-contract centric representation (left) and the transformed sender-centric view (right).

6.1.2 Identifying MEV-Boost payments

To separate PBS-obligated ties from direct ties, we identify *Builder→Validator* transfers attributable to MEV-Boost, via relay and fee-recipient conventions. Across different periods, we achieve coverage of about ~187k out of ~200k estimated payments, and the remaining unmatched cases are compatible with fee-recipient configurations that point directly to the validator.

6.2 USD valuation scheme

USD conversion makes different tokens comparable and enables community analysis in terms of *economic volumes* exchanged, not only in terms of interaction frequency.

For each transaction t, we analyze all token transfer events associated with it. For each event e, we identify:

- the transferred token $a(e)$,
- the transferred quantity $q(e)$, expressed in real units, namely normalized by *decimals*,
- the transaction timestamp $\tau(t)$.

We then retrieve the market price in US dollars, $\mathrm{px}_{\mathrm{USD}}(a(e), \tau(t))$, for the token $a(e)$ on the day of the transaction. The USD value of the transfer event is computed as:

$$\mathrm{usd}(e) = q(e) \cdot \mathrm{px}_{\mathrm{USD}}(a(e), \tau(t)) \,.$$

If the transaction has a wallet sender and a wallet recipient (e.g., $u \to v$), the total USD value is the sum of event values:

$$\mathrm{usd}(t) = \sum_{e \in E(t)} \mathrm{usd}(e).$$

If instead the recipient is a *smart contract*, we separately consider inbound and outbound transfers from the contract within the same transaction t. Let:

- $\mathrm{IN}(t)$ be the set of events whose recipient is the smart contract,
- $\mathrm{OUT}(t)$ be the set of events whose sender is the smart contract.

We estimate the USD volume of the transaction as the maximum of the two valued flows:

$$\mathrm{usd}_{\mathrm{SC}}(t) = \max\left(\sum_{e \in \mathrm{IN}(t)} \mathrm{usd}(e), \sum_{e \in \mathrm{OUT}(t)} \mathrm{usd}(e) \right).$$

This approach yields a conservative but representative estimate of USD volume routed through smart contracts, and avoids value duplication caused by internal transfers.

6.3 Scenarios and metrics, detailed analysis

In this section, we present a complete review of the scenarios analyzed in our *MEV Graph* study. The analysis proceeds in two phases: (i) an initial exploratory analysis of the full private-only graph, without scenario filters, and (ii) the construction of controlled scenarios that isolate MEV relationships with and without PBS payments (MEV-Boost). The analyses are accompanied by Sankey visualizations, heatmaps, and role and community tables, with detailed commentary.

6.3.1 Initial analysis, full private-only graph

Before moving to the specific scenarios, we perform an exploratory analysis of the full MEV ecosystem, considering all private transactions available in the dataset. This provides a global view of structural dynamics before applying restrictions.

Relationships reflect the typical operational MEV flow: Searchers identify opportunities and route them to Builders, which in turn may interact with Validators directly or through hybrid entities such as Builder-Validators. Specifically, we retain the following link types: Searcher→Builder, Searcher→Builder Validator, Searcher→Validator, Builder→Validator, Builder→Builder Validator, and Builder Validator→Validator.

Complete Sankey visualization. The complete Sankey representation (Fig. 6.3) reveals role-to-role flows. The visualization highlights how a small number of Builder and Builder Validator nodes act as critical bridges between roles.

Main component and role distribution. The main component contains 255,347 nodes, dominated numerically by Wallets (152,314) and Senders (91,049), as shown in Table 6.1. This distribution indicates that even within private transactions, specialized MEV actors (Searcher, Builder, Validator) remain a minority.

Table 6.1: Role distribution in the main component of the private-only graph.

Role	Count
Wallet	152,314
Sender	91,049
Validator	7,092
Searcher	4,054
Builder Validator	772
Builder	64

Community structure and modularity. Applying the Leiden algorithm to the full graph produces a partition with exceptionally high modularity, 0.81, indicating strong community structure. This suggests that the MEV graph is not random, but organizes into well-defined clusters. The most significant communities, reported in Table 6.2, display heterogeneous compositions that reflect different organizational regimes. Community 0, with 33,841 members, is wallet-centric, suggesting a cluster oriented toward baseline private activity. The presence of only 196 Searchers indicates limited direct MEV specialization within this cluster.

Community 4 is validator-centric, with 6,954 Validators and 759 Builder Validators, and plausibly represents the infrastructural core of PBS validation and block construction.

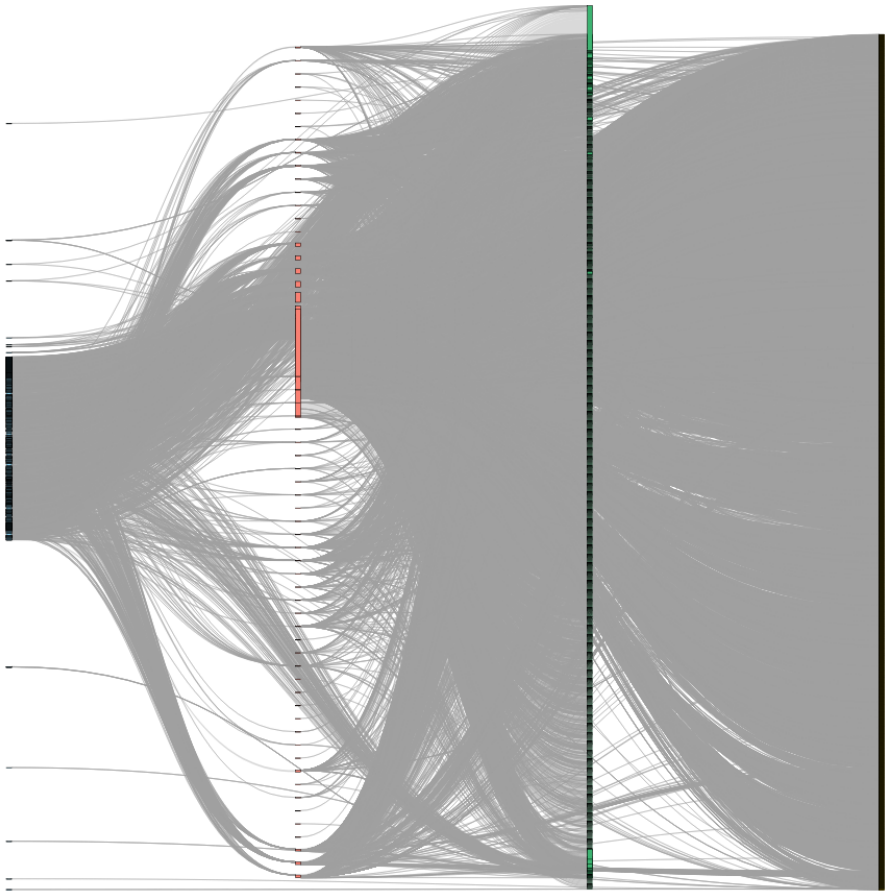


Figure 6.3: Complete Sankey representation of Searcher→Builder (or Builder-Validator)→Validator (or Builder-Validator) relationships in the private-only graph.

Table 6.2: Detailed analysis of major communities in the private-only graph. BV= Builder Validator, V= Validator, B= Builder

Community	Size	Wallet	Sender	Searcher	Other
0	33,841	24,763 (73.2%)	8,878 (26.2%)	196 (0.6%)	BV=3; V=1
1	21,028	17,327 (82.4%)	3,530 (16.8%)	163 (0.8%)	V=8
2	20,393	17,791 (87.2%)	2,459 (12.1%)	137 (0.7%)	V=6
3	19,921	5,402 (27.1%)	12,909 (64.8%)	1,608 (8.1%)	BV=1; B=1
4	18,811	1,281 (6.8%)	9,050 (48.1%)	711 (3.8%)	BV=759; V=6,954; B=54
8	9,747	7,319 (75.1%)	2,181 (22.4%)	243 (2.5%)	BV=2; V=2

Temporal dynamics and cluster persistence. Daily interaction analysis during January 2024, visualized in the heatmap (Fig. 6.4), reveals temporal patterns that support the hypothesis of stable groups. The heatmap shows persistent activity in time for some communities, while others exhibit localized spikes. These patterns suggest the existence of MEV subgroups with distinct

operational behavior, some stable and infrastructural, others more sporadic.

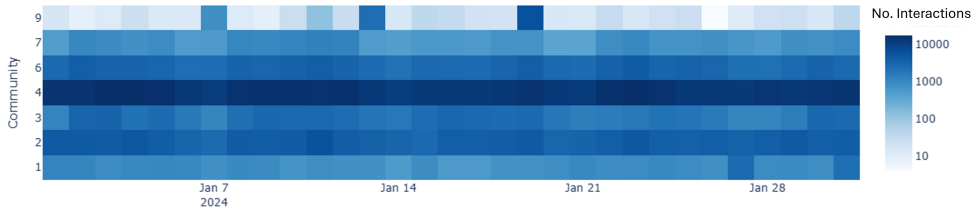


Figure 6.4: Heatmap of daily interactions by community during January 2024.

6.3.2 Scenario 1, MEV-only without PBS payments

After the exploratory analysis, we built the first scenario by focusing exclusively on relationships among specialized MEV actors (Searcher, Builder, Validator), while deliberately excluding all payments attributable to the PBS system. This isolates "voluntary" or direct coordination dynamics, not mediated by MEV-Boost auctions. Excluding PBS payments is crucial. MEV-Boost transfers are "obligated" by the PBS architecture; the winning Builder must pay the Validator. Direct relationships among MEV actors, in contrast, can reveal preferential routing, strategic partnerships, or other coordination patterns that are not mechanically imposed.

Group 1a, MEV actors only

The main component in this scenario is much smaller than the full graph, with 2,332 nodes and an extreme concentration of specialized actors. Searchers dominate (1,995 nodes, 85.5%), followed by Builder Validators (237 nodes, 10.2%), pure Validators (57 nodes, 2.4%), and pure Builders (43 nodes, 1.8%).

This distribution highlights the hierarchical structure of the MEV ecosystem, where many Searchers compete at the opportunity discovery layer, while Builders and Validators are fewer and more specialized. Leiden on this subgraph yields modularity 0.3174, lower than the full graph but still non-trivial. We identify nine main communities (0 to 8), each reflecting a distinct cooperation pattern between Searchers and the execution layer.

MEV interactions and transaction-type patterns. A detailed interaction breakdown for Community 0 shows operational patterns that characterize the cluster. The 5,896 Searcher→Builder swap interactions with Builder payments represent the core activity, consistent with trading and arbitrage specialization.

Community 3 exhibits a different profile. The presence of 303 Builder Validator→Builder interactions without payment is particularly interesting, and may indicate infrastructural coordination or information sharing rather than direct extraction.

Table 6.3: Role distribution across main communities in Scenario 1a.

Community	Size	Searcher	Builder Validator	Validator	Builder
0	779	734 (94.2%)	33 (4.2%)	2 (0.3%)	10 (1.3%)
1	466	424 (91.0%)	39 (8.4%)	0 (0%)	3 (0.6%)
2	391	316 (80.8%)	68 (17.4%)	0 (0%)	7 (1.8%)
3	291	161 (55.3%)	75 (25.8%)	44 (15.1%)	11 (3.8%)
4	149	133 (89.3%)	13 (8.7%)	0 (0%)	3 (2.0%)
5	101	96 (95.0%)	0 (0%)	0 (0%)	5 (5.0%)
6	94	86 (91.5%)	2 (2.1%)	1 (1.1%)	5 (5.3%)
7	8	3 (37.5%)	4 (50.0%)	1 (12.5%)	0 (0%)
8	8	5 (62.5%)	1 (12.5%)	1 (12.5%)	1 (12.5%)

Sankey visualization and operational flows. The Sankey diagram for Scenario 1a (Fig. 6.5) confirms the expected three-layer structure, and reveals strong asymmetries in flow distribution. A small set of Builders aggregates the activity of many Searchers, while the distribution toward Validators appears more fragmented once PBS payments are removed.

USD volume analysis (Fig. 6.6) reveals additional concentration patterns that do not necessarily match interaction-count concentration. Some relatively small communities move large volumes.

Group 1b, adding Wallets

Adding Wallets changes both the size and the structure of the graph. The main component increases to 19,236 nodes, and modularity rises to 0.6758. This increase is methodologically meaningful. Adding nodes could, in principle, dilute structure, but the opposite occurs, suggesting that Wallets act as selective bridges that reinforce community boundaries rather than washing them out.

Table 6.4: Role distribution across selected communities of interest (modularity = 0.6758).

Community	Size	Wallet	Searcher	BV	Validator	Builder
0	5,084	4,977 (97.9%)	104 (2.0%)	1 (0.0%)	2 (0.0%)	0 (0.0%)
1	3,063	1,782 (58.2%)	1,275 (41.6%)	5 (0.2%)	6 (0.2%)	1 (0.0%)
2	2,119	956 (45.1%)	981 (46.3%)	150 (7.1%)	4 (0.2%)	28 (1.3%)
3	1,775	1,629 (91.8%)	118 (6.6%)	0 (0.0%)	24 (1.4%)	4 (0.2%)
7	418	320 (76.6%)	91 (21.8%)	0 (0.0%)	7 (1.7%)	0 (0.0%)
11	194	112 (57.7%)	5 (2.6%)	0 (0.0%)	32 (16.5%)	1 (0.5%)
13	138	40 (29.0%)	80 (58.0%)	0 (0.0%)	44 (31.9%)	5 (3.6%)
15	124	111 (89.5%)	8 (6.5%)	0 (0.0%)	4 (3.2%)	1 (0.8%)
16	107	92 (86.0%)	13 (12.1%)	0 (0.0%)	1 (0.9%)	1 (0.9%)

Community 0 is almost entirely Wallet-based (97.9%), with only a minimal presence of specialized MEV actors. In contrast, Community 2 shows a near-perfect balance between Wallets (45.1%) and Searchers (46.3%), suggesting a cluster where MEV activity is tightly integrated with baseline transaction activity. The non-trivial share of Builder Validators (7.1%) indicates a more integrated operational model, in which discovery, execution, and validation intermediaries co-exist within the same cluster.

