Searcher Competition in Block Building

Akaki Mamageishvili[1]
Offchain Labs, Zurich, Switzerland

Christoph Schlegel
Flashbots, George Town, Cayman Islands

Benny Sudakov
ETH Zürich, Switzerland

Abstract

We study the amount of maximal extractable value (MEV) captured by validators, as a function of searcher (or order flow provider) competition in blockchains with competitive block building markets such as Ethereum. We argue that the core is a suitable solution concept in this context that makes robust predictions that are independent of implementation details or specific mechanisms chosen. We characterize how much value validators extract in the core and quantify the surplus share of validators as a function of searcher competition. Searchers can obtain at most the marginal value increase of the winning block relative to the best block that can be built without their bundles. Dually this gives a lower bound on the value extracted by the validator. If arbitrages are easy to find and many searchers find similar bundles, the validator gets paid all value almost surely, while searchers can capture most value if there is little searcher competition per arbitrage. For the case of passive block-proposers we study, moreover, mechanisms that implement core allocations in dominant strategies and find that for submodular value, there is a unique dominant-strategy incentive compatible core-selecting mechanism that gives each searcher exactly their marginal value contribution to the winning block.

2012 ACM Subject Classification Theory of computation → Algorithmic game theory and mechanism design

Keywords and phrases MEV, Block Building, Searchers, Proposer Builder Separation, Core

Digital Object Identifier 10.4230/LIPIcs.AFT.2024.21

Related Version *Full Version*: https://arxiv.org/abs/2407.07474

1 Introduction

Blockchains that support smart contracts frequently run decentralized finance (DeFi) applications. This in turn gives rise to the phenomenon of miner/maximal extractable value, [5]: Blockchain protocols give *validators*, sometimes called proposers, the right to order transactions for a particular block. This effectively means that validators have a local monopoly to include or exclude transactions or to order transactions in a particular way, in order to generate value for themselves from this privileged position. However, since extracting value from transaction ordering in an optimal way is a difficult task, more specialized actors such as *block builders* and (arbitrage) *searchers* participate in the value extraction process in smart contract blockchains such as Ethereum. Block builders aggregate different arbitrage opportunities, liquidations or "sandwiches" in one block, using transactions from the public mempool, different private mempools, order flow auctions or other transaction sources together with their own transactions. Then, they bid against other builders, to get their block published in a canonical chain of blocks. The bid is paid to the current block proposer. Arbitrage opportunities are typically found by more specialized players, *searchers* and passed

[1] Corresponding author

to the builders. The resulting strategic interaction between searchers, builders and the validator lead to a value distribution of the arbitrage gains between these players. Empirical evidence shows that the majority of the total observable MEV is captured by validators. [2]

In this paper, we analyze the competition between searchers (more generally order-flow providers) in order to explain the value distribution between the proposer (or block builder) and the searchers. Different searchers may capture different value from different opportunities and may or may not find unique opportunities that are not found by competing searchers. This competition and specialization should in turn explain the value capture by different searchers in the MEV supply chain. Our model abstracts away from the intermediate layer of block building where block builders aggregate searcher bundles to blocks and compete in a bidding procedure. However, the ability for different combination of searchers or order flow providers to jointly generate different blocks is captured implicitly by the solution concept we focus on: We consider core allocation where different combinations of players can construct blocks of different values and distribute the value among themselves. The core constraints require that no coalition of players should be able to generate more value for themselves (produce a more valuable block) than they get paid in the realized allocation. Thus, while we generically talk about searchers, our model can for example also capture the case where some searchers act as block builders. The core is a natural solution concept in the context of block building and MEV capture as it captures the competition and possibility of collaboration of different players while abstracting away from any particular mechanisms that would intermediate between searchers, builders and proposers. Thus, it provides robust predictions on the set of plausible value distributions that can arise in any sufficiently competitive block building market. As a next step, one can then look into mechanism that implement particular core allocations through elicitating privately held information of searchers about the value they can generate from different blocks.

Our first result, shows that searchers can at most obtain their marginal contribution to the winning block (the difference in value between the best block that can be built with their transactions and without their transactions) and if value is submodular, then giving each searcher their marginal contribution is in the core.[3] A straightforward calculation shows that in a world with passive block producers, this particular core allocation coincides with the Vickrey-Clark-Growes (VCG) outcome ([11]) and hence can be implemented in dominant strategies (Corollary 5). On the other hand (Proposition 6), no other core-selecting mechanism is dominant-strategy incentive compatible. These result gives additional justification to further study the searcher-optimal point in the core: the value obtained by the validator in this point can be interpreted as the maximal extractable value he can obtain taking into account information rents captured by the searchers.

