Base Fee Manipulation In Ethereum's EIP-1559 Transaction Fee Mechanism

Sarah Azouvi[1] ✉ ⓘ
Unaffiliated, United Kingdom

Guy Goren ✉ ⓘ
Protocol Labs, Israel

Lioba Heimbach ✉ ⓘ
ETH Zurich, Switzerland

Alexander Hicks ✉ ⓘ
UCL, United Kingdom

─── **Abstract** ───

In 2021 Ethereum adjusted the transaction pricing mechanism by implementing EIP-1559, which introduces the *base fee* — a network fee that is burned and dynamically adjusts to the network demand. The authors of the Ethereum Improvement Proposal (EIP) noted that a miner with more than 50% of the mining power could be incentivized to deviate from the honest mining strategy. Instead, such a miner could propose a series of empty blocks to artificially lower demand and increase her future rewards. In this paper, we generalize this attack and show that under rational player behavior, deviating from the honest strategy can be profitable for a miner with less than 50% of the mining power. We show that even when miners do not collaborate, it is at times rational for smaller miners to join the attack. Finally, we propose a mitigation to address the identified vulnerability.

1 Introduction

Ethereum occupies a central role in the blockchain and decentralized application landscape. Not only does Ethereum's market capitalization currently exceed $225 billion [4], but it is also the leading platform for smart contracts and decentralized finance. Thus, Ethereum has revolutionized the way people envision and interact with blockchain technology.

The proposal of Ethereum Improvement Proposal #1559 (EIP-1559) [3] in April 2019 and its later deployment on Ethereum's mainnet in August 2021, mark a significant milestone for the Ethereum network. EIP-1559 reshaped Ethereum's transaction fee mechanism and remains in place to this day. One of the main goals of EIP-1559 is to simplify the bidding process by reducing the need for complex fee estimation algorithms while ensuring that the mechanism is *incentive compatible* — the best strategy for all players is to follow the protocol as intended. In particular, the mechanism should be both truthful and not incentivize miner or user bribes, as articulated by Roughgarden [19]. EIP-1559 has been shown to be incentive compatible [19, 20] when miners are assumed to be *myopic*, i.e., short-sighted in the sense

─────────────

[1] The authors of this work are listed alphabetically.

that they are maximizing only their immediate profits. Furthermore, an empirical study [16] concludes that EIP-1559 succeeded in making fees easier to estimate and in reducing delays.

EIP-1559 represents a departure from the previous *first-price auction* system, wherein users submit bids and pay the exact amount of their bid if their transaction was included. Importantly, the entirety of the fees paid by users is awarded to miners. This is no longer the case under EIP-1559 which introduced a base fee, a portion of the transaction fee that is burned and not awarded to miners. The base fee varies according to the fill rate of blocks — a proxy for network demand. Blocks exceeding a predefined target size increase the base fee and blocks below this target size decrease the base fee. Miners are then compensated for creating blocks through a block reward and user tips, i.e., user fees exceeding the base fee.

Considering the substantial financial value being traded on the Ethereum network, it is highly probable that profitable and rational deviations will occur when possible. Thus, it is essential for EIP-1559 to be incentive compatible as the protocol guarantees rely on participants following the intended behavior. Deviations from the intended behavior are considered attacks on the systems. As we will see, the dynamic nature of the base fee opens the door for possible rational attacks by miners. Miners have control over the fill rate of blocks and may thus choose to mine emptier blocks in order to decrease the base fee and increase their future profits, i.e., tips paid on top of the base fee by users. This genre of attack strategies has been acknowledged in Ethereum's EIP-1559 proposal and has been explored under various assumptions in previous research works [19, 11].

This paper presents an analysis of the potential for minority attacks on Ethereum's EIP-1559 transaction fee mechanism under the conservative assumption of a steady demand curve. Our results show that the mechanism is vulnerable to such a 20% minority attack.[2] Additionally, we show that smaller miners may be incentivized to join in on the attack. We provide general insights into when deviating from the prescribed strategy is rational and note that, due to the nature of our model and assumptions, the results are applicable in a wide range of scenarios. We also explain how the attack can be initiated by an Ethereum user (rather than a miner), i.e., show the incentives of users to enact bribes. Finally, to address the identified vulnerability, we propose a mitigation and evaluate its effectiveness through simulations.

2   Basic block reward mechanism in Ethereum

Block proposals. Whether Ethereum's blockchain relies on proof-of-work (PoW) or proof-of-stake (PoS) as sybil resistance, it relies on a leader election to determine the proposer of the next block. To be precise, in a PoW blockchain, miners compete to solve a computational puzzle, and the likelihood of a miner being chosen is based on their share of the network's computational power. In a PoS blockchain, on the other hand, miners stake amounts of the blockchain's native currency to participate and are randomly selected to create a new block with probability proportional to the amount they stake. In both cases, the ideal process is memoryless so each leader election is independent of the previous one. Our model covers both PoW and PoS, hence, the results of this paper apply to both types of blockchains.

When a miner is chosen to propose a block, they select a set of pending transactions to include in the next block and broadcast it to the network. Upon successful inclusion of their block in the blockchain, the miner receives two rewards: a fixed reward of newly minted currency (Ether in Ethereum's case), and a variable reward from the transaction fees of the

---

[2] An individual entity with more than 30% staking power currently exists in the Ethereum network [9].

included transactions. The block reward is fixed, regardless of the block's content or the miner that proposes it, so our analysis focuses on the potential for miners to increase their revenue via transaction fees. Therefore, we do not consider the block reward.

Transaction fees under EIP-1559. Ethereum transactions involve a set of instructions that are carried out by miners when the transaction is added to the blockchain. To prevent users from overwhelming the network with bogus transactions, users must pay *transaction fees*. These fees should reflect the amount of computational resources needed to execute the instructions, measured in units of *gas* and priced in *Gwei*. Note that 1 Gwei $\triangleq 10^{-9}$ Ether.

The EIP-1559 transaction fee includes a *base fee*, which is paid per unit of gas and varies to balance the supply of gas with the demand for gas. To be exact, the base fee $b[i]$ for block i is determined from the base fee $b[i-1]$ and size $s[i-1]$ of the previous block as follows:

$$b[i] \triangleq b[i-1] \cdot \left(1 + \phi \cdot \frac{s[i-1] - s^*}{s^*}\right). \tag{2.1}$$

Thus, the base fee is determined by comparing the size (measured in consumed gas units) of the previous block to a target block size s^*. If the block is larger than the target size, it indicates high demand for gas, and the base fee is increased to decrease demand. Conversely, if the block size is smaller than the target, it indicates low demand for gas, and the base fee is decreased to increase demand. The sensitivity of the base fee to the size of the previous block is determined by the adjustment parameter ϕ that is currently set to $\frac{1}{8}$ on Ethereum. We note that $b[0]$ was set to 1 Gwei initially and that the maximum valid block size is $2s^*$.