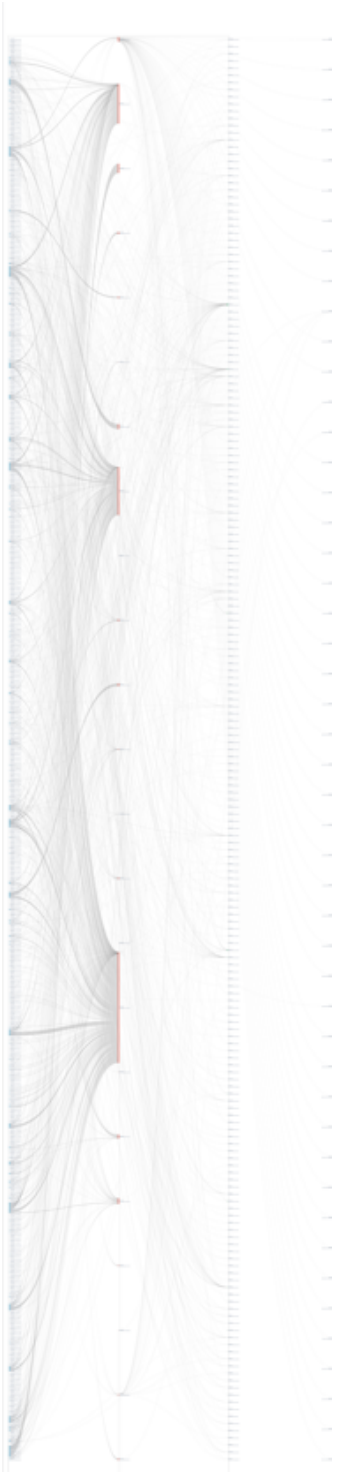


Figure 6.5: Sankey diagram for MEV-only without PBS payments (Scenario 1a).

A detailed interaction analysis for Community 2 confirms the richness of this integrated cluster. Swap interactions Searcher→Builder with Builder payments dominate, followed by arbitrage transactions. Particularly notable are Searcher→Wallet swaps without payment, which can reflect internal transfers within the cluster or cluster-internal routing.

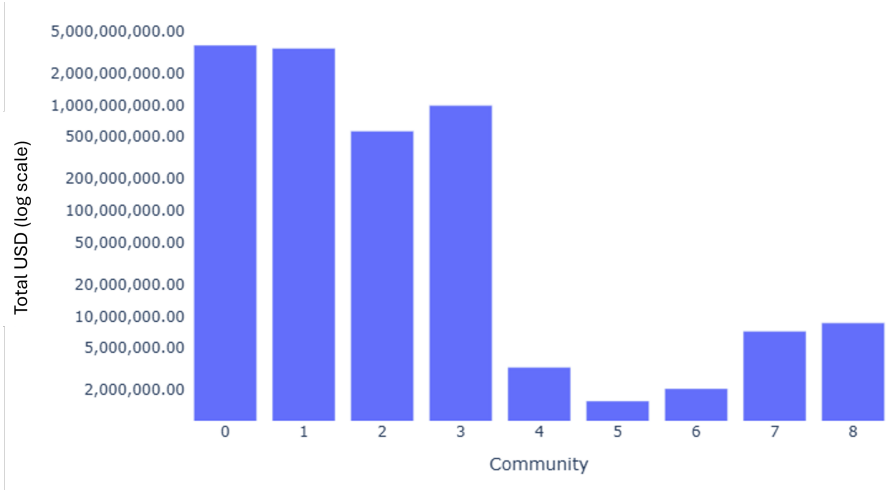


Figure 6.6: USD volume analysis by community in Scenario 1a.

Community 11 exhibits a coordination-heavy profile characterized by Builder Validator→Builder interactions without payment, suggesting a cluster specialized in infrastructural coordination rather than direct extraction. This pattern may reflect partnerships or shared operational infrastructure.

MEV-type diversification and operational specialization. Including Wallets not only increases interactions, it also diversifies their types. The distribution of MEV types (Fig. 6.7) shows high prevalence of swaps and arbitrage.

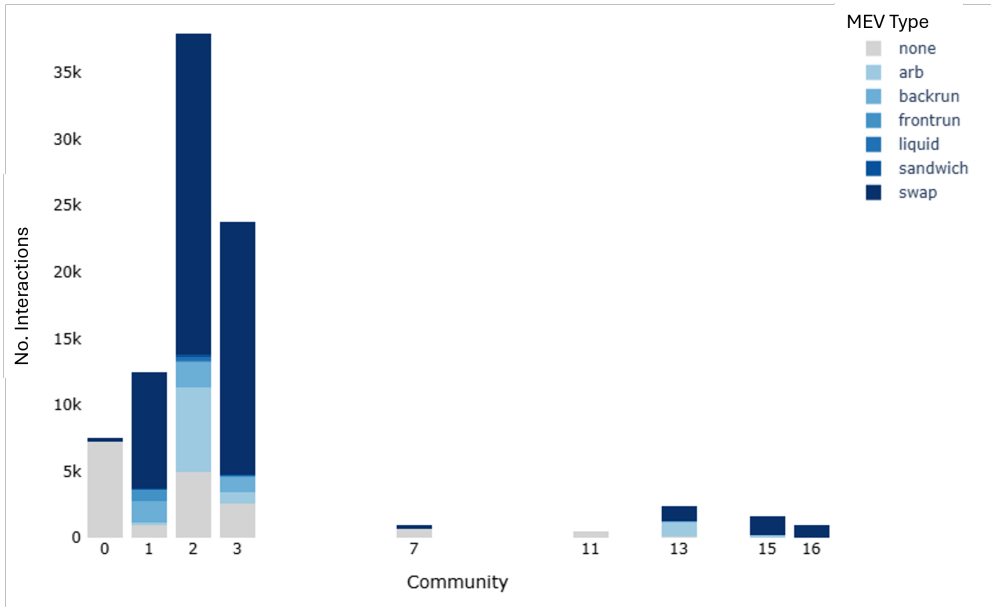


Figure 6.7: Distribution of MEV types by community in Scenario 1b.

Temporal dynamics (Fig. 6.8) show more regular and persistent daily patterns than Scenario 1a, consistent with the idea that Wallet inclusion stabilizes MEV activity by spreading participation across a broader base.

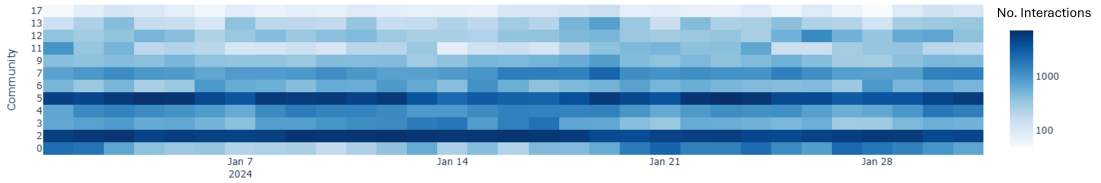


Figure 6.8: Daily interaction patterns in Scenario 1b.

Group 1c, full inclusion with Wallets and Senders

The final expansion of Scenario 1 includes all actor types: Searcher, Builder, Validator, Wallet, and Sender. This configuration yields the largest and structurally richest direct-tie graph, with 38,596 nodes and modularity 0.6987, the highest observed among the MEV-only scenarios. Such high modularity indicates that the full MEV ecosystem exhibits strong community structure even under direct-tie semantics.

Table 6.5: Community structure in Scenario 1c with all actors included.

C.	Size	Sender	Wallet	Searcher	BV	Validator	Builder
0	7,335	6,386 (87.1%)	172 (2.3%)	420 (5.7%)	276 (3.8%)	47 (0.6%)	34 (0.5%)
1	6,318	918 (14.5%)	5,247 (83.0%)	150 (2.4%)	1 (0.02%)	2 (0.03%)	0 (0%)
2	6,090	5,943 (97.6%)	14 (0.23%)	128 (2.1%)	3 (0.05%)	1 (0.02%)	1 (0.02%)
3	5,393	1,108 (20.5%)	2,405 (44.6%)	1,875 (34.8%)	5 (0.09%)	0 (0%)	0 (0%)
4	3,052	748 (24.5%)	1,544 (50.6%)	682 (22.3%)	5 (0.16%)	73 (2.4%)	0 (0%)
5	2,526	503 (19.9%)	1,842 (72.9%)	165 (6.5%)	10 (0.40%)	0 (0%)	6 (0.24%)
8	919	177 (19.3%)	0 (0%)	3 (0.33%)	234 (25.5%)	0 (0%)	0 (0%)
9	603	455 (75.5%)	41 (6.8%)	91 (15.1%)	13 (2.2%)	2 (0.33%)	1 (0.17%)
11	409	288 (70.4%)	112 (27.4%)	6 (1.5%)	1 (0.24%)	2 (0.49%)	0 (0%)
14	193	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (0.52%)	0 (0%)

The MEV-type distribution in the full scenario (Fig. 6.9) confirms the predominance of swaps and arbitrage. The USD volumetric analysis (Fig. 6.10) shows concentration patterns consistent with the previous cases.

6.3.3 Scenario 2, PBS payments only

After examining "voluntary" MEV dynamics in Scenario 1, we now adopt the opposite approach: we retain only ties attributable to Proposer Builder Separation (PBS), namely MEV-Boost payments in Ethereum. This isolates the procedural structure of the post-Merge PBS architecture. The motivation is to observe MEV through the PBS lens. Scenario 1 captures direct, potentially preferential relationships; Scenario 2 highlights ties that are obligated by PBS mechanics. Comparing the two is crucial to distinguish emergent organization from protocol-induced structure.

Group 2a, MEV actors only with PBS payments

The resulting structure differs sharply from Scenario 1. The nominal number of nodes increases (10,081), and the PBS-only graph shows modularity 0.3252,

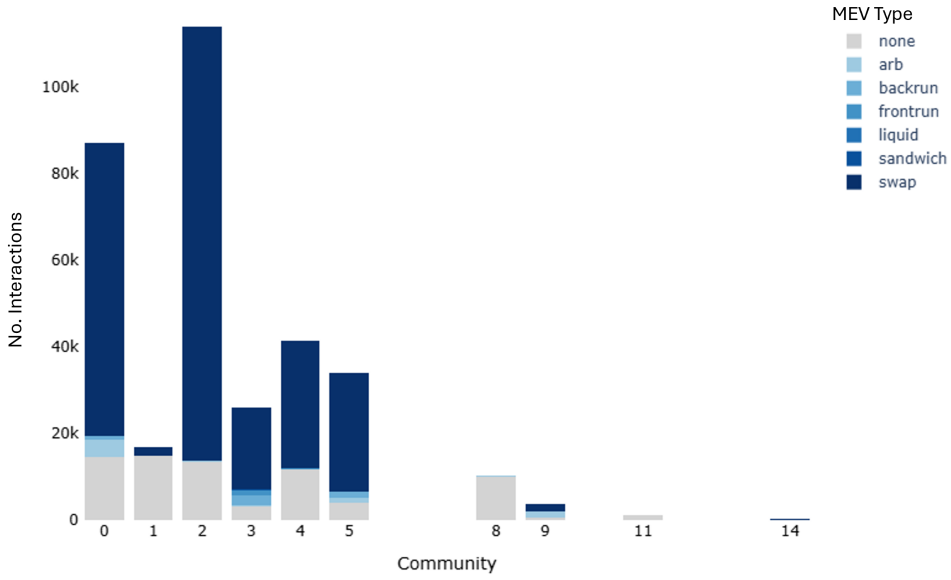


Figure 6.9: MEV-type distribution in Scenario 1c.

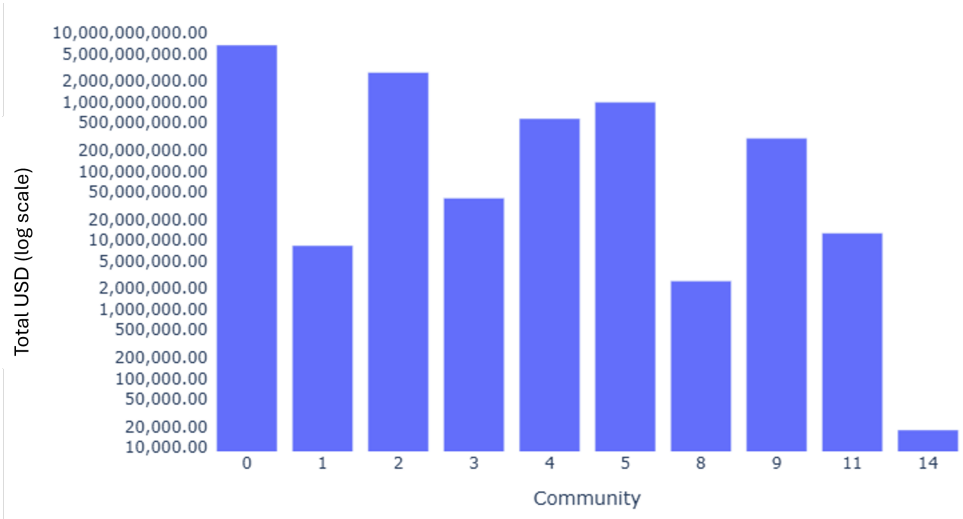


Figure 6.10: USD volume concentration in the full MEV ecosystem (Scenario 1c).

indicating weaker cohesion than the rich direct-tie variants. This reduction suggests that PBS payments create functional links, but do not necessarily produce the same community cohesion as discretionary routing.