As a next step, we consider a stochastic model of searcher competition to study the value distribution between the searchers and the validator. Different (arbitrage or MEV) opportunities have a fixed probability of being found by a searcher. The more searchers find an opportunity, the more value can be extracted by the validator and the less value can be extracted by the searchers. We show that if the probability of finding an opportunity is bounded from below by $p > \log(n)/n$, where n is the number of searchers actively searching in the strategy, with high probability the validator captures all value in all core allocations.

[2] See `https://www.galaxy.com/insights/perspectives/distribution-of-mev-surplus/`.

[3] While transactions submitted to a builder might exhibit significant complementarities, submodularity can be argued to be a reasonable assumption in a sufficiently consolidated market. Order flow providers with complementary flow have an incentive to integrate to capture more value together. Thus, the realizable value might be submodular in the contributions of the different (consolidated) players.

On the other hand, if the probability of sucessfully searching is very low, $p \in \Theta(1/n)$, then with positive constant probability, searchers will be able to capture all value in the searcher-optimal core allocation.

Related Literature

The phenomenon of miner extractable value has first been documented in [5]. Another early contribution on front-running in decentralized finance is [6]. MEV has been documented by a variety of public dashboards and data sets, see for example `https://mevboost.pics/`, `https://libmev.com/`. Recently, also the magnitude of non-atomic arbitrage has been empirically investigated, where searchers realize one lag of a trade on chain and one on a different domain, see e.g.[8] and `https://dune.com/flashbots/lol`.

The block building market structure that has evolved in Ethereum since the change to proof of stake, has been the topic of several recent contributions: The dashboard `https://orderflow.art/` documents empirically the supply chain through which transaction requests land on chain. [1] studies whether MEV and Proof-of-Stake rewards capture leads to centralization, discussing both validator and builder roles in the market formation. [12] argue that in the current market structure searchers and builders have an incentive to vertically integrate. [4] argue that the builder is prone to centralization. [15] provide descriptive statistics on the level of decentralization on the builder landscape. [3] proposes a dynamic MEV sharing mechanism that the authors argue results into better decentralization and fair allocation.

[2] studies questions of implementation with active block producers. Our model is similar to theirs, but we focus on the question of implied value distribution. Our positive results for dominant-strategy incentive compatible core-selecting mechanisms for passive block builders complement their negative result for active block builders. [2] also study the role of searchers as intermediaries.

The notion of the core was first formally defined in [7]. For an overview of results around the core and submodularity see e.g. [9].

2 Model

There is a finite set $\mathcal{S}$ of searchers that submit transactions for inclusion in the block. Similarly as in [2] we will usually identify searchers with (bundles of) transactions they have sent for inclusion. However, our model also allows for the interpretation that the same searcher (address) sends multiple bundles for inclusion. There is one validator (proposer), denoted by V. For each set of searchers $A \subseteq \mathcal{S}$ there is a finite set of feasible blocks $\mathcal{B}(A)$ that can be built from bundles of transactions submitted by searchers in A. A searcher i generates value $v_i(B)$ from a block B and the validator generates value $v_V(B)$ from block B. Our model can capture externalities (searcher i's realized value may not only depend on her included transactions but also other transactions in the realized block) and active validators/block producers (we may have $v_V(B) \neq 0$). We assume that utility is transferable and the final utility realized by searcher i if block B is realized and she makes a payment of p_i (e.g. to the validator) is $v_i(B) - p_i$.