When creating a transaction under the EIP-1559 mechanism, in addition to the gas limit (g), the user must specify the fee cap (c) which is the maximum fee per gas unit they are willing to pay, and a maximum tip per gas unit (ε) which is the priority fee. The transaction will be included in a block only if the fee cap is greater than or equal to the base fee (b). The total fee paid by the user is $\tilde{g} \cdot \min\{b + \varepsilon, c\}$, where $\tilde{g} < g$ is the actual gas consumed by the transaction.[3] A portion of the fee $\tilde{g} \cdot b$ is burnt and the remaining $\tilde{g} \cdot \min\{\varepsilon, c - b\}$ goes to the miner as a tip. The EIP-1559 mechanism aims for users to bid small tips that only cover a miner's costs [3, 19]. Miners are intended to include all transactions that have a fee cap greater or equal to the base fee and prioritize transactions with higher fees only if the maximal block size (s_{max}) is exceeded. For simplicity, we assume that all transactions are of the same size and have a sufficient gas limit (i.e., $g = \tilde{g} = 1$) to eliminate considerations including knapsack and gas estimation and keep the focus on the core matter. E.g., a large transaction is modeled by multiple smaller transactions.

3 Model and Assumptions

In the following, we present our model and outline our assumptions. We highlight that our model follows Roughgarden's [19] very closely with the exception that we do not restrict miners to be myopic. Specifically, while Roughgarden [19] considers only a miner's immediate profits (myopici), we will also consider her future profits.

Users. We assume that users are rational agents who want their submitted transactions to be processed and included in the blockchain. The cost of having one's transaction (tx) included in the blockchain under EIP-1559 depends on the base fee (b), fee cap (c_{tx}), and the maximum tip (ε_{tx}), which we described in Section 2. Each transaction tx has a value v_{tx} that

---

[3] The gas limit (g) specifies the amount of gas units available for the execution of the transaction. The amount of gas needed to execute a transaction is not always predictable in advance. Moreover, if the needed amount of gas exceeds g, then g gas units are consumed and paid for but the transaction fails.

is private to the user who proposes it, which can be thought of as the maximum price that the user is willing to pay for the transaction tx to be included in the blockchain. We, therefore, take the utility of a user proposing tx at each block to be $u(tx) = v_{tx} - \min\{b + \varepsilon_{tx}, c_{tx}\}$ if tx is included and 0 otherwise. The user will then adjust c_{tx} and ε_{tx} according to her bidding strategy, while b is determined by the size of the past blocks according to Equation 2.1.

Miners. Miners produce blocks that include transactions to be executed. In our model, the miner to propose the next block is chosen at random. Drawings constitute independent experiments. In each drawing, a miner X is chosen with probability p_x, where p_x is the miner's share of the total network power. We assume that the power distribution in the network changes slowly, hence, to simplify the analysis we model the system as having a fixed network power distribution, i.e., miners' network power does not change with time.

A miner has dictatorial power over which transactions to include in the block she produces. We assume that miners are rational agents who wish to maximize their profit and therefore choose which transactions to include in their block according to a strategy that maximizes their profit. Aside from their transaction picking strategy, we assume that miners behave "honestly" – that is, as specified by the protocol. We highlight that by not considering other forms of deviating from the protocol, we strengthen our result, showing that even "practically honest" miners would deviate from EIP-1559.

Collaboration. We adopt a standard buyers-sellers perspective. We assume that miners do not collaborate with each other, as they can be viewed as one miner with more power. Similarly, we assume that users do not collaborate with one another since they are competing for the same scarce resource. However, a user (buyer) and a miner (seller) can communicate and adjust their strategies for mutual benefit. In particular, a (sophisticated[4]) user and a miner will collaborate if it benefits both of them. That is, the user pays fewer fees for her transaction while the miner receives more fees. The collaboration is thus rational behavior, leading to improved outcomes for both.

Steady State. The distribution of transaction values is represented by a demand curve, and the standard demand curve is a decreasing function; the higher the fee is, the fewer transactions are willing to pay it. In our work, we make no assumptions about the shape of the demand curve other than that it is a decreasing function. We consider a system in *steady state*, which we define as a system in which the demand curve does not change over time, i.e., it is the same whenever a miner creates a block. This steady state assumption is good for several reasons: (1) it keeps us on par with Roughgarden's work [19], (2) it simplifies the analysis by removing noisy components that can obscure the core principles, and (3) it is a conservative assumption that strengthen our results in comparison to others. For example, the "steady influx" model — where there is a fixed influx of transactions to the network — assumes that miners are able to manipulate the demand curve in their favor, i.e., artificially increase demand by delaying the inclusion of transactions. Thus, when the attack strategy we will present in Section 4 is beneficial under the *steady state* it will be at least as beneficial under the "steady influx" model, but not vice versa. Hence, our steady demand curve assumption makes our results more general and robust, as it applies to a wider range of adversarial assumptions.

The desired dynamics of EIP-1559 in the steady state are that produced blocks are of target size s^* and that the base fee remains constant at what we refer to as the target base

---

[4] It is expected that sophisticated participants will emerge, actively seeking opportunities to generate excess profits (reduce costs). This expectation is supported by the history of traditional exchange systems, where a multitude of players specialize in high-frequency trading.

fee b^*, i.e., $s[i] = s^*$ and $b[i] = b^*$ for all i. Further, EIP-1559's desired bidding dynamics are for the user's optimal strategy to be honest in reporting the value of a transaction via c_{tx} and to offer a minimal tip. In other words, in the steady state, EIP-1559 should lead to a Nash-equilibrium with the following strategies: (users) honest value-reporting, and (miners) including all txs with $c_{tx} > b^*$. Then a block produced during steady state would include s^* transactions that are each paying $b^* + \varepsilon$ in fees. Thus, $b^* \cdot s^*$ is burned (red area in Figure 1a), and $\varepsilon \cdot s^*$ is received by the miner (blue area in Figure 1a).

4 A Miner's Deviation from the Honest Strategy

In the following, we consider a miner X who controls a proportion p_x of the network's mining power, i.e., the probability that X proposes the next block is p_x.

Honest strategy. The honest strategy for X is always to include the maximum possible number of transactions whose gas fee covers at least the base fee. As we consider a system in steady state, the demand curve does not change over time. Further, the honest user bids $c_{tx} = b^* + \varepsilon$ and $\varepsilon_{tx} = \varepsilon$. Thus, miner X will always propose a block that is exactly the target size s^*. The payout received by miner X for every proposed block is, therefore, $s^* \cdot \varepsilon$, where ε is the tip the miner receives. The costs for the users are $(b^* + \varepsilon)s^*$.