Role distributions reveal a validator-centric architecture. Validators constitute the majority in most large communities, reflecting PBS mechanics, Builders pay Validators, producing many recipient-side nodes.

Community 0 shows extreme validator dominance (81.8%) with limited Searcher presence (7.4%). Community 2 is more balanced, with 40.0% Searchers and 56.6% Validators, consistent with a cluster where PBS activity is more tightly coupled with MEV opportunity discovery.

Table 6.6: Community composition in the PBS-only graph (Scenario 2a).

Community	Size	Searcher	Validator	Builder Validator	Builder
0	4,087	302 (7.4%)	3,343 (81.8%)	421 (10.3%)	20 (0.5%)
1	2,011	302 (15.0%)	1,574 (78.3%)	125 (6.2%)	10 (0.5%)
2	1,952	780 (40.0%)	1,105 (56.6%)	63 (3.2%)	4 (0.2%)
3	621	72 (11.6%)	470 (75.7%)	67 (10.8%)	12 (1.9%)
4	285	85 (29.8%)	185 (64.9%)	14 (4.9%)	1 (0.4%)
6	239	186 (77.8%)	41 (17.2%)	9 (3.8%)	3 (1.3%)
7	236	202 (85.6%)	23 (9.7%)	5 (2.1%)	6 (2.5%)
8	197	56 (28.4%)	123 (62.4%)	17 (8.6%)	1 (0.5%)
9	25	0 (0%)	24 (96.0%)	0 (0%)	1 (4.0%)
12	19	2 (10.5%)	13 (68.4%)	1 (5.3%)	3 (15.8%)
13	11	2 (18.2%)	8 (72.7%)	0 (0%)	1 (9.1%)

Interaction patterns and transactional structure. Community-level interaction summaries emphasize the procedural nature of PBS. Specific patterns highlight how some Builders concentrate searcher flows, even in a PBS-only representation, which can suggest preferential relationships that survive the projection through payment edges.

Sankey visualization and institutional flows. The PBS-only Sankey (Fig. 6.11) reveals a structure dominated by Builder→Validator (Builder→Builder-Validator) flows, with reduced visibility of Searcher activity. This reflects the fact that PBS payments are downstream of block construction, and therefore capture the auction outcome more than the upstream searcher routing.

MEV-type analysis (Fig. 6.12) shows simplification relative to direct scenarios, with high prevalence of transactions classified as “none” alongside “swap”, consistent with payment-focused edges. USD volume analysis (Fig. 6.13) shows extreme concentration, few communities dominate volumes.

Group 2b, PBS with Wallet inclusion

Adding Wallets produces a distinctive effect. The graph expands to 117,626 nodes with modularity 0.8246, the highest observed.

High modularity (0.8246) indicates that Wallets in the PBS context act less as integrative bridges and more as separators, producing extremely segregated clusters.

Table 6.7: Extreme segregation under PBS + Wallets (Scenario 2b).

C.	Size	Wallet	Validator	BV	Searcher	Builder
0	26,018	25,981 (99.9%)	7 (0.03%)	1 (0.004%)	29 (0.1%)	0 (0%)
3	9,787	2,217 (22.7%)	6,281 (64.2%)	673 (6.9%)	561 (5.7%)	54 (0.6%)
4	8,203	6,525 (79.5%)	10 (0.12%)	2 (0.02%)	1,666 (20.3%)	0 (0%)
5	4,451	4,207 (94.5%)	7 (0.16%)	1 (0.02%)	235 (5.3%)	2 (0.04%)
6	3,264	3,151 (96.5%)	7 (0.21%)	3 (0.09%)	101 (3.1%)	2 (0.06%)
17	1,264	1,217 (96.3%)	2 (0.16%)	1 (0.08%)	44 (3.5%)	0 (0%)
18	1,218	929 (76.3%)	67 (5.5%)	4 (0.33%)	218 (17.9%)	0 (0%)

Community 3 is validator-centric (64.2% Validators) but contains a meaningful share of Wallets (22.7%) and Searchers (5.7%), suggesting a PBS integration cluster where multiple actor types co-exist within the same community.

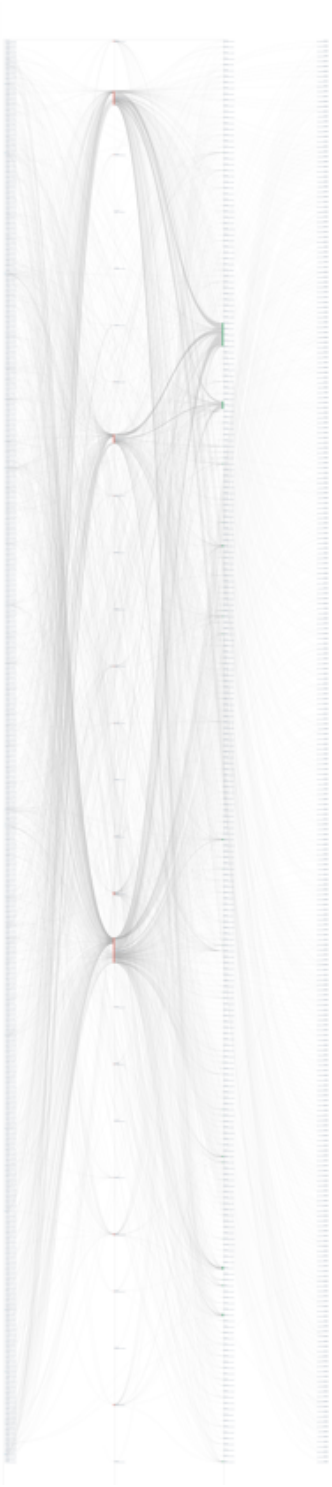


Figure 6.11: PBS architecture in Sankey form, dominance of Builder→Validator flows.

MEV-type analysis (Fig. 6.14) remains dominated by “none” transactions, consistent with PBS payments. Temporal dynamics (Fig. 6.15) are more regular than pure PBS, suggesting that Wallets stabilize PBS activity by distributing participation more evenly over time.

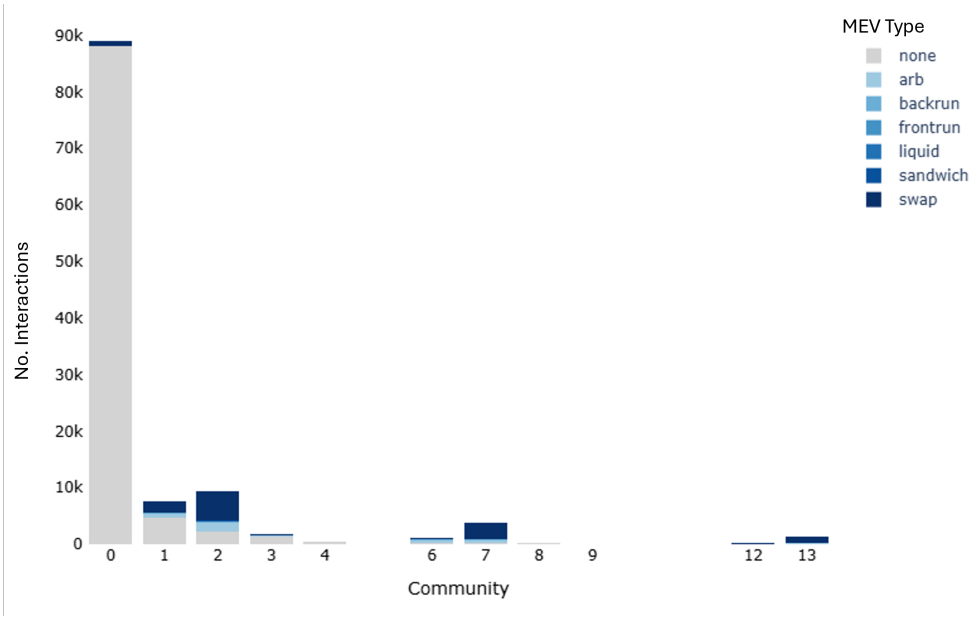


Figure 6.12: Simplification of MEV types under PBS-only ties.

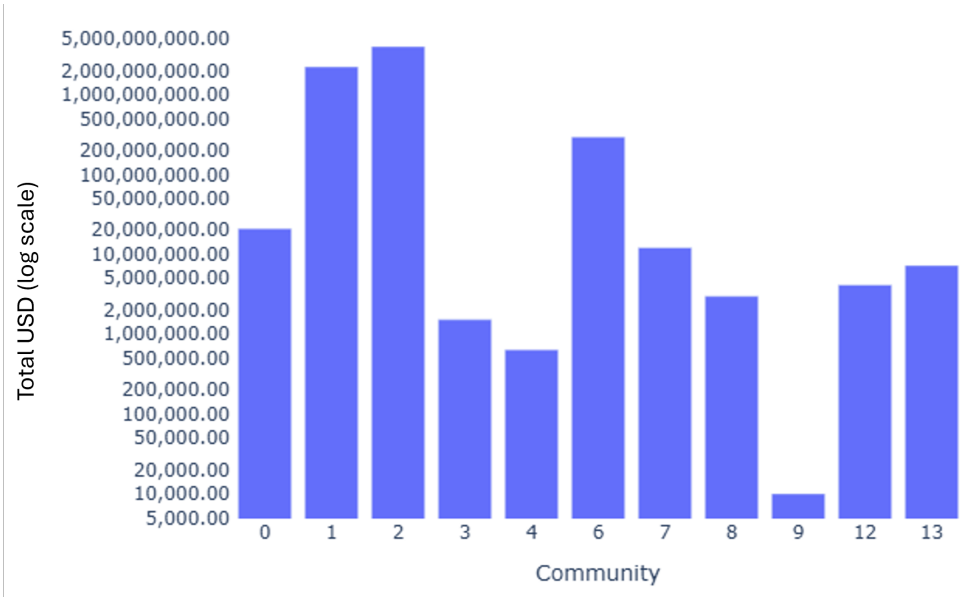


Figure 6.13: Extreme USD volume concentration under PBS-only ties.

Group 2c, full PBS with Wallets and Senders

The final PBS configuration includes all actor types, reaching 147,162 nodes with modularity 0.7743. Although lower than PBS + Wallets, modularity remains very high, indicating that the full PBS ecosystem retains strong community structure even under maximal inclusion.

Community 2 is the most integrated cluster in the full PBS graph, combining Senders (42.1%), Validators (35.4%), Wallets (15.3%), and Builder Validators (5.0%). This composition suggests an operational regime where origin, intermediation, execution, and validation are co-present within a single community.

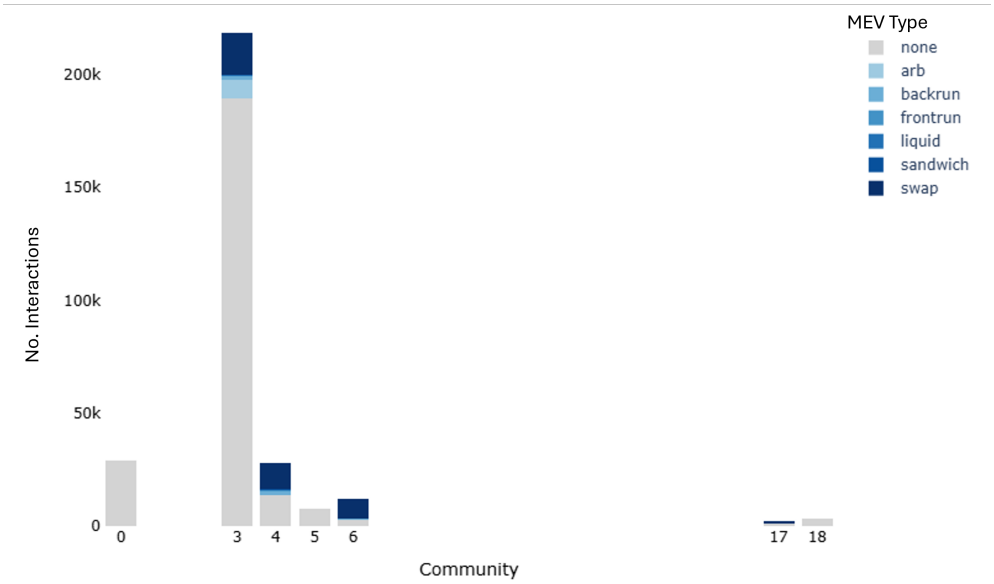


Figure 6.14: MEV-type distribution under PBS + Wallets (Scenario 2b).

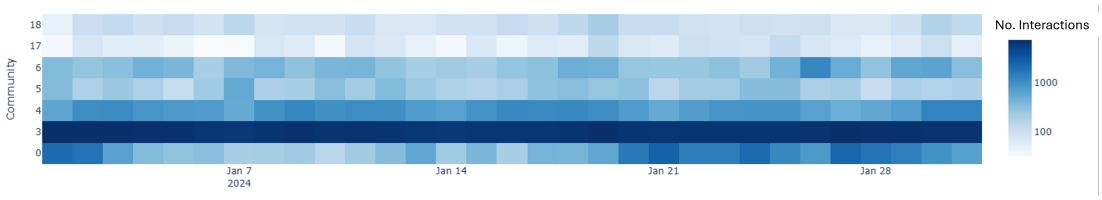


Figure 6.15: Temporal regularization of PBS activity through Wallet inclusion.

Table 6.8: Community architecture under PBS + Wallets + Senders (Scenario 2c).