Since utility is transferable, we can define a coalitional value function $v : 2^{\mathcal{S} \cup \{V\}} \to \mathbb{R}_+$ by

$$v(S \cup V) := \max_{B \in \mathcal{B}(S)} \left(\sum_{i \in S} v_i(B) + v_V(B) \right),$$

and

$$v(S) = 0 \text{ if } V \notin S,$$

i.e. in case the validator is part of the coalition, the coalitional value is the value of the total welfare maximizing block consisting of transactions from searchers in the coalition, and in case the validator is not part of the coalition, no value can be generated, as the validator is necessary to realize a block. It will be useful subsequently to introduce the short hand notation

$$\bar{v}(S) := v(S \cup V)$$

for $S \subseteq \mathcal{S}$ to denote the value that searchers S can generate together with the validator. Observe that by construction, the (collective) value function is monotonic,

$$\bar{v}(A) \leq \bar{v}(B) \quad \text{for all } A \subseteq B \subseteq \mathcal{S};$$

if we receive more bundles to build a block that will increase welfare weakly, since we can always discard submitted bundles when building a block. Moreover, we make the following assumption on the (collective) value function:

Submodularity. Let $A, B \subseteq \mathcal{S}$. Then

$$\bar{v}(A) + \bar{v}(B) \geq \bar{v}(A \cup B) + \bar{v}(A \cap B),$$

Submodularity states that the value of a block we can build from transactions from searchers in A and B, is bounded by subtracting the value of a block we can build from transactions in both A and B from the sum of values we can achieve from building a block with transactions in A and a block with transactions in B.

Submodularity requires that there are not-too-strong complementarities between different submitted bundles. We can justify this assumption in two ways: first, it may be that complementarities are not strong and different MEV opportunities provide value that is mostly independent from other opportunities. Second, it may be that the complementarities are already absorbed by searchers, e.g. in the sense that searchers who provide complementary flow have an incentive to integrate their operations and send their flow together to extract more value. It is worth noting that our upper bound on searcher values holds also for non-submodular value functions, but may be loose in that case.

It is easy to show that for monotone value functions, submodularity is equivalent to requiring decreasing marginal value:

Decreasing Marginal Value. Let $A \subseteq B \subseteq \mathcal{S}$ and $a \in A$. Then

$$\bar{v}(B) - \bar{v}(B \setminus \{a\}) \leq \bar{v}(A) - \bar{v}(A \setminus \{a\}).$$

A direct consequence of submodularity is the following lemma which will be useful subsequently:

▶ **Lemma 1.** *Let $A \subseteq B \subseteq \mathcal{S}$. Then*

$$\bar{v}(B) - \bar{v}(A) \geq \sum_{i \in B \setminus A} (\bar{v}(B) - \bar{v}(B \setminus \{i\})).$$

Proof. We prove the result by induction on $N := |B \setminus A|$. For $N = 0$ the result holds trivially. Now suppose the result hold for $N \geq 0$ and consider the case $N+1$. Let $j \in B \setminus A$. By induction assumption

$$\bar{v}(B \setminus \{j\}) - \bar{v}(A) \geq \sum_{i \in B \setminus (A \cup \{j\})} (\bar{v}(B \setminus \{j\}) - \bar{v}(B \setminus \{i, j\})).$$

Adding $\bar{v}(B) - \bar{v}(B \setminus \{j\})$ on both sides and using submodularity (which implies decreasing marginal value), we obtain

$$\begin{aligned}\bar{v}(B) - \bar{v}(A) &\geq \sum_{i \in B \setminus (A \cup \{j\})} (\bar{v}(B \setminus \{j\}) - \bar{v}(B \setminus \{i, j\})) + \bar{v}(B) - \bar{v}(B \setminus \{j\}) \\ &\geq \sum_{i \in B \setminus (A \cup \{j\})} (\bar{v}(B) - \bar{v}(B \setminus \{i\})).\end{aligned}$$ ◀

An **allocation** is a value distribution $x : \mathcal{S} \cup V \to \mathbb{R}_+$ such that

$$\sum_{j \in \mathcal{S} \cup V} x_j \leq v(\mathcal{S} \cup V).$$

An allocation is in the **core** if the following inequality

$$\sum_{j \in C} x_j \geq v(C) \tag{1}$$

holds for any subset $C \subseteq \mathcal{S} \cup V$ and all value is distributed:

$$\sum_{j \in V \cup \mathcal{S}} x_j = v(\mathcal{S} \cup V).$$

▶ Remark 2. As usual in the formulation of the core, the solution proposes a value allocation without specifying an explicit implementation through a block and payments between the searchers and the validator. Let $B^* \in \mathcal{B}(\mathcal{S})$ be a welfare maximizing block, i.e. $\sum_{i \in \mathcal{S} \cup V} v_i(B^*) = v(\mathcal{S} \cup V)$. Then a core value allocation x can be implemented by realizing the block B^* and requiring that each individual searcher i makes a payment of $p_i = v_i(B^*) - x_i$ to the validator.