Deviation from honest strategy. We continue by outlining a strategy miner X and sophisticated users can employ to manipulate the base fee which results in increasing X's profit and reducing users' costs. When X proposes a block, for which the preceding block was not created by X, she proposes an empty block to reduce the base fee b for subsequent blocks. The miner will receive no payout for this block. Any other consecutive blocks X is chosen to propose, she will propose at target size s^*, profiting from the difference between the targeted and the reduced base fee. To highlight the miner's profit, we assume in the following that users are motivated by slightly reducing their costs. Specifically, users continue to submit transactions with $c_{tx} = b^* + (1-\alpha)\varepsilon$ and $\varepsilon_{tx} = \phi \cdot b^* + (1-\alpha)\varepsilon$, where $0 < \alpha \le 1$. Thus, the miner will receive at least the difference in base fees, i.e., $\phi \cdot b^*$, and the user will pay $\alpha \cdot \varepsilon$ less for her transaction. By making this assumption, we let the miner extract most of the excess profit. In Section 7 we describe the complementary attack where most of the excess value is gained by the users.

Whenever X proposes a s^* sized block with an artificially reduced base fee, X receives

$$s^* \left(\phi \cdot b^* + (1-\alpha)\varepsilon \right), \tag{4.1}$$

where $\phi \cdot b^*$ is the base fee reduction and α represents the proportional reduction of the tip paid by the users. For simplicity, we assume the attack finishes whenever miner X's consecutive turns as proposer finishes. The deviation can only become more profitable if X can continue the attack after an honest proposer, thus, our results apply without the simplification and their generality is not reduced.

In Theorem 1 we compute the expected reward of X following the honest strategy, as well as the aforementioned described deviation from the honest strategy. By comparing the payout of consecutive turns of X in both strategies, we find that it is rational also for a miner with less than 50% of the power to deviate from the honest strategy.

▶ **Theorem 1.** *In expectation, it is rational for miner X to deviate from the honest strategy, if*

$$p_x > \frac{\varepsilon}{\phi \cdot b^* + (1-\alpha)\varepsilon}.$$

Proof. See Appendix A. ◀
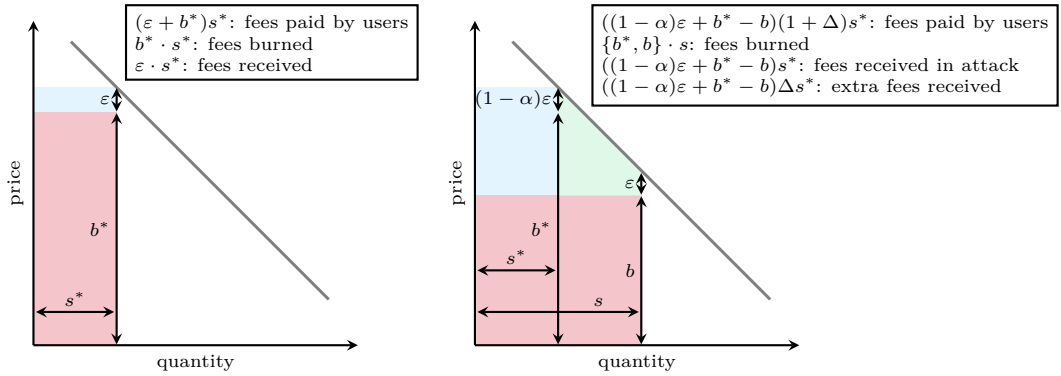
(a) Sample steady state demand curve, s^* transactions are willing to pay $b^*+\varepsilon$, where b^* is the base fee that corresponds to the target size. The red area is the amount of fees being burned, whereas the blue area are the fees received by the miner.

(b) Sample steady state demand curve with lowered base fee b. If the base fee was lowered to $b < b^*$, the demand is increased to $s \geq s^*$. An honest miner will fill up the block with all transactions, paying at least the base fee. The red area indicates the amount of fees being burned, the blue area the fees the attacking miner X receives, and the green area is the additional fees received by the honest miner Y for including all transactions. The scaling factor Δ dictates the size of the green area.

Figure 1 Example demand curves for Ethereum transactions. The quantity (x-axis) indicates the number of transactions willing to pay the transaction fee (price shown y-axis).

To better illustrate, when it is rational behavior for miner X to deviate from the honest strategy, we plot the relative difference between the expected reward of the attack and the honest strategy in Figure 2. For many realistic parameter configurations, it is rational behavior for miner X to perform the attack and thereby manipulate the base fee. In Figure 2a, we plot the profitability of the attack in comparison to the honest strategy dependent on X's mining power p_x. We set $\phi = 1/8, \varepsilon/b^* = 1/25$, and $\alpha = 0.5$. Notice that even miners with a mining power of less than 0.3 are expected to profit from executing the attack. To underline the expected profitability of the attack, even for small miners, we vary the ratio between the tips and the base fee (ε/b^*) in Figure 2b and set $p_x = 0.3$. Additionally, we vary α, the share of the tips the users keep for themselves, in Figure 2c, and again set $p_x = 0.3$. We conclude that there are multiple realistic parameter configurations for which the outlined attack is profitable, even for a miner with less than 50% of the mining power. Thus, as individual Ethereum pools control in excess of 30% of the staking power [9] such an attack is realistic.

5   The Attack's Effect on Other Miners

In the following, we consider a scenario where a miner X with staking power p_x ($0 < p_x < 1$) exists for which it is rational behavior to perform the base fee manipulation studied in Section 4. We then analyze the effect of a miner Y with staking power p_y, where $0 < p_y < p_x$, observing that X performs the base fee manipulation attack. More precisely, we study whether it is rational behavior for Y to join the attack partially, i.e., Y will always propose blocks at target size and thereby help keep the base fee artificially low in Section 5.1. Additionally, we will also study when Y would rationally join the attack entirely, i.e., Y also proposes empty blocks when the base fee is not already artificially lowered in Section 5.2. We remark that, throughout, we always assume that miners X and Y do not collaborate.

5.1   Joining the Attack

Consider a miner Y that observes X performing the attack outlined from Section 4. Miner Y
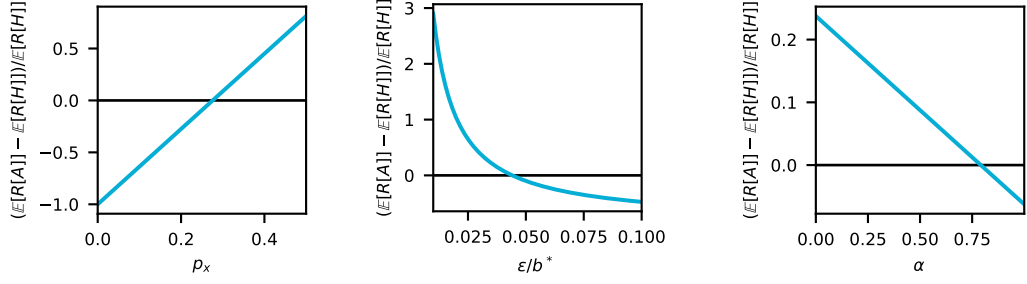
(a) Relative difference shown as a function of p_x. We set $\alpha = 0.5$, $\varepsilon/b^* = 1/25$.

(b) Relative difference shown as a function of ε/b^*. We set $p_x = 0.3$, $\alpha = 0.5$.