C.	Size	Sender	Wallet	Validator	Searcher	BV	Builder
0	27,275	1,266 (4.6%)	25,980 (95.2%)	9 (0.03%)	19 (0.07%)	1 (0.004%)	0 (0%)
2	12,569	5,286 (42.1%)	1,920 (15.3%)	4,452 (35.4%)	233 (1.9%)	626 (5.0%)	51 (0.4%)
3	10,777	101 (0.94%)	10,652 (98.8%)	5 (0.05%)	18 (0.17%)	1 (0.009%)	0 (0%)
4	9,904	1,921 (19.4%)	6,406 (64.7%)	7 (0.07%)	1,569 (15.8%)	1 (0.01%)	0 (0%)
5	9,669	6,868 (71.0%)	348 (3.6%)	2,116 (21.9%)	232 (2.4%)	100 (1.0%)	5 (0.05%)
6	8,267	1,573 (19.0%)	6,382 (77.2%)	7 (0.08%)	304 (3.7%)	1 (0.01%)	0 (0%)
7	8,254	582 (7.1%)	5,696 (69.0%)	40 (0.48%)	1,934 (23.4%)	1 (0.01%)	1 (0.01%)
9	4,899	481 (9.8%)	3,659 (74.7%)	28 (0.57%)	729 (14.9%)	0 (0%)	1 (0.02%)
11	3,701	352 (9.5%)	3,307 (89.4%)	33 (0.89%)	8 (0.22%)	1 (0.03%)	0 (0%)
12	3,688	583 (15.8%)	3,012 (81.7%)	13 (0.35%)	76 (2.1%)	3 (0.08%)	1 (0.03%)
13	3,014	564 (18.7%)	2,232 (74.1%)	41 (1.4%)	174 (5.8%)	2 (0.07%)	1 (0.03%)
17	1,513	1,484 (98.1%)	0 (0%)	1 (0.07%)	0 (0%)	0 (0%)	1 (0.07%)

Temporal dynamics (Fig. 6.16) are stable and consistent with a mature PBS environment.

6.3.4 Comparative analysis across scenarios

The comparison between Scenarios 1 and 2 reveals fundamental differences in the organizational architecture of the MEV ecosystem that go beyond simple numeric scale or USD volumes.

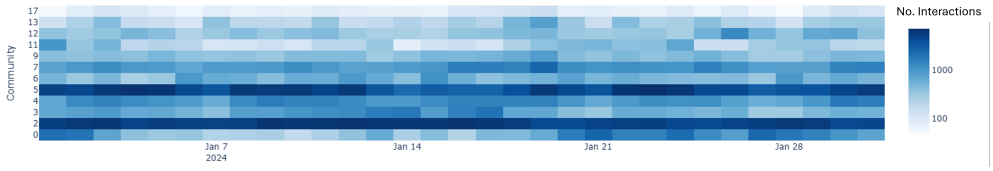


Figure 6.16: Stable daily interaction patterns under mature PBS (Scenario 2c).

Modularity evolution and structural cohesion

Modularity evolution across scenarios is a key result.

Table 6.9: Comparative modularity evolution across all scenarios.

Scenario	Nodes	Modularity
1a: MEV only	2,332	0.3174
1b: MEV + Wallet	19,236	0.6758
1c: Full MEV	38,596	0.6987
2a: PBS only	10,081	0.3252
2b: PBS + Wallet	117,626	0.8246
2c: Full PBS	147,162	0.7743

The progression 1a→1b→1c in Scenario 1 shows a steady increase in modularity, suggesting that progressively including non-specialized actors strengthens community structure rather than diluting it. This pattern supports the hypothesis that the MEV ecosystem is intrinsically structured along functional lines that extend beyond specialized actors.

In Scenario 2, the progression 2a→2b→2c shows an extreme peak (0.8246 in PBS + Wallets) followed by a reduction (0.7743 in full PBS). This suggests that PBS achieves maximal segregation at an intermediate inclusion level, and then becomes slightly more integrated once the full sender layer is included, while remaining highly organized.

Temporal patterns and operational stability

Temporal dynamics differ qualitatively. Scenario 2 (PBS) shows more volatile and opportunistic patterns at the specialized level, consistent with ties driven by slot-level auction outcomes. Scenario 1 (direct ties) shows more regular and predictable patterns, especially once Wallets and Senders are included, consistent with persistent participation and repeated relationships.

Strategic diversification and MEV innovation

MEV type comparisons reveal a trade-off between strategic diversity and procedural regularity. Direct-tie scenarios (Scenario 1) show higher diversity and greater presence of complex strategies, including backrun, frontrun, and arbitrage. PBS scenarios show a prevalence of non-MEV-like transactions tied to

Builder→Validator payments, reflecting the procedural nature of the retained edges.

6.3.5 Intra- and extra-community relations

The analysis of interactions across communities shows that connections are not uniformly distributed between internal and external ties. In both scenarios, the majority of relations are *extra-community*, indicating a structural tendency to connect outward. In Scenario 1 (without PBS payments), the intra-community share averages around 40%, with values ranging from about 20% up to above 70% in specific cases. In Scenario 2 (with PBS payments), the intra-community share decreases. In most cases internal interactions remain below 30%, while external relations stably exceed 80%. Overall, Scenario 2 therefore exhibits a more polarized distribution, with a clear prevalence of *extra-community* ties and reduced internal cohesion.

6.3.6 Discussion

This section connects the scenario evidence explicitly to **RQ3**, namely, "Which stable groups and interest clusters form within the MEV ecosystem, how do PBS-mediated relationships differ from direct transfers, and what is the degree of centralization among extracting actors".

First, the results support the existence of *stable groups and interest clusters*. The full private-only graph exhibits very high modularity (0.81) and persistent daily activity patterns for several communities (Fig. 6.4), which is difficult to reconcile with random mixing. Scenario 1 strengthens this interpretation, because modularity increases as the user layer is included, 0.3174 in the specialized-only view, 0.6758 with Wallets, and 0.6987 with Wallets and Senders. This monotonic increase indicates that the specialized MEV layer is not isolated from the user substrate. Instead, it is embedded in structured participation patterns in which private senders and wallets cohere around recurring interaction surfaces. Operationally, this is consistent with stable routing habits, repeated counterparties, and cluster-specific execution pathways, rather than one-off opportunistic interactions distributed uniformly across the ecosystem.

Second, the comparison clarifies how *PBS-mediated relationships differ from direct transfers*. Scenario 2 is structurally different from Scenario 1 because edges are procedurally obligated by the PBS architecture. In the specialized view, PBS-only modularity (0.3252) remains modest, and communities are validator-centric (Table 6.6), consistent with payments that reflect auction outcomes more than discretionary coordination. At the same time, once Wallets are included, modularity becomes extremely high (0.8246 in Scenario 2b), and remains high under full inclusion (0.7743 in Scenario 2c). This shows that strong community separation can also arise from procedural connectivity when embedded in the wallet-level substrate. Therefore, for **RQ3**, the key takeaway is not that "high modularity implies collusion", but that *the meaning of modularity depends on edge semantics*. Under direct ties, increasing modularity is more

suggestive of stable relationships and interest clusters. Under PBS-only ties, high modularity is more naturally interpreted as segregation induced by the institutional payment backbone plus heterogeneous wallet participation.

Third, the Sankey visualizations speak directly to *centralization among extracting actors*. In Scenario 1a and in the full private-only view (Fig. 6.3, Fig. 6.5), many Searchers route activity toward a comparatively small execution layer, and a small number of Builder or Builder Validator nodes appear as bridges between Searchers and Validators. This pattern is consistent with centralization in block construction and ordering mediation. Even when Searchers are numerous, the construction layer constitutes a structural bottleneck, which implies that extracting power can concentrate even if opportunity discovery remains competitive. Scenario 2 further emphasizes a centralized payment backbone through dominant Builder→Validator flows (Fig. 6.11), reflecting a procedural centralization in which a small set of Builders repeatedly interacts with a large validator recipient set.

Taken together, these findings answer **RQ3** as follows. The MEV ecosystem exhibits stable clusters that persist over time and remain visible under direct-tie semantics, consistent with the presence of interest regions and repeated counterparties. PBS-mediated edges encode a different social meaning, they impose a procedural payment network that can generate high segregation when embedded in wallet participation, but that is less informative about discretionary coordination among Searchers and Builders. Finally, centralization is most visible in the construction layer, Sankey flows and bridge roles show that a restricted set of builders and hybrid entities mediate a disproportionate share of MEV-relevant routing and payment relationships.

7

Cross-Chain MEV and Multihop Arbitrage

The preceding chapters studied how MEV is realized *within* a single chain, focusing on post-Merge Ethereum, Solana, and the role of PBS, private order flow, and off-chain coordination in shaping inclusion and ordering outcomes. This chapter moves one level up and studies **cross-chain MEV**, namely extractable value that arises when liquidity and price discovery are distributed across multiple execution environments, including L1s and rollups, that are connected by bridges.

This chapter answers **RQ4 (Cross-chain dynamics)**: *How does MEV manifest across chains, what is the feasibility and frequency of multi-hop cross-chain arbitrage, and how do bridge latency and costs shape profitability and execution time*. The organizing intuition is that multi-chain growth increases *liquidity fragmentation*, and fragmentation creates persistent price discrepancies. At the same time, realizing these discrepancies requires infrastructure, fast execution, and coordination across heterogeneous systems. In practice, the dominant ecosystems in our measurement window are either proof-of-stake chains, or rollups with centralized sequencing that still rely on highly specialized off-chain infrastructure for transaction production and execution. In both cases, the core mechanism that turns theoretical opportunities into realized MEV is not purely on-chain; it is mediated by off-chain routing, search, and timing.

Empirically, the chapter introduces a conservative detection framework for **sequence dependent arbitrage** over more than two hops, and evaluates it on a year-long multi-chain dataset [79]. The core empirical result is stark and directly relevant for RQ4: multihop sequence-dependent arbitrage exists and can be observed in the wild, but it is *extremely rare*, and its feasibility degrades rapidly with hop count due to compounding bridge latency, costs, and exposure to market drift.

7.1 Why cross-chain MEV is structurally different

Cross-chain MEV differs from single-chain MEV for one main reason: cross-chain execution is typically *non atomic*. On a single chain, many profitable strategies can be executed in an atomic way, either within one transaction, or via bundles that ensure all-or-nothing inclusion. Across chains, strategies must traverse interoperability layers, and the ability to guarantee atomicity across independent ledgers is limited. This changes the feasibility, risk, and the meaning of "realized" MEV.

7.1.1 Fragmentation, interoperability, and the rise of cross-chain arbitrage

The growth of L1s and rollups fragments liquidity and price formation across many AMMs that quote the same economic assets in different execution environments. Bridges reduce user friction by enabling asset transfers and cross-chain communication, but they also create new profit surfaces. When similar assets trade across chains, discrepancies can persist due to heterogeneous latency, fees, congestion, and execution policies, producing arbitrage opportunities that fall under the broader umbrella of cross-chain MEV [131].

From the perspective of realized MEV, fragmentation is a double-edged force. It increases the number of theoretical opportunities, because it creates more venues where prices can diverge. It also strengthens the role of infrastructure, because taking these opportunities requires searching over a larger state space, fast submission over multiple systems, and careful management of latency and cost constraints. In other words, fragmentation expands opportunity surfaces, while cross-chain mechanics constrain feasibility.

7.1.2 Sequence independent versus sequence dependent arbitrage

Following the taxonomy used by Öz et al. [131], we distinguish two cross-chain arbitrage families.

- **Sequence independent arbitrage (SIA).** The arbitrageur holds inventory on multiple chains (inventory arbitrage), and can execute the two legs in either order. This decoupling reduces the role of bridge latency during execution, but requires inventory management and periodic rebalancing.
- **Sequence dependent arbitrage (SDA).** The arbitrageur *bridges* assets such that outputs of earlier steps fund later steps (bridge arbitrage). This enforces an ordered, cross-chain execution, and directly exposes the strategy to bridge latency and bridge fee structures.

Figure 7.1 and Figure 7.2 visualize the two mechanisms. In *SIA*, the arbitrageur can execute legs in either order because inventory is held on both

chains, while in *SDA*, the execution is constrained by bridging, because outputs of earlier steps fund later steps.

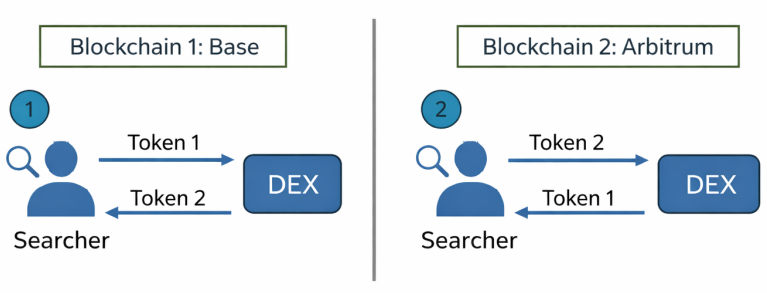


Figure 7.1: Example of *sequence independent arbitrage* (SIA). The arbitrageur holds inventory on both chains, so the two swap legs can be executed without the use of a bridge.

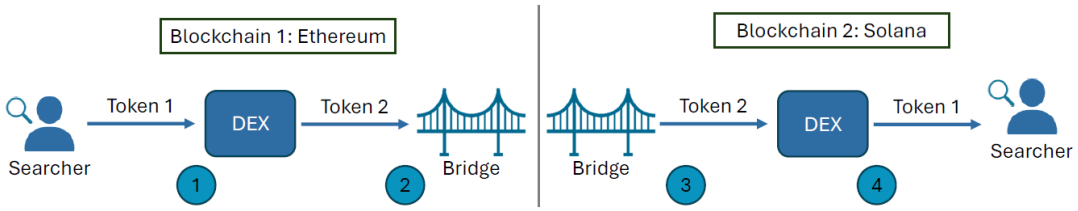


Figure 7.2: Example of *sequence dependent arbitrage* (SDA). The arbitrageur bridges assets so that outputs of earlier steps fund later steps, enforcing an ordered cross-chain execution.