Immediately from the requirement that without the validator no value can be realized, it follows that the validator getting all gains is in the core.

▶ **Observation 3.** *The core is always non-empty. The allocation where $x_V = v(\mathcal{S} \cup V)$ and $x_i = 0$ for each $i \in \mathcal{S}$ is in the core.*

3 Analysis

Our first main result states each searcher can at most capture their marginal contribution to the realized block, and any allocation that gives each searcher at most their marginal contribution is in the core. In particular, this implies that there is a searcher optimal core allocation (giving each searcher exactly their marginal contribution to the realized block) and a validator optimal allocation (giving the validator all realized value).

▶ **Proposition 4.** *An allocation is in the core if and only*

$$0 \leq x_i \leq \bar{v}(\mathcal{S}) - \bar{v}(\mathcal{S} \setminus \{i\})$$

for each $i \in \mathcal{S}$ and $x_V = \bar{v}(\mathcal{S}) - \sum_{i \in \mathcal{S}} x_i$.

Proof. First we show that $x_i \leq \bar{v}(\mathcal{S}) - \bar{v}(\mathcal{S} \setminus \{i\})$ for each i. Suppose not. In that case

$$x_V + \sum_{j \in \mathcal{S} \setminus \{i\}} x_j = \bar{v}(\mathcal{S}) - x_i < \bar{v}(\mathcal{S} \setminus \{i\}).$$

The previous inequality contradicts the core-stability of x which requires inequality (1) to hold for $C = \mathcal{S} \setminus \{i\} \cup V$. The lower bound on x_i is trivial. It remains to show that all allocations with $0 \leq x_i \leq \bar{v}(\mathcal{S}) - \bar{v}(\mathcal{S} \setminus \{i\})$ are in the core. Let x be a vector satisfying these inequalities. Let $A \subseteq \mathcal{S}$ and observe that

$$x_V + \sum_{j \in A} x_j = \bar{v}(\mathcal{S}) - \sum_{i \in \mathcal{S} \setminus A} x_i \geq \bar{v}(A),$$

where the last inequality follows from Lemma 1. Thus, the core inequalities (1) are satisfied for all coalitions $A \cup V$ with $A \subseteq \mathcal{S}$. For coalitions without the validator the core inequalities are trivially satisfied, as all agents get non-negative value in x. ◀

Implementation

In reality, the values that searchers obtain from different blocks is private information to them. However, in a world with passive block producers[4], we can implement the extreme point in the core that gives maximal value to searchers in dominant strategies: Consider the case where the block producer is passive i.e. $v_V(B) = 0$ for each block B. Observe that VCG-payments in this problem are defined by

$$p_i := \max_{B \in \mathcal{B}(\mathcal{S} \setminus \{i\})} \sum_{j \neq i} v_j(B) - \sum_{j \neq i} v_j(B^*) = \bar{v}(\mathcal{S} \setminus \{i\}) - \sum_{j \neq i} v_j(B^*),$$

for each $i \in \mathcal{S}$ where B^* is the welfare-optimal block that can be produced. A straightforward calculation shows that,

$$v_i(B^*) - p_i = \sum_{j \in \mathcal{S}} v_j(B^*) - \bar{v}(B \setminus \{i\}) = \bar{v}(\mathcal{S}) - \bar{v}(\mathcal{S} \setminus \{i\}),$$

i.e. the searcher-optimal core outcome coincides with the VCG outcome. We obtain the following corollary of Proposition 4:

▶ **Corollary 5.** *Under submodular value and with passive block-producers, the searcher-optimal core-outcome can be implemented in dominant strategies.*

On the other hand, it is straightforward to see that other selections from the core that not always select the VCG outcome are not dominant-strategy incentive compatible.

▶ **Proposition 6.** *Under submodular value and with passive block-producers, any core-selecting mechanism that is not always choosing the searcher-optimal outcome in the core is not dominant-strategy incentive compatible.*

---

[4] We know from [2] that with active block producers implementation of non-trivial solutions is not possible. This in turn resembles previous results from other context where negative results prevail if incomplete information in a two-sided markets is on both sides of the market [10,]. In particular, for assignment games [14], which can be re-interpreted as the special case of our model where the validator has additively separable value, there is no mechanism that implements a core-allocation in dominant strategies, for any domain of valuations when there is at least one profile of valuations for which a core allocation that gives positive value to some searcher exists. For a proof of this "folk theorem" see e.g. [13].