(c) Relative difference shown as a function of α. We set $p_x = 0.3$, $\varepsilon/b^* = 25$.

■ **Figure 2** Relative difference between the expected reward of attack and honest strategy as a function of p_x (cf. Figure 2a), ε/b^* (cf. Figure 2b) and α (cf. Figure 2c). We set $\phi = 1/8$, as set in Ethereum. It is rational for X to perform the attack whenever the relative difference is positive. The equations to produce these graphs can be found in the proof of Theorem 1 in Appendix A.

is selected as the proposer for a block that follows X's turn as proposer, i.e., the base fee is currently artificially lowered to $(1 - \phi)b^*$. We analyze whether it is rational for Y to follow the honest strategy, i.e., propose the largest block possible and thereby increase the base fee again, or to join the attack and continue keeping the base fee artificially low.

Honest strategy. We first describe the honest strategy. When it is miner Y's turn to propose a block at an artificially lower base fee, miner Y proposes a block with the most transactions possible. Note that the number of transactions is restricted both by the demand at the current base fee (b), as well as the maximum block size, which is twice the target size ($2s^*$). We now consider the demand curve that is drawn in Figure 1b. The demand at price $b^* + \varepsilon$, where b^* is the target base fee price and ε the tip, corresponds to a block of target size s^*. Miner Y will propose a block with the artificially lowered base fee $b = (1 - \phi)b^*$. The demand at this new price $b + \varepsilon$ is represented as s in Figure 1b. Recall, that we make no assumptions about the shape of the demand curve. To make our results stronger, we consider the best case for the honest strategy. Namely, due to the increased demand, Y can propose a block of maximum size (and reap the resulting extra rewards). That is, $s = 2s^*$, after X's turn. Thus, the payout for miner Y proposing a block of size $2s^*$ is given by

$$s^* \left(\phi \cdot b^* + (1 - \alpha)\varepsilon \right) (1 + \Delta), \tag{5.1}$$

where $s^* \left(\phi \cdot b^* + (1 - \alpha)\varepsilon \right)$ is the payout the attacking miner X would receive (cf. Equation 4.1) and $\Delta \in [0, 1]$ is a scaling factor that dictates how much additional rewards miner Y received for mining a maximum size block. While $\Delta = 0$ would indicate that Y earns exactly as much as X, i.e., all extra transactions are capped at exactly the base fee, $\Delta = 1$ would indicate that miner Y earns twice the rewards of X, i.e., all extra transactions are capped at the highest possible price of $b^* + \varepsilon$.

After miner Y proposes the block, she will be chosen to propose the next block with a probability of p_y. We continue with the approximation from before, i.e., the effect of a full block after an empty one leads approximately to the same point on the demand curve $-$ $(s^*, b^* + \varepsilon)$. We note that this approximation is accurate up to a term in $O(\phi^2 b^*)$ where $b \in \Theta(\phi b^*)$. Thus, miner Y proposes a block of size s^* and is awarded $s^* \cdot \varepsilon$. From thereon out, she will continue doing so until her consecutive turns as a proposer stops.

With probability p_x, miner X will interrupt Y's turn as proposer. Miner X will start the attack again and propose an empty block to lower the base fee. Note that our approximation of the base fee returning to steady-state levels is the best case for Y's honest strategy and,
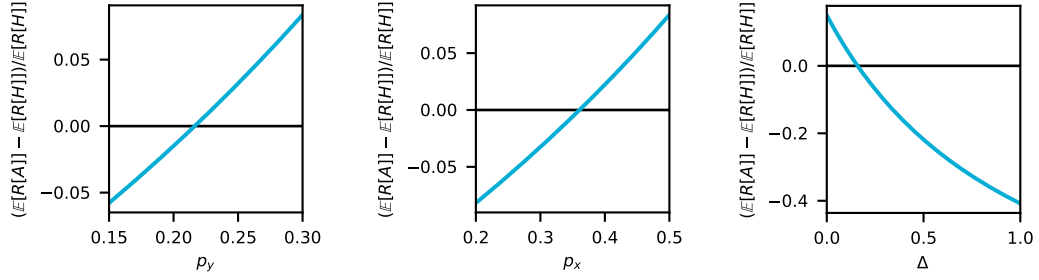
(a) Relative difference shown as a function of p_y. We set $p_x = 0.3$, $\Delta = 0.2$.

(b) Relative difference shown as a function of p_x. We set $p_y = 0.6 p_x$, $\Delta = 0.2$.

(c) Relative difference shown as a function of Δ. We set $p_x = 0.3$, $p_y = 0.18$.

■ **Figure 3** Relative difference between the expected reward of attack and honest strategy as a function of p_y (cf. Figure 3a), p_x (cf. Figure 3b) and Δ (cf. Figure 3c). We set $\phi = 1/8$, as implemented in Ethereum, $\varepsilon/b^* = 1/25$, $\alpha = 0.5$ in all plots. It is rational for Y to join the attack whenever the relative difference is positive. The equations to produce these graphs can be found in the proof of Theorem 2 in Appendix A.

thus, makes our results stronger. For all consecutive blocks proposed by X, miner X will propose target size blocks to keep the base fee b (i.e., $(1-\phi)b^*$). If miner Y is again selected as a proposer after X's turn, miner Y will proceed with her previously outlined strategy.

At any point, with a probability of $1 - p_y - p_x$, the consecutive turn of the two miners finishes. Note that for Y's honest strategy analysis, we are only interested in these consecutive turns of the two miners as proposers, as we analyze the expected payout of the same sequences for the deviation from the honest strategy which we outline in the following.

Deviation from honest strategy. We now describe a strategy Y can employ to join the attack she observes. When it is Y's turn to propose a block and the base fee is artificially lowered by X, miner Y will propose a block at the target size. Equivalently to the payout received by miner X for such a block (cf. Section 4), miner Y's reward for proposing the block is given by

$$s^* \left(\phi \cdot b^* + (1 - \alpha)\varepsilon \right). \tag{5.2}$$

Following the block proposed by Y, miner Y is again selected to propose a block with probability p_y and miner X with probability p_x. As long as the two miners have an uninterrupted sequence of block proposals, they both keep the base fee artificially low by always proposing target size blocks. Once their consecutive turn as proposers ends, we consider the attack finished. In Theorem 2, we show that it can be profitable for such a miner Y with a mining power p_y to join the attack, i.e., keep the base fee artificially low if she sees a miner X with a mining power $p_x > p_y$ perform the attack. Importantly, this is without assuming collaboration between the two miners.