This chapter concentrates on **SDA**, because SDA is exactly where RQ4 becomes binding: bridge time and costs mechanically shape feasibility, execution time, and profitability, and these frictions compound as the number of hops increases.

7.1.3 From two-hop to multihop execution

Most prior empirical work studies cross-chain arbitrage as a two-hop phenomenon, where a swap on one chain is linked to a swap on another chain, optionally connected by a bridging leg [131]. The natural generalization is **multihop SDA**, where an actor traverses more than two chains, repeatedly alternating swaps and bridges. Multihop is not simply "more steps". It compounds three constraints that are central to realized MEV:

- **Latency compounding.** SDA requires waiting for bridges, and each additional bridge stretches the execution window.
- **Cost compounding.** Multiple swaps plus multiple bridge fees increase the break-even discrepancy needed for profit.

- **Risk compounding.** Longer windows increase exposure to price drift, liquidity changes, and operational failures.

The goal of this chapter is therefore to operationalize an explicit multihop SDA model, detect executed instances in large-scale data, and quantify how quickly feasibility collapses as hop count grows.

7.2 Data and measurement scope

We analyze cross-chain activity using Flipside Crypto data over the period **September 1, 2023 to August 31, 2024**, aligned with the time window adopted by Öz et al. [131]. The dataset comprises **2,490,877,965** unique swaps and **34,835,278** bridge transactions across **12** blockchain networks: Ethereum, Arbitrum, Near, Polygon, BSC, Solana, Blast, Osmosis, Avalanche, Optimism, Base, and Gnosis. Bridge coverage spans **45** protocols, including, among others, Wormhole, Axelar, Stargate, Multichain, LayerZero, and Symbiosis [79].

This broad scope matters for RQ4. Multihop paths are, by definition, cross-ecosystem and cannot be reconstructed from a single chain trace. Moreover, the set of chains included in the dataset reflects a modern, infrastructure-rich environment. Many of the chains are proof-of-stake systems, and the rollups rely on specialized sequencers and bridging infrastructure. This is precisely the kind of environment where realized MEV is expected to be facilitated by off-chain services, even when the underlying ledgers remain transparent.

7.3 Multihop SDA model

7.3.1 Path definition

We model an n hop SDA arbitrage as an ordered sequence of $2n-1$ transactions that alternates swaps and bridges:

$$A = \{t_1, t_2, \ldots, t_{2n-1}\},$$

where t_{2i-1} is a swap on chain C_i, and t_{2i} is a bridge from C_i to C_{i+1}, for $i = 1, \ldots, n$

Each transaction t_i is represented as:

$$t_i = (h_i, \tau_i, S_i, R_i, C_i, X_{in}^{(i)}, X_{out}^{(i)}, v_{in}^{(i)}, v_{out}^{(i)}, \alpha_i, \beta_i),$$

where h_i is the transaction hash, τ_i is timestamp, S_i and R_i are sender and receiver, C_i is chain identifier, $X_{in}^{(i)}$ and $X_{out}^{(i)}$ are input and output tokens, $v_{in}^{(i)}$ and $v_{out}^{(i)}$ are the corresponding USD values, and (α_i, β_i) indicates whether the transaction is a swap or a bridge, with $\alpha_i + \beta_i = 1$.

The main challenge in detecting multihop SDA in a billion-scale swap universe is avoiding spurious linkages, naive matching can easily connect unrelated

transactions that happen to be close in time or value. Since the goal of this chapter is a conservative feasibility measurement, we impose a strict interpretability constraint and retain only paths in which each swap produces an *effective token change*. Concretely, we discard sequences where a swap leaves the input and output token unchanged, because such steps do not reflect economic conversion and can arise from wrappers, accounting artifacts, or protocol-specific mechanics. This restriction reduces recall, but it substantially increases precision and yields a small set of high-confidence multihop SDA executions suitable for measuring existence and frequency.

7.3.2 Detection algorithm

For any two consecutive transactions (t_i, t_{i+1}) in a candidate path, we enforce five constraints that implement the sequential logic of SDA.

Constraint 1, time continuity

$$\tau_i < \tau_{i+1} \leq \tau_i + \Delta t,$$

with a conservative choice $\Delta t = 5$ minutes. This window is chosen to make detected sequences temporally plausible while limiting spurious matches. Relaxing Δt to 10 or 15 minutes would increase recall, but at the cost of admitting false positives from coincidental price and value alignment across unrelated transactions. The five minute choice therefore reflects a deliberate trade off in favor of precision over recall, which is appropriate for establishing existence and characterizing true multihop SDA.

Constraint 2, value continuity

$$v_{in}(t_{i+1}) \in [0.98 \cdot v_{out}(t_i),\ v_{out}(t_i)].$$

We allow up to a 2% discrepancy to account for fees and execution approximations. The interval is asymmetric because value tends to decrease across hops due to swap and bridge fees. The 2% threshold is therefore used as a conservative tolerance that captures small losses across linked steps while reducing coincidental matches between unrelated transactions.

Constraint 3, token continuity

$$X_{out}(t_i) = X_{in}(t_{i+1}).$$

Outputs of step i must fund inputs of step $i + 1$.

Constraint 4, actor consistency

Actor identity must remain consistent with swap to bridge and bridge to swap transitions:

$$S_{i+1} = \begin{cases} S_i, & \text{if } t_i \text{ is swap and } t_{i+1} \text{ is bridge,} \\ R_i, & \text{if } t_i \text{ is bridge and } t_{i+1} \text{ is swap.} \end{cases}$$

This captures that a bridge changes the relevant address context, because the receiver on the destination chain becomes the effective controller of the next swap leg.

Constraint 5, cross-chain consistency

Chain identity must behave as expected:

$$C_{i+1} = \begin{cases} C_i, & \text{if } t_i \text{ is swap and } t_{i+1} \text{ is bridge,} \\ \neq C_i, & \text{if } t_i \text{ is bridge and } t_{i+1} \text{ is swap.} \end{cases}$$

Swap to bridge occurs on the same chain, bridge to swap changes chains.

We define an indicator function $\phi(t_i, t_{i+1}) \in \{0, 1\}$ that equals 1 if the pair satisfies constraints 1 to 5, and 0 otherwise. A path A is accepted if all consecutive pairs are valid:

$$\Phi(A) = \prod_{i=1}^{2n-2} \phi(t_i, t_{i+1}) = 1.$$

This produces an explicit and auditable rule for extracting multihop SDA candidates from the full transaction universe.

We compute a deliberately simple notion of **gross profit**:

$$\Pi(A) = v_{out}(t_{2n-1}) - v_{in}(t_1).$$

This captures the realized USD value change along the detected execution. It does not subtract all costs, such as gas on each chain, bridge fees, slippage, and opportunity costs, and this limitation is central in the interpretation of feasibility and profitability, as discussed later.

7.4 Results

The gross profit metric $\Pi(A)$ used throughout this section is computed as the USD value difference between the final output and the initial input of the detected path. It does not subtract per-chain gas costs, bridge protocol fees, slippage on individual swaps, or opportunity costs. Net profitability for individual detected paths should therefore be assumed to be lower than reported, and in some cases negative after costs. This limitation is discussed in detail in Section 7.5.

We recall that **RQ4** asks whether multi-hop, sequence-dependent cross-chain arbitrage is not only theoretically possible but empirically observable at scale, and how feasibility degrades with hop length under real bridge latency and cost frictions.

Applying the detector to the full dataset yields only **10** multihop SDA instances in one year, comprising **8** three-hop and **2** four-hop sequences, with **no** five-hop or six-hop cases [79]. This immediately answers the frequency component of RQ4 for this conservative detector: multihop SDA is empirically

observable, but it is not a common mechanism of realized cross-chain MEV at the scale of billions of swaps.

A useful comparative anchor comes from two-hop measurement work. Öz et al. report a large population of executed cross-chain arbitrages, including SDA connected by bridging, over the same overall period, indicating that cross-chain arbitrage is economically meaningful at the two-hop level [131]. Our multihop result, therefore, isolates a sharp boundary in the feasibility frontier: moving from two-hop to three-plus-hop SDA triggers a dramatic collapse in observed execution frequency.

The multihop study hypothesizes that feasibility declines according to a power law as hop count grows, reflecting compounding frictions. In the observed range, statistical comparisons favor the power law specification over an exponential alternative, with slightly better fit under AIC and RMSE for the power law model [79]. While the sample is necessarily small at higher hop counts, the pattern is consistent with the mechanism implied by RQ4: additional bridges multiply delays and costs, shrinking the set of opportunities that survive long enough to be realized.

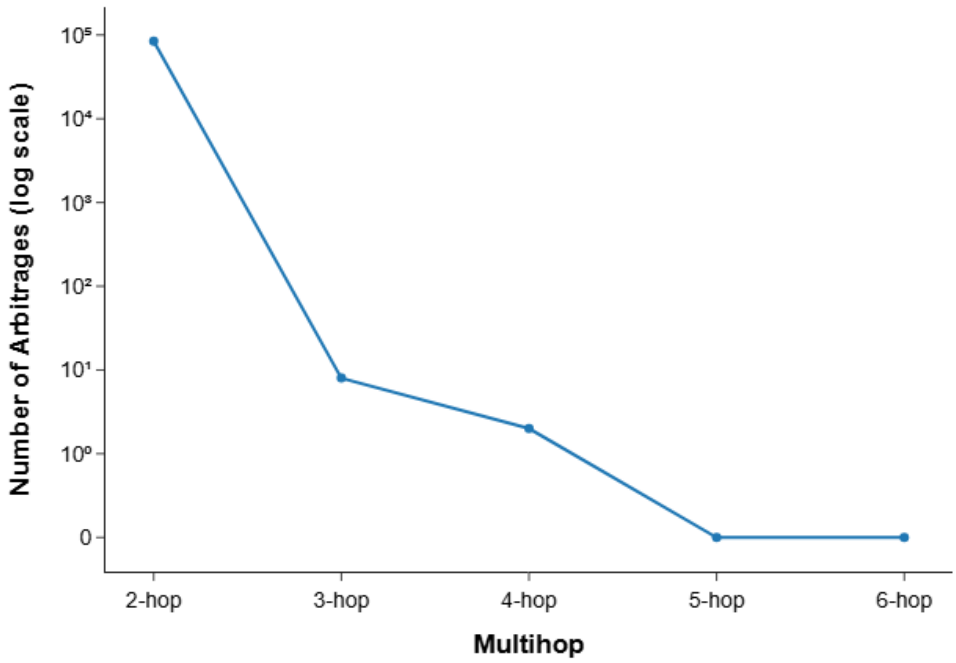


Figure 7.3: Log scale distribution of identified cross-chain SDA arbitrages as a function of multihop complexity. The two-hop reference point is derived from Öz et al. [131], while three-hop and four-hop detections are from our conservative n hop detector [79].

Execution time provides a direct mechanistic link between bridges and feasibility. For the eight detected three-hop paths, the mean duration is approximately **434.1 seconds**, about **7.2 minutes** [79]. For the two four hop paths,

observed durations are **776 seconds** (about 12.9 minutes) and **617 seconds** (about 10.3 minutes) [79]. SDA is sequential; the arbitrageur must wait for bridge completion before executing later legs, and the opportunity can decay during this waiting time due to price convergence, volatility, and competitive pressure.

These durations are not a minor implementation detail. They define the economic environment in which multihop SDA operates. In particular, even when a discrepancy exists at time t_1, the relevant market state at t_{2n-1} can be substantially different, and, therefore, multihop SDA converts a price discrepancy into a bet on persistent mispricing under latency.

Among the ten detected cases, **6** yield positive gross profit. Profits range from small values, for example, on the order of cents or dollars, up to a few hundred USD, while some sequences incur large losses, including an observed three-hop loss of approximately **$255** [79]. Table 7.1 reports the detected three-hop executions, including their chain paths, token sequences, durations, and gross profits. For completeness, Table 7.2 summarizes the two observed four-hop executions, which have longer durations and comparatively small but positive gross profits in this sample.

Chain path	**Dur. (s)**	**Tokens**	**Profit ($)**
Base→Eth→Base	646	USDC/BAL	32.78
Base→Opt→Base	490	HOP/WETH	0.43
Arb→Opt→Base	311	ARB/WETH/HOP	-255.07
Opt→Base→Arb	466	WETH/HOP	264.04
Base→Avax→Base	448	USDC/axlUSDC	21.40
Arb→Poly→Arb	412	USDT/USDC/WBTC	-0.17
Blast→Arb→Eth	458	WETH/ezETH	-8.17
Poly→Arb→Poly	242	WMATIC/WBTC/WETH	-0.02

Table 7.1: Detected three-hop SDA paths, duration, token sequence, and gross profit, using the conservative n hop detector [79].

Chain path	**Dur. (s)**	**Profit ($)**
Optimism→Base→Optimism→Base	776	20.69
Arbitrum→Optimism→Base→Arbitrum	617	1.61

Table 7.2: Detected four hop SDA paths, duration, and gross profit, using the conservative n hop detector [79].