Proof. Suppose for reported value functions $(v_i)_{i\in\mathcal{S}}$, a block B^* and payments $(p_i)_{i\in\mathcal{S}}$ are chosen by the mechanism. By core-stability and Proposition 4, we have $0 \leq v_i(B^*) - p_i \leq v(\mathcal{S}) - v(\mathcal{S}\setminus\{i\})$ for searcher i, or equivalently

$$v_i(B^*) \geq p_i \geq \bar{v}(\mathcal{S}\setminus\{i\}) - \sum_{j\neq i} v_j(B^*)$$

Suppose for the sake of contradiction that the last inequality is strict and i reports different values $\tilde{v}_i$ with $\tilde{v}_i(B) = v_i(B)$ for B blocks not including transactions by i, and $\tilde{v}_i(B) \leq \tilde{v}_i(B^*)$ for blocks including transactions by i (so that B^* is still optimal) and

$$p_i > \tilde{v}_i(B^*) > \bar{v}(\mathcal{S}\setminus i) - \sum_{j\neq i} v_j(B^*).$$

Note that this change in value preserves the submodularity of the coalitional value function that the same block B^* is optimal and that the payment $\tilde{p}_i$ for searcher i now satisfies

$$p_i > \tilde{v}_i(B^*) \geq \tilde{p}_i \geq \bar{v}(\mathcal{S}\setminus i) - \sum_{j\neq i} v_j(B^*),$$

which is stricly less. Therefore i gains from misreporting. ◀

The corollary and previous proposition motivate to further study the searcher-optimal point in the core. The value that a (passive) builder/validator obtains in this point is the maximal extractable value taken into account information rents that searchers can capture. In the next sections, we study the particular case where the value of a block is derived from independent (MEV) opportunities for which different searchers compete. For that case, we derive results on when competition lets the core collapses to one point (in which the validator captures all value) and when lack of competition allows searchers to generate positive value in the searcher-optimal core allocation.

4 Independent Bundles and Competing Searchers

In this section, we consider the special case of additively separable value where the value of a block is the sum of values derived from the individual (bundles) of transactions. Moreover, we look at a scenario where multiple searchers may compete for the same (arbitrage) opportunities so that their submitted bundles possibly "clash" with bundles submitted by other searchers.[5] We denote opportunities by $\mathcal{A}$ and now can identify a block by a matrix $B = (B_{ij})_{i\in\mathcal{A},j\in\mathcal{S}}$ where $B_{ij} = 1$ if searcher j's bundle competing for opportunity i is included, $B_{ij} = 0$ if it is not included. We require that $\sum_{j\in\mathcal{S}} B_{ij} \leq 1$ so that multiple clashing bundles cannot be included in the same block. We can then write searcher j's value from block B as

$$v_j(B) = \sum_{i\in\mathcal{A}} v_{ij} B_{ij},$$

where v_{ij} is the value extracted by searcher j from opportunity i. We can add additional constraints such as a capacity constraint on the total number of bundles. In the unconstrained case, where all blocks are feasible, we have

$$v(S\cup V) = \max_{B\in\mathcal{B}(S)} \sum_{i\in\mathcal{A}} \sum_{j\in\mathcal{S}} v_{ij} B_{ij} = \sum_{i\in\mathcal{A}} \max_{j\in S} v_{ij}. \tag{2}$$

[5] This matches the reality of searching where often different searchers compete in the same strategy and find conflicting bundles among which the block builder chooses the most profitable one and includes it in the block while discarding the less profitable competing bundles. Around one third of submitted bundles to Ethereum block builders "clash".

To get an intuition how the core looks like in this case, we start this section with few examples and observations. Previously, in Observation 3 we had observed that giving all value to the validator is always in the core. In the opposite direction, if there is no searcher competition, i.e. no two searchers find the same opportunity, the validator can receive no value in the core.

▶ **Observation 7.** *There are examples of core allocations where the validator receives* 0.

Proof. Suppose that for each opportunity $i \in \mathcal{A}$ there is at most one searcher submitting a bundle of positive value, $v_{ij} > 0$ for at most one j. Then, the value allocation with $x_j = \sum_{i:v_{ij}>0} v_{ij}$ for each searcher j and $x_V = 0$ is in the core. ◀

Next, we give an example where the maximum payment to the validator is enforced in the core.