▶ **Theorem 2.** *In expectation, it is rational for a miner Y to deviate from the honest strategy and join X in keeping the base fee low, if*

$$p_y > \frac{\Delta((1-\alpha)\varepsilon + \phi \cdot b^*)}{(1-\alpha)\Delta \cdot \varepsilon + (1+\Delta)\phi \cdot b^* - \alpha \cdot \varepsilon}.$$

Proof. See Appendix A.   ◀

To better gauge when it is profitable for miner Y to join the attack, we plot the relative difference between the expected return from the outlined attack and the expected return from the honest strategy in Figure 3. We vary the mining powers of miner Y (cf. Figure 3a) and miner X (cf. Figure 3b), as well as Δ (cf. Figure 3c). In all three plots we set

$\phi = 1/8, \varepsilon/b^* = 1/25$, and $\alpha = 0.5$. In Figure 3a, we set $\Delta = 0.2$ and notice that even a miner Y with a mining power slightly larger than 0.2 would be inclined to join the attacking miner X with a mining power of 0.3 in manipulating the base fee. We observe in Figure 3b that a miner Y that is only six-tenths of the size of miner X would also join the attack even if miner X only controls less than 40% of the mining power. Finally, in Figure 3c, we show the dependency of the attack's profitability for miner Y as a function of Δ, i.e., the additional payout received by miner Y following the honest strategy when proposing a full block after the attack by X. Notice that while it is rational for a miner Y to join the attack for small Δ's, this is not the case for larger Δ's. We remark that this is due to the rewards from the full block mined if Y follows the honest strategy being very significant for a large Δ. Nevertheless, our results show that for realistic parameter configurations, it is rational for a miner Y to join the attack she sees a larger miner X perform — even without assuming collaboration between the two.

5.2 Join and Initiate the Attack

In addition to only joining the attack, it is also possible for miner Y, observing X continuously performing the base fee manipulation, to also propose an empty block whenever she proposes a block with the target base fee (b^*), knowing that X will aid her in keeping the base fee low subsequently. We analyze, in the following, when it is more profitable for miner Y to join X's attack in her entirety, as opposed to remaining honest.

Honest strategy. The honest strategy for miner Y is identical to that described in Section 5.1. However, now it is also important to mention that anytime miner Y proposes a block with base fee b^*, i.e., whenever Y proposes a block that does not follow X's attack, Y will propose a target size (s^*) block. For proposing such a block, miner Y will receive $s^* \cdot \varepsilon$.

Deviation from honest strategy. In the deviation from the honest strategy, Y will propose an empty block whenever she mines a block where the base fee corresponds to the target base fee b^*. Then with probability p_y, miner Y will also propose the next block and will profit from the difference between the base fee and the target base fee by mining a target size block. On the other hand, with probability p_x, miner X will propose the next block and will also mine a target size block – keeping the base fee low. With Theorem 3, we show that it can even be profitable for a miner Y to commence the attack if she knows that a larger miner X will help her in keeping the base fee artificially low.

▶ **Theorem 3.** *In expectation, it is rational for a miner Y to deviate from the honest strategy and lower the base fee, if*

$$p_y > \frac{\varepsilon(1 - p_x)}{(1 - \alpha)\varepsilon + \phi \cdot b^*(1 - p_x) + (1 + \Delta)\varepsilon\alpha p_x - \Delta p_x(\varepsilon + \phi \cdot b^*)}.$$

Proof. See Appendix A. ◀

To better understand the parameter configuration under which it would be rational for Y to also initiate the attack, we plot the relative difference of the reward of the attack in comparison to the reward of the honest strategy in Figure 4. Note that we again set $\phi = 1/8, \varepsilon/b^* = 1/25$, and $\alpha = 0.5$. Figure 4a, where we also set $p_x = 0.3$ and $\Delta = 0.2$, shows that it can be profitable for a miner Y to initiate the attack, i.e., mine an empty block to lower the base fee, knowing that miner X will support her in keeping the base fee artificially low. Notice though that the threshold where it is rational for Y to also start the attack is reached later in comparison to the threshold where it is rational for miner Y to only join the attack (cf. Figure 3a). A similar picture paints itself when we look at the profitability of
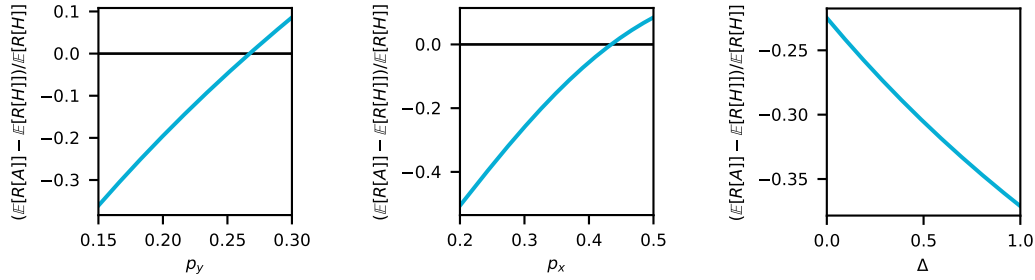
(a) Relative difference shown as a function of p_y. We set $p_x = 0.3$, $\Delta = 0.2$.

(b) Relative difference shown as a function of p_x. We set $p_y = 0.6p_x$, $\Delta = 0.2$.

(c) Relative difference shown as a function of Δ. We set $p_x = 0.3$, , $p_y = 0.18$.

■ **Figure 4** Relative difference between the expected reward of attack and honest strategy as a function of p_y (cf. Figure 4a), p_x (cf. Figure 4b) and Δ (cf. Figure 4c). We set $\phi = 1/8$, as implemented in Ethereum, $\varepsilon/b^* = 1/25$, $\alpha = 0.5$ in all plots. It is rational for Y to initiate new attacks in addition to X whenever the relative difference is positive. The equations to produce these graphs can be found in the proof of Theorem 3 in Appendix A.

initiating the attack for Y as a function of X's mining power in Figure 4a. Again we see that for a miner Y, it can be profitable to start the attack only with the knowledge that X will aid her in keeping the base fee low. Finally, in Figure 4c, we show the relative difference between the expected profit of the outlined attack for Y and the honest strategy. For the chosen parameter configuration, $p_x = 0.3$ and $p_y = 0.18$, the attack would actually never be profitable regardless of Δ. We, thus, summarize that while it is only rational for Y to initiate the attack for a reduced set of parameters, it is remarkable that this is even the case. By starting the attack, Y is taking a loss for a larger miner X just based on the knowledge that she will be supported by X in keeping the base fee low. Astonishingly, the collaboration of the two miners is not required.

6   Possible Mitigations

To mitigate the attacks described in Sections 4 and 5, we focus on addressing the deviation of player X, as if X does not deviate from the honest strategy, and neither does Y. We start by examining the trivial mitigation of reducing ϕ. Theorem 1 shows that decreasing ϕ by a factor of β will require that p_x is approximately β times larger for a deviation to be profitable. For example, if Ethereum were to decrease its current ϕ of $1/8$ to $1/16$, it would require p_x to be approximately twice as large for a deviation to be profitable. This approach can be effective during stable periods, but it might not be able to adjust quickly enough to changes in demand. Further, Leonardos et al. [14] show how the value of ϕ determines the trade-off between a base fee that adjusts too quickly (which can lead to chaotic behavior) and one that adjusts too slowly to fulfill its purpose.