Even with only ten cases, chain participation hints at where multihop SDA becomes feasible. Base appears in **5 out of 8** three hop instances, followed by Arbitrum (**4**) and Optimism (**3**) [79]. These ecosystems combine high activity with relatively low per-transaction costs compared to Ethereum L1, and they are tightly connected by bridges, creating a dense interoperability substrate. This is consistent with the broader cross-chain literature, where rollup ecosystems and

high-throughput environments are emphasized as key sites for realized cross-chain MEV due to connectivity and infrastructure maturity [131].

For the two detected four-hop paths, both are executed by the same address, suggesting that actors capable of multihop SDA may be a small, specialized subset with sufficient infrastructure, monitoring, and operational sophistication [79]. While this is not a strong statistical claim, it aligns with the qualitative view that cross-chain MEV favors vertical integration and specialized operational stacks.

7.5 Discussion

The results provide a direct synthesis for **RQ4**. Cross-chain MEV is facilitated by liquidity fragmentation and by an interoperability layer that connects many chains, but the *realization* of cross-chain value depends on whether execution can complete before the opportunity dissipates. Prior measurement shows that two-hop cross-chain arbitrage can be frequent and economically meaningful [131]. In contrast, the evidence in this chapter indicates that multihop SDA sits beyond a practical feasibility threshold under current latency and cost regimes.

The primary mechanism is **latency compounding**. In SDA, later legs are funded by earlier outputs, so bridge time cannot be bypassed during execution. Each additional hop increases the time window in which prices can converge, and volatility can erase the discrepancy. The observed durations in our detections are measured in minutes rather than seconds, and four-hop executions stretch toward the ten-minute range, which makes the opportunity set fragile by construction. This also clarifies why infrastructure matters. Many of the involved ecosystems are proof-of-stake chains or rollups supported by specialized off-chain routing and execution stacks, which can accelerate *local* submission and inclusion, but cross-chain settlement and bridge finality still impose nontrivial delays that dominate at higher hop counts.

A second mechanism is **cost compounding**. Even when per-hop fees are moderate, multiple swaps plus multiple bridge fees create a cumulative cost burden, raising the break-even discrepancy needed for profitability. This selects against multihop SDA relative to alternatives. Actors can often capture similar economic opportunities through fewer hops, or through sequence-independent arbitrage that relies on inventory and periodic rebalancing rather than sequential bridging. As a result, many potential multihop paths are plausibly arbitraged away before they become executable sequences.

A third mechanism is **risk compounding**. Multihop SDA converts arbitrage into a sequential exposure problem, longer sequences increase exposure to drift across multiple markets, slippage, and operational uncertainty. This is consistent with the mixed profitability distribution observed in the detected cases, including several positive profits but also meaningful losses. Together with the count collapse from two-hop to three-plus-hop, the pattern is consistent with a rapid decline in feasibility as hop count increases, as discussed in the chapter's modeling results.

These interpretations must be read together with the study's conservative design. The detector is intentionally precision-oriented, and its counts should be treated as a lower bound rather than an exhaustive census. In particular, the strict matching constraints and interpretability filters can miss valid strategies that rely on aggregators, multi-path routes, or longer execution windows than the adopted continuity bound. Profit accounting is also incomplete. The gross profit metric $\Pi(A)$ is computed from USD-denominated swap values and does not subtract all costs, including per-chain gas, bridge fees, slippage, and opportunity costs. Finally, coverage and observability remain limited by the underlying data sources. Some bridges and chains can be underrepresented, and private execution or opaque routing can further reduce visibility, again biasing the measurement toward conservative lower bounds.

Within these limits, the empirical conclusion is robust in the sense required by **RQ4**. Using a year-long dataset spanning billions of swaps and tens of millions of bridge transfers across 12 blockchains and 45 bridge protocols, and applying the conservative n hop SDA detector, we observe only 10 realized multihop SDA instances, 8 three hop and 2 four hop, and none beyond [79]. Multihop SDA therefore exists and can be observed in the wild, but it is exceedingly rare at the scale of our dataset. The chapter refines the thesis narrative on realized MEV. Fragmentation expands opportunity surfaces across ecosystems, but the realized boundary is shaped by interoperability frictions. Under current bridge latency and cost regimes, two-hop cross-chain arbitrage can be a meaningful component of MEV, while multihop SDA remains beyond the feasibility frontier for most actors because latency, costs, and risk scale with hop count [79, 131].

8

Conclusions and Future Work

This thesis investigated how maximal extractable value, its enabling infrastructure, and users' routing choices jointly shape the efficiency, distribution, and observability of value extraction in contemporary blockchains, with Ethereum as the main empirical case and cross-chain settings as an extension. Across the thesis, the guiding idea has been that MEV is not only a property of DeFi or smart contract rules, it is also a property of market microstructure, namely how order flow is propagated, aggregated, and sequenced, and of governance, namely which actors acquire discretion over ordering and how that discretion is constrained.

The central contribution is an integrated view that connects, first, measurement of MEV mediation and private routing, second, behavioral adaptation by users and searchers, third, the role of proof of stake with proposer builder separation and relays as an institutional layer, and fourth, the emergence of cross-chain arbitrage as an additional frontier where extraction can be composed across execution domains. Empirically, the thesis combined large-scale transaction-level datasets, on-chain traces and token transfers, and external labeling sources, together with graph-based representations of economic relationships, to quantify prevalence, identify patterns, and test robustness under conservative assumptions.

8.1 Answers to the research questions

RQ1, How prevalent is MEV mediation and private routing, and how does it change across consensus regimes and execution pipelines. The first research question concerned the prevalence of MEV mediation and its relation to private routing, across proof of work and proof of stake, with proposer builder separation. The thesis answered this question by introducing an operational definition of mediation that is consistent with the observable execution pipeline in each regime, and by quantifying the share of blocks and transactions attributable to mediated execution.

The main insight is that the Merge in Ethereum (transition from proof of work to proof of stake) and the subsequent consolidation of proposer builder

separation correspond to a structural change in the execution pipeline. In the post-Merge window, relayed and builder constructed blocks become dominant, and the block production process becomes increasingly mediated by off-chain components that coordinate order flow, block assembly, and delivery. This is not merely an engineering convenience, it is a market structure outcome, because it concentrates the points at which order flow is aggregated and sequenced, and therefore where MEV opportunities can be detected and converted into revenue.

A second insight is methodological. Measuring private routing is unavoidably constrained by what can be observed. The thesis therefore adopted conservative matching strategies that are biased toward undercounting private transactions rather than overcounting them. Under these conservative choices, the observed trends still show that private routing is not an edge case, it is a systematic component of the modern Ethereum execution pipeline. This supports the interpretation that "going private" is not only a response to adversarial public propagation, but also a new normal in which users, wallets, and integrators increasingly treat private submission as a default safety and latency choice.

A third insight links prevalence to roles. The analysis highlighted that mediation is not evenly distributed across participants. Instead, a limited set of relays, builders, and affiliated searcher ecosystems mediate a large share of executed order flow, and this concentration is visible in interaction graphs built from repeated economic relationships. The practical consequence is that changes in mediation are changes in who has effective discretion over sequencing.

RQ2, Ordering and favoritism. The second research question examined whether intra-block ordering systematically favors MEV-related activity, and whether private channel submissions and direct payment to the block builder flows obtain priority beyond what fees alone would imply. The thesis answered this question by analyzing ordering outcomes at the granularity of blocks, relating observed positions and adjacency patterns to MEV labels, and by testing whether the presence of private submission and explicit builder compensation is associated with preferential placement that cannot be explained by standard fee signals.

The main conclusion is that mediated execution under proposer builder separation creates a practical pathway for ordering based on relationships rather than on transparent fee competition. When searchers can attach explicit compensation to builders, either through dedicated payment transactions or in contract-mediated transfers, inclusion and ordering become less dependent on winning an open priority race and more dependent on negotiated guarantees. In this setting, the relevant economic signal is not only the user-paid fee, but the total surplus shared within the off-chain pipeline. This mechanism structurally favors strategies that benefit from precise ordering, because they can purchase reliability directly from the sequencing locus.

A second conclusion is that the empirical evidence is consistent with systematic prioritization of MEV-related flows, especially in blocks produced through relayed and builder-mediated channels. The thesis does not interpret this as a

purely technical artifact, but as an equilibrium of the mediated market, blocks are assembled by actors with strong incentives to internalize off-chain side payments and to select bundles that maximize joint revenue. As a result, malicious or adversarial strategies, such as sandwich attacks, can be executed more reliably, because the same coordination mechanism that reduces uncertainty for benign optimization also reduces uncertainty for harmful extraction.

Finally, the thesis discussed signals that are consistent with coordinated behavior across roles. Repeated patterns of preferential placement, recurring payment relationships, and the persistence of the same counterparties across windows provide circumstantial evidence that some searcher builder validator triplets behave as stable economic coalitions rather than as independent competitors. While direct proof of collusion is not generally observable on chain, the convergence of ordering outcomes and persistent payment links is consistent with a market in which cooperation, exclusivity, or favoritism can arise as rational responses to mediated execution.

RQ3, Social structure and centralization. The third research question shifted from ordering outcomes to the social structure of extraction. It asked which stable groups and interest clusters form within the MEV ecosystem, how PBS-mediated relationships differ from direct transfers, and what degree of centralization characterizes extracting actors. The thesis answered this question by constructing interaction graphs from repeated economic relationships, distinguishing PBS-mediated links from direct value transfers, and then identifying stable clusters that persist across time windows.

The core insight is that the MEV ecosystem is not well described as a flat field of anonymous competitors. Instead, it exhibits a stable structure, a relatively small set of builders, relays, and tightly connected searcher ecosystems that repeatedly interact, and these interactions form clusters that are robust to windowing and to basic perturbations. This stability indicates that extraction is organized around durable relationships, shared infrastructure, and repeated counterparties, rather than being purely opportunistic at the transaction level.

A second insight is that PBS-mediated relationships have distinct structural properties relative to direct transfers. Direct transfers reveal explicit money flow ties, but PBS-mediated execution adds an additional layer where ordering influence can be exercised without leaving the same kind of direct transfer footprint. The thesis showed that these mediated links concentrate around a limited set of hubs, and that these hubs act as coordination points through which a large share of MEV-related activity is routed. This implies that centralization can increase even if on-chain transfers appear diffuse, because the effective discretion over sequencing is mediated by a smaller set of block construction intermediaries.

The practical implication is a form of soft centralization. The protocol remains permissionless, yet the extraction and ordering landscape is shaped by a network of intermediaries whose market shares and relationships concentrate influence. This matters for security and governance because it determines who can consistently observe, simulate, and act on order flow, and it determines how quickly strategies and counter-strategies propagate across the ecosystem.

RQ4, Cross-chain dynamics. The fourth research question extended the analysis beyond a single chain and focused on how MEV manifests across chains, the feasibility and frequency of multi-hop cross-chain arbitrage, and how bridge latency and costs shape profitability and execution time. The thesis answered this question by framing cross-chain extraction as a sequence-level phenomenon, where the relevant unit is a multi-domain action path rather than a single transaction, and by evaluating feasibility constraints induced by bridge design, confirmation rules, and cost structure.

The main finding is that cross-chain MEV is fundamentally a coordination problem under heterogeneous timing. Profitability depends not only on price discrepancies, but also on whether an actor can execute a sequence across domains before the discrepancy closes, while paying bridge fees and bearing finality risk. Bridge latency and costs therefore act as first-order constraints, they bound which opportunities are actionable, they shape the distribution of realized profits, and they determine the typical execution time of successful sequences.

Empirically, the thesis also shows that the multi-hop arbitrages detected in the dataset arise within the Ethereum ecosystem, namely across proof of stake chains and Ethereum-aligned Layer 2 domains. This concentration is informative. It suggests that realized cross-chain extraction is not simply a generic consequence of having many chains, but is strongly shaped by where MEV strategies and infrastructure can reliably operate. In practice, the ability to realize multi-hop opportunities depends on architectures that support predictable execution, low coordination overhead, and integration with the same searcher and builder ecosystems that already specialize in single-chain MEV. In this sense, cross-chain MEV should be interpreted as an extension of established MEV strategies into domains whose execution and bridging architectures make realization feasible, rather than as a uniform phenomenon across all chains.

Finally, we argue that cross-chain MEV intensifies governance and design questions. Extraction can exploit timing mismatches and finality asymmetries across domains, and these asymmetries are shaped by bridge protocols and interoperability choices. As a result, evaluating cross-chain fairness and security requires reasoning not only about ordering within a chain, but also about the cross-domain execution fabric that connects chains into a single economic surface.

8.2 Synthesis of insights

Across these research questions, three synthesis insights emerge.

First, ordering power is increasingly mediated. Under PBS, ordering outcomes can reflect negotiated guarantees and off-chain compensation, and direct payment to the block builder flows provides a concrete mechanism through which MEV-related activity can obtain priority beyond what transparent fee signals would predict.

Second, extraction has a social structure. Stable clusters and repeated economic relationships indicate that the MEV ecosystem contains durable coali-

tions and hubs. PBS-mediated links can concentrate effective discretion even when direct transfers alone would understate centralization, implying that measuring decentralization requires measuring the mediation layer.

Third, MEV is becoming multi-domain. Cross-chain opportunities are shaped by latency, cost, and finality heterogeneity, making sequence-level measurement and modeling essential. Bridge design and interoperability protocols, therefore, become part of the MEV threat model and part of the governance surface.

Regulatory and governance implications. The evidence presented in this thesis raises a question that falls outside the scope of empirical measurement but is nonetheless consequential for protocol design and public policy: *could regulators legitimately exercise oversight over transaction ordering and value extraction in proof-of-stake systems*? The post-Merge PBS architecture creates identifiable, persistent entities, builders, relays, and large staking providers, that operate like financial intermediaries and that, in some jurisdictions, may already fall under existing market-abuse or broker-dealer frameworks [70]. The ESMA report on MEV explicitly calls for continued supervisory attention and characterizes existing countermeasures as "incomplete" [44]. Whether oversight mechanisms can be designed that are compatible with the decentralization goals of permissionless blockchains remains an open governance research question, and one that the empirical record of this thesis directly informs [3].