▶ **Observation 8.** *There are examples where the validator gets the full value of the winning block in any core allocation.*

Proof. Suppose for each opportunity $i \in \mathcal{A}$ at least two searchers submit the same highest value bundle (or no searcher finds the opportunity). Then for each searcher j we have $x_j = v(\mathcal{S}) - v(\mathcal{S} \setminus \{j\}) = 0$. The claim follows from Proposition 4. ◀

For a vector of non-negative numbers X, let $SH(X)$ denote the second highest coordinate of it. Then, let M denote the following sum:

$$M := \sum_{i \in \mathcal{A}} M_i := \sum_{i \in \mathcal{A}} SH((v_{ij})_{j \in \mathcal{S}}). \tag{3}$$

Additive value function as defined in (2) are submodular. Thus, we obtain the following special case of Proposition 4.

▶ **Corollary 9.** *The allocation in which the validator receives $x_V = M$ and searcher j receives $x_j = \sum_{i:j \in argmax_{j\in\mathcal{S}} v_{ij}} (v_{ij} - M_i)$ is in the core.*

This particular core allocation is the worst for the validator and the best for searchers and can be implemented, see Corollary 5, in dominant strategies by a generalized second price auction for bundles.

Stochastic Model

In this section, we analyze the core when searchers find bundles with some probability and success is independent across searchers and opportunities. Let v_{ij} be a binary random variable, which is 1 with probability p and 0 with probability $1 - p$.[6] Thus, p measures how easy it is to find an arbitrage (bundle).

▶ **Proposition 10.** *Let $n := |\mathcal{S}|$. If $p > \frac{2 \log n}{n}$ and $m := |A| < n$, then the validator receives the entire block value with high probability in any core allocation.*

[6] Generalizations of the subsequent results to non unit value that can be different for different opportunities are straightforward. The only assumption needed is that searchers conditional on finding the same opportunity generate the same value from it. We could also accommodate heterogeneous value from the same opportunity as long as the noise is sufficiently bounded.

Proof. We show that there are at least 2 searchers who have positive value each arbitrage. This can be done using direct computation of probabilities and an application of the union bound inequality. Consider the following sum $Y_i := \sum_{j \in \mathcal{S}} v_{ij}$. Thus, Y_i is the number of searchers that find opportunity i. Note that Y_i is a Binomial random variable with parameters n and p, that is $Y_i \sim Bin(n,p)$.

$$\begin{aligned} P[Y_i < 2] &= P[Y_i = 0] + P[Y_i = 1] = (1-p)^n + \binom{n}{1}(1-p)^{n-1}p \\ &\leq e^{-pn} + npe^{-p(n-1)} \leq \frac{1}{n^2} + 2\log n \frac{1}{n^2} \leq \frac{2\log n}{n^2}, \end{aligned} \tag{4}$$

where the first inequality is obtained from the well known inequality: $1 - x \leq e^{-x}$ for any $x > 0$. By the union bound, we have:

$$P[\text{at least one } Y_i < 2] \leq mP[Y_1 < 2] \leq n \cdot \frac{\log n}{n^2} = \frac{\log n}{n}.$$

The last inequality is by (4). For any $\varepsilon > 0$ there is a n large enough so that $\frac{\log n}{n} < \varepsilon$ and therefore

$$P[Y_i \geq 2 \text{ for any } i] \geq 1 - \varepsilon.$$

Applying Proposition 4 shows the claim of the proposition. ◀

On the other hand, if the probability of discovering opportunities shrinks sufficiently fast in the number of searchers, then searchers can capture value with positive non-vanishing probability:

▶ **Proposition 11.** *If $p \in \Theta(\frac{1}{n})$ then with positive probability that is constant in n the validator receives 0 in the searcher-optimal core allocation.*

Proof. For each opportunity $i \in \mathcal{A}$, with constant probability, there is exactly one searcher that has a positive value, i.e., $Y_i = 1$. Namely,

$$P[Y_i = 1] = \binom{n}{1}(1-p)^{n-1}p \to \frac{1}{e}\Theta(1),$$

as $n \to \infty$. Then, $P[Y_i = 1 \text{ for any } i] \to \Theta\left(\frac{1}{e^m}\right)$ as $n \to \infty$. That is, searchers complement each other in finding different arbitrages. From Proposition 4, with probability $\Theta\left(\frac{1}{e^m}\right)$, we have $M = 0$. ◀

▶ Remark 12. The previous results discuss the value distribution for scenarios where the block value is expected to be positive. If probability shrinks faster than $1/n$ as n grows, then with high probability the produced block has value 0. However, conditional on the block having positive value, all value is captured by searchers in the searcher-optimal core allocation.