The following question appears in EIP1559 FAQ [2]: *"Won't miners have the incentive to collude to push down the base fee by making all their blocks less than half full?"* In response to the question Buterin proposes this mitigation: *"Divert half of the collected base fees, that would otherwise be burned, into a special pool. Whenever a miner mines a new block add to her block reward a $1/8192$ portion of the amount in that pool. This will incentivize miners to maintain a higher base fee."* One might falsely presume that this solution, to the above-posed question, might also be used to solve the deviation we outline in Section 4. However, the two attacks are inherently different as the one we outline does not require miners to collude. Thus, this proposal does not solve our deviation. The added cost of the attack, i.e., the lost

revenue from half of the base fee reduction, is distributed among many while the attacker's expected revenue remains unchanged. In the long run, this proposal only minimally reduces the attacker's expected revenue by $b^* \cdot 2^{-13}$, which is typically significantly less than the expected rewards and, therefore, ineffective.

Another straw man to consider is to use the average of the previous W block sizes instead of just the size of the last block to measure demand in the base fee update rule. This method may appear to be promising because it reduces the effect of an empty block on the following block by a factor of W, and its effect on the rate of adjustment (embodied in ϕ) is easily accounted for, adding an adjustment delay within $O(W)$. However, this mitigation fails to mitigate the attack, as it increases the opportunity for X to profit from later blocks that are within a W distance from the empty block, thereby increasing the expected profitability of the deviation. For example, if we use a two-block window (i.e., $W = 2$) and $\phi = 1/8$, the base fee is reduced by a smaller factor of only $\phi/W = -1/16$ to $b_1 = 15b^*/16$ as desired. Nevertheless, even if the ensuing block is completely full (i.e., $s_1 = 2s^*$), the base fee for the block following it does not increase and remains $b_2 = b_1 = 15b^*/16$ which means an additional profit opportunity for the deviation that compensates X for the reduced factor. As a result, the deviation is not mitigated and is actually exacerbated.

We propose the following mitigation, to use a geometric sequence as weights to average the history of block sizes. Formally, for $q \in (0, 1)$ denote

$$s_{avg}[i] \triangleq \frac{1-q}{q} \sum_{k=1}^{\infty} q^k \cdot s[i - k + 1] = (1 - q) \cdot s[i] + q \cdot s_{avg}[i - 1], \qquad (6.1)$$

and replace $s[i]$ in Equation 2.1 by $s_{avg}[i]$ to get the following base-fee update rule

$$b[i + 1] = b[i] \cdot \left(1 + \phi \cdot \frac{s_{avg}[i] - s^*}{s^*} \right). \qquad (6.2)$$

We note that $b[0]$ would be initialized to 1 Gwei initially and that s_{avg} would be initialized to be the size of the first block after the transition. By applying the update rule in Equation 6.2, we reduce the effect of the empty block on the ensuing block by a factor of $(1 - q)$ and discount its effects on later blocks exponentially fast (with base q). For example, using the same parameters as before with $\phi = 1/8$, if we set $q = 1/2$, after X's empty block (at slot τ), the base fee will be reduced by a factor of $(1 - \phi(1 - q)) = (1 - 1/16)$ to $b[\tau + 1] = 15b^*/16$. Now, however, making the same assumption as before (i.e., $s[\tau + 1] = 2s^*$), the base fee for the next block, $b[\tau + 2]$, will be $495b^*/512$. This decreases the potential profit margin of block $\tau + 2$ from $b^*/16$ to $b^*/32 + b^*/512$, which is almost a factor of 2 reduction. Therefore, the added profit opportunity in future blocks is not enough to compensate for the lost potential in the immediately ensuing block. As a consequence, the attack is considerably mitigated.

The above mitigation method has two additional properties: (i) its computation and space complexity are both in $O(1)$, and (ii) it gradually phases out the impact of a single empty block without causing significant fluctuations. To reason about the effect our proposal has on response times we use the following methodology. Suppose that the demand suddenly changes and the new (desired) steady state should be reached at a new point with a base fee that is β times higher than the current base fee (b^*). Denote by T the number of consecutive full blocks required to reach the new base fee. We use T as a function of β to characterize the response time of a fee-setting mechanism to sudden changes in demand.

EIP-1559 as it is currently implemented will take $T(\beta) = \lceil \log_{(1+\phi)} \beta \rceil$ blocks to reach the new base fee βb^*. With our mitigation, it takes slightly longer. To be precise, in Appendix B we show the exact delay $T(\beta)$ of our mitigation proposal is the smallest integer T that satisfies $\prod_{k=1}^{T} (1 + \phi(1 - q^k)) \geq \beta$. We further plot the $T(\beta)$ for EIP-1559 and our mitigation

in Appendix B (cf. Figure 10).

6.1 Simulations

In order to gauge the effectiveness of the proposed mitigation, we conducted simulations that compare the excess profit of an attacker (profit gained through deviation minus profit gained through honest mining) under EIP-1559 with and without the mitigation. To account for the probabilistic nature of the attack (profits in expectation only), we calculated the average of the results for each data point over 10,000 runs, each using a different random seed. Each simulation begins with X mining an empty block, the next blocks are mined by X with probability p_x per block and by an honest miner with probability $1 - p_x$. We assume that X will then always mine target-size blocks, while the honest miners will mine blocks at twice the target size. We keep track of the payout received by X and declare the attack finished if the base fee has recovered to 99% of its target size. We perform 10,000 runs for both the base fee evolution according to EIP-1559, as well as our mitigation. Simultaneously, we keep track of the payout X would have received for the same random seed if she had followed the honest strategy. The results of the simulation are depicted in Figures 5 and 6.
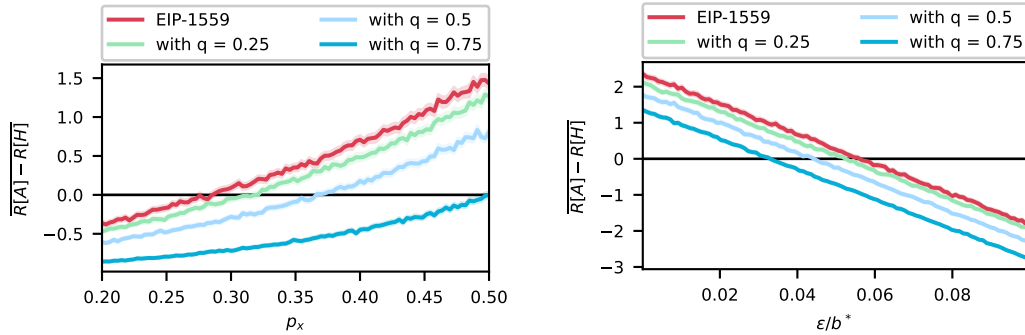
(a) Difference of mean attack and mean honest return as a function of p_x for EIP-1559 and the proposed mitigation for $q \in \{1/4, 1/2, 3/4\}$. We set $\varepsilon/b^* = 1/25$.