8.3 Limitations

The thesis adopted measurement strategies that prioritize correctness over maximal recall, especially when inferring private routing, classifying strategies, or attributing roles. This implies that several quantitative results should be read as lower bounds rather than complete enumerations. Moreover, actor classification is inherently imperfect, because identities and operational arrangements can change across time and infrastructures.

Another limitation concerns dependence on external data sources for labels and for parts of the execution pipeline. External sources can be incomplete, temporally limited, or affected by outages. The thesis addressed this by validating results across windows, excluding anomalous periods where appropriate, and checking consistency across multiple signals, but residual uncertainty remains.

Finally, cross-chain analysis remains the most challenging setting. Even with high-quality per-chain data, bridges, off-chain components, and heterogeneous finality create observational gaps. The thesis therefore presents cross-chain analysis as both an empirical contribution where feasible and a roadmap for scalable multi-domain measurement.

8.4 Future work

Several directions follow naturally.

A first direction is a deeper ordering of accountability in mediated pipelines. This includes measurement methods that separate benign optimization from exploitative favoritism, and instrumentation that makes relay and builder selection more transparent without forcing disclosure of sensitive proprietary details.

A second direction is richer social structure measurement. Longitudinal cluster stability, cross-layer graphs that combine direct transfers with mediated execution signals, and robustness analyses could improve confidence about persistent coalitions and the degree of centralization among extracting actors.

A third direction is cross-chain sequence modeling at scale. A unified dataset of swaps, bridges, and traces across multiple chains, combined with sequence-aware detection, would enable systematic measurement of multi-hop arbitrage and would quantify how bridge latency and costs shape realized profitability and execution time.

Bibliography

[1] Aestus. Aestus Documentation Relay. https://aestus.live/, 2022. Accessed on December, 5th, 2023.

[2] Aestus. Aestus Relay. https://mainnet.aestus.live/, 2022. Accessed on December, 5th, 2023.

[3] Raphael Auer, Jon Frost, and Jose María Vidal Pastor. Miners as Intermediaries: Extractable Value and Market Manipulation in Crypto and DeFi. BIS Bulletin 58, Bank for International Settlements, 2022.

[4] Kushal Babel, Philip Daian, Mahimna Kelkar, and Ari Juels. Clockwork Finance: Automated Analysis of Economic Security in Smart Contracts. In *2023 IEEE Symposium on Security and Privacy (SP)*, pages 2499–2516, 2023.

[5] Pierre Bagourd et al. Quantifying MEV on Layer 2 Networks. *arXiv preprint arXiv:2309.00629*, 2023.

[6] Bank for International Settlements. Central Bank Digital Currencies: Foundational Principles and Core Features. Technical report, Bank for International Settlements, 2020. Joint report by BIS and seven central banks.

[7] Bank for International Settlements. The Future of the Monetary System. Annual economic report, chapter 3, Bank for International Settlements, 2022.

[8] Paul Baran. On Distributed Communications. Technical Report RM-3420-PR, RAND Corporation, 1964.

[9] Apolline Blandin, Johannes Cloots, Kilian Hussain, Michel Rauchs, Rasheed Saleuddin, and Bryan Wu. 3rd Global Cryptoasset Benchmarking Study. Technical report, Cambridge Centre for Alternative Finance, 2020.

[10] Blocknative. MEV-Boost & the Blocknative Relay. https://www.blocknative.com/blog/get-started-mev-boost, 2022. Accessed on December, 5th, 2023.

[11] bloXroute. bloXroute Relay Documentation. https://docs.bloxroute.com/apis/mev-solution/mev-relay-for-validators, 2022. Accessed on December, 5th, 2023.

[12] bloXroute Labs. bloXroute Ethical Relay Shutdown. https://twitter.com/bloXrouteLabs/status/1690065892778926080, 2023. Accessed on December, 5th, 2023.

[13] bloXroute Max Profit. bloXroute Max Profit Relay. https://bloxroute.max-profit.blxrbdn.com/, 2022. Accessed on December, 5th, 2023.

[14] bloXroute Regulated. bloXroute Regulated Relay. https://bloxroute.regulated.blxrbdn.com/, 2022. Accessed on December, 5th, 2023.

[15] builder0x69.eth. Relayooor Relay Information. https://twitter.com/builder0x69/status/1585287608858451970, 2022. Accessed on December, 5th, 2023.

[16] builder0x69.eth. Relayooor Shutdown. https://twitter.com/builder0x69/status/1639029339743297537, 2022. Accessed on December, 5th, 2023.

[17] Vitalik Buterin. Ethereum: A Next-Generation Smart Contract and Decentralized Application Platform. https://ethereum.org/en/whitepaper/, 2014. White paper; Accessed 2025-10-06.

[18] Vitalik Buterin. Proof-of-stake: next-generation consensus, explained. https://vitalik.eth.limo/general/2020/11/06/pos2020.html, 2020. Overview of PoS design, finality, weak subjectivity, and slashing.

[19] Vitalik Buterin and Virgil Griffith. Casper the Friendly Finality Gadget. In *Financial Cryptography and Data Security Workshops*, 2017. Preprint version widely circulated in 2017.

[20] Phil Champagne. *The Book of Satoshi: The Collected Writings of Bitcoin Creator Satoshi Nakamoto*. e53 Publishing, 2014.

[21] David Chaum. Blind Signatures for Untraceable Payments. In *Advances in Cryptology, Proceedings of CRYPTO '82*, pages 199–203. Springer, 1983. Accessed on November, 3rd, 2025.

[22] Tianyang Chi, Ningyu He, Xiaohui Hu, and Haoyu Wang. Remeasuring the Arbitrage and Sandwich Attacks of Maximal Extractable Value in Ethereum. *arXiv preprint arXiv:2405.17944*, 2024.

[23] Nicolas Christin. Traveling the Silk Road: A Measurement Analysis of a Large Anonymous Online Marketplace. In *Proceedings of the 22nd International World Wide Web Conference (WWW)*, pages 213–224, 2013.

[24] Coindesk. Is Ethereum's Censorship Problem Taking a Turn? https://www.coindesk.com/tech/2022/12/21/is-ethereums-censorship-problem-taking-a-turn/, 2022. Accessed on December, 5th, 2023.

[25] Coindesk. Blocknative to Suspend MEV-Boost Relay After Economics Fail to 'Materialize'. https://www.coindesk.com/tech/2023/09/26/blocknative-to-suspend-mev-boost-relay-after-economics-fail-to-materialize/, 2023. Accessed on December, 5th, 2023.

[26] CoinGecko. Global Cryptocurrency Market Capitalization. https://www.coingecko.com/, 2024. Time series of aggregate cryptoasset market capitalization; Accessed on November, 14th, 2025.

[27] Eric Conner, Rick Dudley, Ian Norden, Andrew Ashworth, and Vitalik Buterin. EIP-1559: Fee market change for ETH 1.0 chain. https://eips.ethereum.org/EIPS/eip-1559, 2019.

[28] Wei Dai. b money. https://nakamotoinstitute.org/library/b-money/, 1998. Accessed on November, 3rd, 2025.

[29] Philip Daian, Steven Goldfeder, Tyler Kell, Yuan Li, Xueyuan Zhao, Iddo Bentov, Lorenz Breidenbach, and Ari Juels. Flash Boys 2.0: Frontrunning in Decentralized Exchanges, Miner Extractable Value, and Consensus Instability. In *2020 IEEE Symposium on Security and Privacy (SP)*, pages 910–927. IEEE, 2020.

[30] Christian Decker and Roger Wattenhofer. Information Propagation in the Bitcoin Network. In *13th IEEE International Conference on Peer-to-Peer Computing (P2P)*, pages 1–10, 2013.

[31] Gilles Deleuze and Félix Guattari. *A Thousand Plateaus: Capitalism and Schizophrenia*. University of Minnesota Press, 1987. Transl. Brian Massumi; original French 1980.

[32] EigenPhi. MEV Analytics Dashboard. https://eigenphi.io/, 2024. Daily statistics on MEV profits, costs, and attack types; Accessed on November, 13th, 2025.

[33] Shayan Eskandari, Seyedehmahsa Moosavi, and Jeremy Clark. SoK: Transparent Dishonesty: Front-running Attacks on Blockchain. *arXiv preprint arXiv:1902.05164*, 2019.

[34] Ethereum Foundation. The Merge. https://ethereum.org/en/upgrades/merge/, 2022. Official documentation of Ethereum's transition to proof of stake.

[35] Ethereum Foundation. Ethereum Proof-of-Stake Consensus Specifications. https://github.com/ethereum/consensus-specs, 2023. Living specification repository for the Beacon Chain and PoS.

[36] Ethereum Improvement Proposals. EIP-1559: Fee market change for ETH 1.0 chain. https://eips.ethereum.org/EIPS/eip-1559, 2019. Accessed on December, 27th, 2025.

[37] Ethereum.org. Maximal Extractable Value (MEV). https://ethereum.org/en/developers/docs/mev/, 2023. Accessed on December, 5th, 2023.

[38] ethereum.org. Gas and fees. https://ethereum.org/developers/docs/gas/, 2025. Accessed on December, 27th, 2025.

[39] Ethereum.org. Proposer builder separation (PBS). https://ethereum.org/roadmap/pbs/, 2025. Accessed on December, 23rd, 2025.

[40] ethereum.org. Transactions. https://ethereum.org/developers/docs/transactions/, 2025. Accessed on December, 27th, 2025.

[41] Etherscan. Ethereum Block 0 Genesis Block. https://etherscan.io/block/0, 2015. Accessed 04 November 2025.

[42] Etherscan. Etherscan APIs documentation. https://docs.etherscan.io/, 2024. Accessed on December, 5th, 2024.

[43] European Central Bank. Stablecoins, Crypto-Assets and the European Financial System. https://www.ecb.europa.eu/, 2023. ECB publication on crypto-assets and stablecoins.

[44] European Securities and Markets Authority. Maximal Extractable Value and On-Chain Market Integrity. Technical report, ESMA, 2025. ESMA report on MEV and DeFi markets.

[45] Bitcoin Block Explorer. Genesis Block (Block 0) and embedded message. https://www.blockchain.com/explorer/blocks/btc/0, 2009. Access details of the coinbase message.

[46] Hal Finney. Bitcoin and me. https://bitcointalk.org/index.php?topic=155054.0, 2013. Forum post.

[47] Flashbots. Quantifying MEV – Introducing MEV-Explore v0. https://medium.com/flashbots/mev-explore-v0-quantifying-mev-7e43a1d02e3, 2021. Blog post announcing MEV-Explore and early MEV measurements.

[48] Flashbots. Flahbots MEV-Boost - Introduction. https://docs.flashbots.net/flashbots-mev-boost/introduction, 2022. Accessed on December, 5th, 2024.

[49] Flashbots. Flashbots. https://www.flashbots.net/, 2022. Accessed on December, 5th, 2024.

[50] Flashbots. Flashbots Relay. https://boost-relay.flashbots.net/, 2022. Accessed on December, 5th, 2023.

[51] Flashbots. Flashbots Relay API. https://flashbots.github.io/relay-specs/#/, 2022. Accessed on December, 5th, 2023.

[52] Flashbots. MEV-Boost: Merge-Ready Proposer-Builder Separation. https://boost.flashbots.net/, 2022. Description of MEV-Boost and PBS design on Ethereum.

[53] Flashbots. MEV-Boost Risks and Considerations. https://docs.flashbots.net/flashbots-mev-boost/architecture-overview/risks, 2022. Accessed on December, 5th, 2024.

[54] Flashbots. Flashbots MEV-Blocks API. https://blocks.flashbots.net/, 2023. Accessed on December, 5th, 2023.

[55] Flashbots. Mempool Dumpster. https://mempool-dumpster.flashbots.net/index.html, 2024. Accessed on December, 5th, 2024.

[56] Flashbots. MEV-Inspect Overview. https://docs.flashbots.net/flashbots-data/mev-inspect-py/overview, 2024. Accessed on December, 5th, 2024.

[57] Arthur Gervais, Ghassan O. Karame, Karl Wüst, Vasileios Glykantzis, Hubert Ritzdorf, and Srdjan Capkun. On the Security and Performance of Proof of Work Blockchains. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security (CCS)*, pages 3–16, 2016.

[58] William Gibson. *Neuromancer*. Ace Books, 1984.

[59] Sasha Gogol et al. Cross-Rollup MEV: Non-Atomic Arbitrage Across Layer-2 Blockchains. *arXiv preprint arXiv:2406.02172*, 2024.

[60] Alexander Gramlich, Lewis Gudgeon, Ariah Klages-Mundt, Andreea Minca, Daniël Reijsbergen, and Roni Michaly Shneer. Maximal extractable value: Current understanding, categorisation and open research questions. *Journal of Economic Surveys*, 2024. Early view; volume, issue and pages to be added.

[61] Tivas Gupta, Mallesh M Pai, and Max Resnick. The centralizing effects of private order flow on proposer-builder separation, 2023.

[62] Florian Haffke. Technical Analysis of Established Blockchain Systems. Master's thesis, Technical University of Munich, 2017.

[63] Leon Heimbach et al. Non-Atomic Arbitrage in Decentralized Finance. *arXiv preprint arXiv:2401.01622*, 2024.

[64] Lioba Heimbach, Lucianna Kiffer, Christof Ferreira Torres, and Roger Wattenhofer. Ethereum's proposer-builder separation: Promises and realities. In *Proceedings of the 2023 ACM on Internet Measurement Conference*, IMC '23, page 406–420, New York, NY, USA, 2023. Association for Computing Machinery.