As remarked in footnote 5, block builders observe around $\sim 1/3$ of submitted bundles clashing. We can use this number to get approximate values of the parameters in our model: we use as $n = 125$ which is the number of addresses that have placed at least 2 bundles in Ethereum blocks within the last 30 days prior to writing this paper (excluding addresses that have landed only one bundle gives us a crude way to identify addresses that have a high chance of being the main address used by a searcher) according to the website `https://libmev.com/`. Then

$$2/3 \approx P[Y_i < 2] = (1-p)^{125} + 125(1-p)^{124}p \Rightarrow p \approx 1\%.$$

Capacity Constraints

The previous model can be easily adapted to the case of capacity constraints on blocks where a block can only contain up to a fixed number of transactions, denoted by K. Note that submodularity is maintained when adding a capacity constraint. Thus, Proposition 4 naturally extends to the case of capacity constraints. For the model with independent bundles, we now can consider the value function

$$\bar{v}^K(S) = \max_{B \in \mathcal{B}(S)} \sum_{i \in \mathcal{A}} \sum_{j \in \mathcal{S}} v_{ij} B_{ij} = \max_{A' \subseteq \mathcal{A}, |A'| \leq K} \sum_{i \in A'} \max_{j \in \mathcal{S}} v_{ij}.$$

It follows immediately that Corollary 9 holds with

$$M^K := \max_{A' \subseteq \mathcal{A}, |A'| \leq K} \sum_{i \in A'} M_i$$

and $x_j^K := \sum_{i: j \in \mathcal{A}' \cap \operatorname{argmax}_j v_{ij}} (v_{ij} - M_i)$.

Next, we consider how the bounds on searcher and validator value capture for the stochastic model change with capacity constraints on the block. We assume that only a constant fraction of possible opportunities can be accommodated. With a bound on the block size, the probability threshold above which the validator captures all value in all core allocation becomes lower, and now matches the corresponding threshold from Proposition 11 where the probability is positive (Proposition 11 still holds with a bound on the block size).

▶ **Proposition 13.** *Let $n := |\mathcal{S}|$. assume the block has capacity to include $(1-\alpha)m$ transactions where $m := |\mathcal{A}|$ and $1/m < \alpha < 1$ is a constant. Then, there is a decreasing function ϕ such that if $p \in \omega(\frac{\phi(\alpha)}{n})$, the validator gets the entire block value with high probability in any core allocation.*

Proof. We show that there are at least 2 searchers who have positive value each arbitrage. As in the proof of Proposition 10 defining $Y_i := \sum_{j \in \mathcal{S}} v_{ij}$, we obtain

$$P[Y_i < 2] \leq e^{-pn} + npe^{-p(n-1)}. \tag{5}$$

Now consider the probability that for more than αm indices we have $Y_i < 2$. This is bounded by

$$P[Y_i < 2 \text{ for at least } \alpha m \text{ indices i}] \leq \binom{m}{\alpha m} P[Y_1 < 2]^{\alpha m} \leq \left(\frac{1}{\alpha}\left(e^{1-pn} + pne^{1-p(n-1)}\right)\right)^{\alpha m},$$

where the last inequality uses the well-known inequality $\binom{a}{b} \leq \left(\frac{ae}{b}\right)^b$ and the previously obtained inequality (5). Choosing $\phi(\alpha)$ to satisfy

$$(1 + \phi(\alpha))e^{-\phi(\alpha)} = \frac{\alpha}{e},$$

we have for $p \in \omega(\phi(\alpha)/n)$ that for any $\epsilon > 0$ there is a n such that

$$\frac{e^{1-pn} + pne^{1-p(n-1)}}{\alpha} < \epsilon$$

and therefore

$$P[Y_i \geq 2 \text{ for at least } (1-\alpha)m \text{ indices } i] \geq 1 - \varepsilon^{\alpha m} > 1 - \varepsilon.$$

Applying Proposition 4 shows the claim of the proposition. ◀

▶ Remark 14. The previous result is tight in the sense that the matching result Proposition 11 which holds for $p \in \Theta(1/n)$, still holds for the case of capacity constraints on the block.