(b) Difference of mean attack and mean honest return as a function of ε/b^* for EIP-1559 and the proposed mitigation for $q \in \{1/4, 1/2, 3/4\}$. We set $p_x = 0.4$.

Figure 5 Profitability of the attack under EIP-1559 and our proposed mitigation. We plot the mean attack profitability (cf. Figures 5a and 5b) along with the 95% confidence interval. We set $\phi = 1/8$, $\alpha = 0.5$, $s^* = 1$.

Figures 5a and 5b show X's profit from attacking as a function of p_x and the ratio ε/b^* for EIP-1559 with and without the mitigation for $q \in \{1/4, 1/2, 3/4\}$. The results decisively demonstrate the benefit of our mitigation; it becomes much harder for X to profit from attacking. Finally, Figure 6 illustrates the effect of the mitigation (with $q \in \{1/4, 1/2, 3/4\}$) on the configurations at which X will attack. Most notably, the green area represents a set of configurations in which X had attacked without the mitigation and will be attacking no longer. The value of the proposed approach is evident from the results.

7   User perspective

Until now, we have approached the topic from the miners' point of view. Considering the users as first-class citizens, and observing the attack through their eyes, contributes a new perspective on the results.

(a) $q = 1/4$ **(b)** $q = 1/2$ **(c)** $q = 3/4$
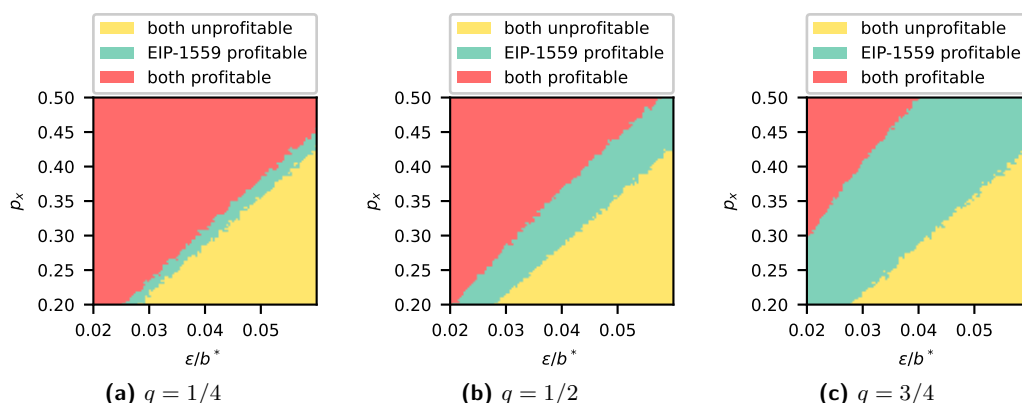
Figure 6 Attack profitability under EIP-1559 and the proposed mitigation for $q \in \{1/4, 1/2, 3/4\}$ as a function of p_x and ε/b^*. Importantly, the green area shows where the mitigation can prevent the attack but EIP-1559 cannot. We set $\phi = 1/8$, $\alpha = 0.5$, $s^* = 1$.

Instead of the miners taking it upon themselves to initiate the attack, we can imagine users who wish to pay lower costs coordinating the attack. Let u be a user (or group of users) that has transactions with a g amount of gas. Assume the other users naively follow the strategy of the desired equilibrium. That is, they bid an honest valuation with an ε tip. The attacker's strategy is as follows: u bribes the miner of the current block (no matter the miner's power) to propose an empty one instead. Any bribe larger than $s^*\varepsilon$ will suffice. Consequently, the base fee reduces in the next block. If the other users naively continue to bid with an ε tip (or they are simply slow to react), u can guarantee the inclusion of her transactions with any tip larger than ε — making the attack profitable whenever $g\phi b^* > s^*\varepsilon$.

8   Discussion

Myopic vs. non-myopic. Roughgarden's analysis [19] of the EIP-1559 protocol suggests that, in the steady state, the honest strategies of the users and miners are incentive compatible. Our opposing conclusion stems from a single different assumption, namely, [19] considers miners that only care for immediate profits (referred to as myopic), while we consider non-myopic miners that do not disregard future profits. There are valid reasons to model miners as myopic. For example, the proposing turns of a very small miner are sporadic, and accounting for rare future profits is negligible in comparison to the prize at hand. However, there are strong reasons for modeling miners as non-myopic. Measurement studies have shown that the distribution of mining power follows a power law rather than a uniform distribution [8, 9], i.e., the biggest miners control significant portions of the mining power. Moreover, the "shorter vision" of smaller miners is accounted for under our model. Our quantitative results show that as a miner's size decreases, the lesser the profitability of the deviations. Finally, from a conceptual perspective, miners are typically players of high stakes (the minimum stake in Ethereum is 32 ETH which is currently over $50,000), and participation also requires some expertise, planning, and locking of assets. Therefore, it does not seem appropriate to consider these players as ones that neglect future considerations.

Additional observations. The deviations described in this paper have an interesting property; they appear to have a win-win-win outcome. The attacker profits, the non-attacking miners profit, and the users profit by paying a cheaper total gas fee. But not all is rosy; there is a hidden cost involved. In order for users to benefit, they must diligently follow the miners' actions and compute the appropriate response. This increase in complexity for

the users is in opposition to one of the main goals of EIP-1559 – simplifying the bidding mechanism and eliminating the need for complex fee estimation algorithms. As a result, sophisticated users take precedence and push naïve users to the back of the line.

Although it is desirable that the leader election process be unpredictable, in practice, this is not the case. Currently, in Ethereum, implementation considerations led to miners knowing their own proposing slots 32–62 blocks in advance [6]. This predictability clearly favors the attacker, who no longer needs to lose tips for the probability of winning more. Instead, the miner only attacks when it is guaranteed to mine at least two blocks in a row. The predictability issue has risen with the move to PoS and was not present in PoW Ethereum.[5]

9   Related work

Blockchain transaction fee mechanism. Huberman et al. [12] provide an early analysis of Bitcoin's first-price auction. An in-depth exploration of miner manipulation in transaction fee mechanisms was first explored by Lavi et al. [13]. While these works study the first-price auction used in Bitcoin and originally in Ethereum as well, our work studies the incentives for miners to deviate from the EIP-1559 protocol currently used in Ethereum.

Stability of the base fee. As Leonardos et al. [14] show in their theoretical analysis, a key parameter in the base fee mechanism is the adjustment parameter ϕ. In particular, they show that stability is not guaranteed depending on system conditions (e.g., a congested network), especially if ϕ is set to too high a value. However, there are also conditions under which the base fee may have bounded oscillations or even converge, providing stability. In another work, Leonardos et al. [15] show that despite the short-term chaotic behavior on the base fee, the long-term average block size is close to the target size.