[65] Helius. Solana MEV Report: Trends, Insights, and Challenges. https://www.helius.dev/blog/solana-mev-report, 2025. Accessed on December, 27th, 2025.

[66] Maurice Herlihy. Cross-Chain Deals. *Proceedings of the VLDB Endowment*, 12(9):1139–1152, 2019.

[67] Eric Hughes. No Subject. https://mailing-list-archive.cryptoanarchy.wiki/archive/1992/09/fdf9c19e77ec3f1a9bbc6bc19266d565b89d19dbd0ad369f5a2e800af3fc9558/, 1992. Accessed on November, 3rd, 2025.

[68] Eric Hughes. A Cypherpunk's Manifesto. https://cdn.nakamotoinstitute.org/docs/cypherpunk-manifesto.txt, 1993. Accessed on November, 3rd, 2025.

[69] Paul Janicot and Claude Vinyas. Private MEV RPCs: Benchmark Study. arXiv preprint arXiv:2505.19708, 2025.

[70] Yunjia Ji and James Grimmelmann. Regulatory Implications of MEV Mitigations. *Cornell Law Studies Research*, 26(7), 2024. Available at SSRN: https://ssrn.com/abstract=4831566.

[71] Jito Labs. Introducing the First Solana MEV Dashboard. https://www.jito.wtf/blog/introducing-the-first-solana-mev-dashboard/, 2022. Accessed on December, 23rd, 2025.

[72] Jito Labs. Jito Block Engine Expands Access to All Solana MEV Traders. https://www.jito.wtf/blog/jito-block-engine-expands-access-to-all-solana-mev-traders/, 2023. Accessed on December, 23rd, 2025.

[73] Jito Labs. Jito documentation. https://docs.jito.wtf/, 2025. Accessed on December, 23rd, 2025.

[74] Steven Levy. Crypto Rebels. *Wired*, 1993. Accessed on November, 3rd, 2025.

[75] Metzdowd Cryptography Mailing List. Announcement of Bitcoin on the cryptography mailing list. https://www.metzdowd.com/pipermail/cryptography/2008-October/014810.html, 2008. Mailing list archive.

[76] Xingyu Lyu, Mengya Zhang, Xiaokuan Zhang, Jianyu Niu, Yinqian Zhang, and Zhiqiang Lin. An Empirical Study on Ethereum Private Transactions and the Security Implications. *arXiv preprint arXiv:2208.02858*, 2022.

[77] Davide Mancino, Alberto Leporati, Marco Viviani, and Giovanni Denaro. Exploiting Ethereum after "The Merge": The Interplay between PoS and MEV Strategies. In *Proceedings of the Italian Conference on Cybersecurity (ITASEC 2023)*, 2023.

[78] Davide Mancino, Alberto Leporati, Marco Viviani, and Giovanni Denaro. A role and reward analysis in off-chain mechanisms for executing MEV strategies in Ethereum proof-of-stake. *Distrib. Ledger Technol.*, June 2024.

[79] Davide Mancino, Alberto Leporati, Marco Viviani, and Giovanni Denaro. Bunny Hops and Blockchain Stops: Cross-Chain MEV Detection With N-Hops. *arXiv preprint arXiv:2511.17527*, 2025.

[80] Davide Mancino, Alberto Leporati, Marco Viviani, and Giovanni Denaro. Decentralization or favoritism? an analysis of ethereum transactions and maximal extractable value strategies. In *2025 IEEE International Conference on Blockchain and Cryptocurrency (ICBC)*, pages 1–9, 2025.

[81] Manifold. SecureRpc Relay. https://mainnet-relay.securerpc.com/, 2022. Accessed on December, 5th, 2023.

[82] Herbert Marcuse. *One-Dimensional Man*. Beacon Press, 1964.

[83] Huned Materwala, Shraddha M. Naik, Aya Taha, Tala Abdulrahman Abed, and Davor Svetinovic. Maximal extractable value in decentralized finance: Taxonomy, detection, and mitigation. *IEEE Transactions on Services Computing*, 18(6):4386–4407, 2025.

[84] Timothy C. May. The Crypto Anarchist Manifesto. https://groups.csail.mit.edu/mac/classes/6.805/articles/crypto/cypherpunks/may-crypto-manifesto.html, 1988. Accessed on November, 3rd, 2025.

[85] Adi Mazor and Omer Rottenstreich. On the Feasibility of Cross-chain Arbitrage in Decentralized Exchanges. In *16th International Conference on Communication Systems and Networks (COMSNETS 2024)*, pages 543–549, 2024.

[86] Conor McMenamin. Sok: Cross-domain mev, 2023.

[87] MetaMask. MetaMask developer documentation. https://docs.metamask.io/services/reference/ethereum/json-rpc-methods/, 2024. Accessed on December, 5th, 2024.

[88] Satoshi Nakamoto. Bitcoin: A Peer to Peer Electronic Cash System. https://bitcoin.org/bitcoin.pdf, 2008. Accessed on November, 3rd, 2025.

[89] Arvind Narayanan, Joseph Bonneau, Edward Felten, Andrew Miller, and Steven Goldfeder. *Bitcoin and Cryptocurrency Technologies*. Princeton University Press, 2016.

[90] Eden Network. Eden Network Relay. https://relay.edennetwork.io/info, 2022. Accessed on December, 5th, 2023.

[91] Eden Network. Eden Network Relay Documentation. https://www.edennetwork.io/, 2022. Accessed on December, 5th, 2023.

[92] Mathieu Obadia et al. MEV Explore: Quantifying Maximal Extractable Value in Ethereum. https://mev.explore.flashbots.net/, 2021. Analytics dashboard and documentation on MEV.

[93] Stéphane Obadia. Unity is Strength: A Formalization of Cross-Domain Maximal Extractable Value. *arXiv preprint arXiv:2112.01472*, 2021.

[94] Nathaniel Popper. *Digital Gold: Bitcoin and the Inside Story of the Misfits and Millionaires Trying to Reinvent Money*. Harper, 2015.

[95] Kaihua Qin, Liyi Zhou, Anton Afonin, Laura Lazzaretti, and Arthur Gervais. Quantifying Blockchain Extractable Value: How dark is the forest? *arXiv preprint*, arXiv:2101.05511, 2021. First posted January 2021.

[96] Kaihua Qin, Liyi Zhou, and Arthur Gervais. Quantifying Blockchain Extractable Value: How Dark is the Forest? In *2022 IEEE Symposium on Security and Privacy (SP)*, pages 198–214. IEEE, 2022.

[97] A. Ramiro and R. de Queiroz. Cypherpunk. *Internet Policy Review*, 11(2), 2022.

[98] Ameer Rasheed et al. The Evolution of the MEV Ecosystem in Ethereum. arXiv preprint arXiv:2404.07986, 2024. Preprint on the evolution of MEV infrastructure.

[99] Christian Raun, Daniel Golovnev, Lukas Schmid, and Roger Wattenhofer. Leveraging Machine Learning for Bidding Strategies in MEV Auctions. https://eprint.iacr.org/2023/1281, 2023. Accessed on December, 23rd, 2025.

[100] Agnostic Relay. Agnostic Relay. https://agnostic-relay.net/, 2022. Accessed on December, 5th, 2023.

[101] Agnostic Relay. Agnostic Relay Documentation. https://www.gnosis.io/blog/agnostic-relay-a-credibly-neutral-tool, 2022. Accessed on December, 5th, 2023.

[102] r/ethereum user community. Miners Front-running? https://www.reddit.com/r/ethereum/comments/2d84yv/miners_frontrunning/, 2014. Reddit discussion on potential miner frontrunning in early Ethereum.

[103] Dan Robinson and Georgios Konstantopoulos. Ethereum is a Dark Forest. https://www.paradigm.xyz/2020/08/ethereum-is-a-dark-forest, 2020. Accessed on December, 23rd, 2025.

[104] Tim Roughgarden. Transaction Fee Mechanism Design for the Ethereum Blockchain: An Economic Analysis of EIP-1559. https://www.cs.columbia.edu/~roughgarden/papers/eip1559.pdf, 2020.

[105] Caspar Schwarz-Schilling, Fahad Saleh, Thomas Thiery, Jennifer Pan, Nihar Shah, and Barnabé Monnot. Time is Money: Strategic Timing Games in Proof-of-Stake Protocols. *arXiv preprint arXiv:2305.09032*, 2023.

[106] SecureRPC. SecureRpc Relay Documentation. https://securerpc.com/, 2022. Accessed on December, 5th, 2023.

[107] Kristin Kåsa Sjursen, Hoang Anh Meng, and James Chiu. Towards Quantifying Cross-Domain Maximal Extractable Value for Blockchain Decentralisation. In *Information and Communications Security*, Lecture Notes in Computer Science, 2023.

[108] Solana. Solana developer documentation. https://solana.com/docs, 2025. Accessed on December, 23rd, 2025.

[109] Solana. Solana Status, Incident History. https://status.solana.com/history, 2025. Accessed on December, 5th, 2025.

[110] Solscan. Solscan documentation. https://docs.solscan.io/, 2025. Accessed on December, 23rd, 2025.

[111] Solscan. Solscan Pro API documentation. https://pro-api.solscan.io/pro-api-docs/v2.0, 2025. Accessed on December, 23rd, 2025.

[112] Neal Stephenson. *Snow Crash*. Bantam Books, 1992.

[113] Bruce Sterling. *Mirrorshades: The Cyberpunk Anthology*. Arbor House, 1986.

[114] Nick Szabo. Formalizing and Securing Relationships on Public Networks. *First Monday*, 2(9), 1997. Accessed on November, 3rd, 2025.

[115] Nick Szabo. Bit gold. https://unenumerated.blogspot.com/2005/12/bit-gold.html, 2005. Accessed on November, 3rd, 2025.

[116] Titan. Titan Global Relay. https://global.titanrelay.xyz/, 2024. Accessed on December, 4th, 2024.

[117] Christof Ferreira Torres, Ramiro Camino, and Radu State. Frontrunner Jones and the Raiders of the Dark Forest: An Empirical Study of Frontrunning on the Ethereum Blockchain. In *30th USENIX Security Symposium (USENIX Security 21)*, pages 1343–1359, 2021.

[118] Christof Ferreira Torres et al. Rolling in the Shadows: An Empirical Study of MEV on Ethereum Layer-2 Rollups. *arXiv preprint arXiv:2405.00138*, 2024.

[119] V. A. Traag, L. Waltman, and N. J. van Eck. From Louvain to Leiden: Guaranteeing well-connected communities. *Scientific Reports*, 9(1):5233, 2019.

[120] TripleA. Cryptocurrency Ownership Data 2024. https://triple-a.io/crypto-ownership/, 2024. Accessed on November, 14th, 2025.

[121] Ultrasound. Ultrasound Relay. https://relay.ultrasound.money/, 2022. Accessed on December, 5th, 2023.

[122] Lukas Wahrstätter, Monnef Marcus, et al. Time to Bribe: Measuring Block Construction Markets. *arXiv preprint arXiv:2305.16468*, 2023.

[123] Ye Wang, Yan Chen, Haotian Wu, Liyi Zhou, Shuiguang Deng, and Roger Wattenhofer. Cyclic Arbitrage in Decentralized Exchanges. *arXiv preprint arXiv:2105.02784*, 2022.

[124] Gavin Wood. Ethereum: A Secure Decentralised Generalised Transaction Ledger (Yellow Paper). Technical report, Ethereum Foundation, 2014. Accessed 2025-10-06.

[125] Anatoly Yakovenko. Solana: A new architecture for a high performance blockchain (Proof of History). https://solana.com/solana-whitepaper.pdf, 2018. White paper; Accessed 2025-10-06.

[126] Sen Yang, Fan Zhang, Ken Huang, Xi Chen, Youwei Yang, and Feng Zhu. Sok: Mev countermeasures. In *Proceedings of the Workshop on Decentralized Finance and Security*, DeFi '24, page 21–30, New York, NY, USA, 2024. Association for Computing Machinery.

[127] Alexei Zamyatin, Mustafa Al-Bassam, Dionysis Zindros, Eleftherios Kokoris-Kogias, Pedro Moreno-Sanchez, Aggelos Kiayias, and William J.

Knottenbelt. SoK: Communication Across Distributed Ledgers. In *Financial Cryptography and Data Security*, volume 12675 of *Lecture Notes in Computer Science*, pages 3–36, 2021.

[128] Zeromev. Zeromev. https://zeromev.org/, 2024. Accessed on December, 5th, 2024.

[129] Peilin Zheng, Zibin Zheng, Jiajing Wu, and Hong-Ning Dai. XBlock-ETH: Extracting and Exploring Blockchain Data From Ethereum. *IEEE Open Journal of the Computer Society*, 1:95–106, 2020.

[130] Liyi Zhou, Kaihua Qin, Christof Ferreira Torres, Duc V Le, and Arthur Gervais. High-frequency trading on decentralized on-chain exchanges. In *2021 IEEE Symposium on Security and Privacy (SP)*, pages 428–445, 2021.

[131] Burak Öz, Christof Ferreira Torres, Christoph Schlegel, Bruno Mazorra, Jonas Gebele, Filip Rezabek, and Florian Matthes. Cross-chain arbitrage: The next frontier of mev in decentralized finance, 2025.

La borsa di dottorato cofinanziata con risorse dell'Unione europea-*NextGeneration EU*
Piano Nazionale di Ripresa e Resilienza Missione 4 – Componente 1 – Riforma 4.1 Riforma dei Dottorati - Inv. 4.1
Borse Generici Ricerca PNRR – CUP J11J22002120006