5 Conclusion

We have studied MEV extraction in block building as a function of searcher competition and have argued that the core, and in particular the rule that selects the searcher-optimal point within the core are suitable solution concepts that allow us to make robust prediction about value distribution in MEV extraction. We have further identified a dominant-strategy incentive compatible mechanism, giving searchers their marginal value contribution to the realized block, which would be a theoretically appealing payment mechanism for searchers in the case of passive proposers/builders.

The model and solution concept allowed us to make sense of stylized facts about the Ethereum block building market: validators capture most of the value most of the time, and searchers with unique edge that are less exposed to competition are able to capture significant value. A natural extension of our model for further research would add correlation between different MEV opportunities to our stochastic model. Such an enhanced model would be particularly suitable to study the competition between different block builders and would possibly make theoretical predictions about concentration and value capture in the builder market.

References

1 Maryam Bahrani, Pranav Garimidi, and Tim Roughgarden. Centralization in block building and proposer-builder separation. *CoRR*, abs/2401.12120, 2024. doi:10.48550/arXiv.2401.12120.

2 Maryam Bahrani, Pranav Garimidi, and Tim Roughgarden. Transaction fee mechanism design in a post-mev world. *Cryptology ePrint Archive*, 2024.

3 Pedro Braga, Georgios Chionas, Piotr Krysta, Stefanos Leonardos, Georgios Piliouras, and Carmine Ventre. Who gets the maximal extractable value? A dynamic sharing blockchain mechanism. In Mehdi Dastani, Jaime Simão Sichman, Natasha Alechina, and Virginia Dignum, editors, *Proceedings of the 23rd International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2024, Auckland, New Zealand, May 6-10, 2024*, pages 2171–2173. ACM, 2024. doi:10.5555/3635637.3663097.

4 Agostino Capponi, Ruizhe Jia, and Sveinn Olafsson. Proposer-builder separation, payment for order flows, and centralization in blockchain. *Payment for Order Flows, and Centralization in Blockchain (February 12, 2024)*, 2024.

5 Philip Daian, Steven Goldfeder, Tyler Kell, Yunqi 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 2020, San Francisco, CA, USA, May 18-21, 2020*, pages 910–927. IEEE, 2020. doi:10.1109/SP40000.2020.00040.

6 Shayan Eskandari, Seyedehmahsa Moosavi, and Jeremy Clark. Sok: Transparent dishonesty: front-running attacks on blockchain. *CoRR*, abs/1902.05164, 2019. arXiv:1902.05164.

7 Donald B. Gillies. Solutions to general non-zero-sum games. *Contributions to the Theory of Games IV. Annals of Mathematics Studies*, 40:47–85, 1959.

8 Lioba Heimbach, Vabuk Pahari, and Eric Schertenleib. Non-atomic arbitrage in decentralized finance. *arXiv preprint*, 2024. arXiv:2401.01622.

9 Hervé Moulin. *Axioms of cooperative decision making.* Number 15 in Econometric Society Monographs. Cambridge university press, 1991.

10 Roger B Myerson and Mark A Satterthwaite. Efficient mechanisms for bilateral trading. *Journal of Economic Theory*, 29(2):265–281, 1983.

11 Noam Nisan, Tim Roughgarden, Éva Tardos, and Vijay V. Vazirani, editors. *Algorithmic Game Theory.* Cambridge University Press, 2007. doi:10.1017/CBO9780511800481.

12 Mallesh Pai and Max Resnick. Structural advantages for integrated builders in mev-boost. *arXiv preprint*, 2023. `arXiv:2311.09083`.

13 J David Pérez-Castrillo and Marilda Sotomayor. *Two Folk Manipulability Theorems in the General One-to-one Two-sided Matching Markets with Money.* FEA/USP, 2013.

14 Lloyd S. Shapley and Martin Shubik. The assignment game i: The core. *International Journal of Game Theory*, 1971.

15 Sen Yang, Kartik Nayak, and Fan Zhang. Decentralization of ethereum's builder market. *CoRR*, abs/2405.01329, 2024. `arXiv:2405.01329`.