Reijsbergen et al. [18] empirically show that a stable base fee may not tell the whole story, however, as even in cases where the base fee remains relatively stable (e.g., between 25 and 35 Gwei) and the block size is on average the target block size, block sizes can fluctuate wildly (as explained by Leonardos et al. [14]), impacting miner revenue. One reason for this is that the currently used value for the adjustment parameter ($\phi = \frac{1}{8}$) is too low during periods where the demand rises sharply (i.e., the base fee does not increase quickly enough) but also too high when demand is stable (inducing fluctuations in block size). Therefore, [18] suggests making ϕ variable based on the demand. Their work does not consider bribes and is complementary to ours. Their suggested mitigation to the stability issue is to have ϕ adaptive according to an Additive Increase Multiplicative Decrease (AIMD). Since we do not vary ϕ (but instead update the base fee update rule), an interesting experiment would be to combine our mitigation technique (averaging $s[i]$ geometrically) with their AIMD ϕ setting and examine the results on data from the real world.

Ferreira et al. [7] show that although the first-price auction utilized under EIP-1559 is incentive-aligned for miners, it provides a bad user experience. In particular, they observe bounded oscillation of the base fee in experiments when bidders all associate the same value with their transaction. While their work studies the stability of the base fee with myopic miners, we study the manipulability of the base fee in the presence of non-myopic miners.

Manipulability of the base fee. Manipulation of the base fee has been a concern since EIP-1559 was proposed [1] as it is straightforward to notice that the base fee could be

---

[5] Since the randomness in PoW does not depend on a peer-to-peer communication source, it does not have a predictability concern. The predictability issue is a result of implementation constraints for the cryptographic protocols of Verifiable Random Functions (VRFs).

manipulated downwards by a miner that intentionally mines empty blocks.

The EIP as listed on Ethereum's Github repository acknowledges the possibility of miners mining empty blocks but determines that such a deviation from an honest mining strategy would not lead to a stable equilibrium as other miners would benefit from this (i.e., benefiting from the reduced base fee without the opportunity cost of mining an empty block) [3]. Their belief was therefore that executing such a strategy would require a miner to control more than half the hashing power (the document precedes Ethereum's switch to PoS).

In his exposition of EIP-1559 [19], Roughgarden considered this in the case of a 100% miner (or any miner with greater than 50% of the mining power) that would drive the base fee down to 0 by mining empty blocks then, in order to maximize their revenue, switching to either mining target size blocks in perpetuity, maintaining the base fee at 0 and essentially reverting back to a first price auction, or mining sequences of under full and overfull blocks (relative to the target size) according to the demand. Roughgarden restricts himself to this case and does not otherwise consider non-myopic miners, assuming that with a high enough level of decentralization, the probability of any miner being elected to propose a block two times in a row was too low for any miner to consider strategies over multiple blocks.

Similar to us, a concurrent work by Hougaard and Pourpouneh [11] also relaxes the myopic assumption and proves that non-myopic miners would be incentivized to deviate even if they are a minority. However, their analysis relies on several assumptions, that while legitimate, provide the attacker with advantageous conditions in comparison to the more conservative assumptions made in this paper. In particular, [11] relies on miners knowing the exact parameters of the demand distribution of the users (which itself is restricted to each user drawing from a uniform distribution), as well as on miners being able to manipulate the demand curve in favor of the future blocks; inducing artificial future congestion by not including transactions in the current block and letting them accumulate. We do not assume any specific demand curve, only a steady demand curve, and do not rely on miners inducing congestion, which makes our attack more profitable. Although we do assume that users act rationally and will adapt their strategy if benefits them, while [11] assumes users are passive and do not adapt their strategies in a rational manner.

MEV. Miner/Maximal Extractable Value (MEV) has gained significant attention in the blockchain research community in recent years [5, 21, 17]. Similarly to this work, MEV strategies enable a miner to accrue excess profits in comparison to naïve mining. However, in MEV the value comes from analyzing the actual data in the transactions, whereas, in this work, the miner's excess profit comes from manipulating the EIP-1559 mechanism. Therefore, many of the suggested mitigations against MEV [10] (such as obscuring transaction data until inclusion) will not work against our attack.

10   Conclusion and Future Work

In this paper, we demonstrated that even under very conservative assumptions (steady state, miners cannot induce congestion, unknown demand functions), there are strong incentives for minority miners to deviate from the EIP-1559 protocol. Furthermore, we showed that once an attack begins, previously honest miners' rational response may be to join the deviation and even sometimes initiate new attacks — worsening the problem rather than improving it.

To mitigate the problem, we suggested using a weighted average with the weights being a geometric series. This direction seems promising as it has several desirable properties and trade-offs (e.g., balancing attack mitigation with low additional response delays). However, further research rigorously analyzing it in a broader context is warranted.

Many questions remain open regarding EIP-1559. To name but a few: While the suggested deviation captures the essence of the problem, devising an optimal deviation strategy is left as an open problem. What is the right abstraction to evaluate the benefits and shortcomings of EIP-1559 in comparison to a first-price auction? How does it differ from PoW to PoS? On a broader level, questions concerning adaptations of auction mechanisms for a blockchain use case have the potential to be both intellectually challenging and practically important and will hopefully receive more attention from the academic community.

References

1 Tim Beiko. EIP-1559 Community Outreach Report. `https://medium.com/ethereum-cat-herders/eip-1559-community-outreach-report-aa18be0666b5`, 2020. Accessed: 2023-01-25.

2 Vitalik Buterin. EIP 1559 FAQ. `https://notes.ethereum.org/@vbuterin/eip-1559-faq#Won%E2%80%99t-miners-have-the-incentive-to-collude-to-push-down-the-BASEFEE-by-making-all-their-blocks-less-than-half-full`, 2021. Accessed: 2023-01-26.

3 Vitalik Buterin, Eric Conner, Rick Dudley, Matthew Slipper, Ian Norden, and Abdelhamid Bakhta. Fee market change for ETH 1.0 chain. `https://github.com/ethereum/EIPs/blob/master/EIPS/eip-1559.md`, 2019. Accessed: 2020-09-28.

4 CoinMarketCap. Today's Cryptocurrency Prices by Market Cap. `https://coinmarketcap.com/`, 2017. Accessed: 2023-07-21.

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)*, pages 910–927. IEEE, 2020.

6 Ethereum. Phase 0 – Honest Validator. `https://github.com/ethereum/consensus-specs/blob/dev/specs/phase0/validator.md#lookahead`, 2022. Accessed: 2023-01-27.

7 Matheus VX Ferreira, Daniel J Moroz, David C Parkes, and Mitchell Stern. Dynamic posted-price mechanisms for the blockchain transaction-fee market. In *Proceedings of the 3rd ACM conference on Advances in Financial Technologies*, pages 86–99, 2021.

8 Adem Efe Gencer, Soumya Basu, Ittay Eyal, Robbert Van Renesse, and Emin Gün Sirer. Decentralization in bitcoin and ethereum networks. In *Financial Cryptography and Data Security: 22nd International Conference, FC 2018, Nieuwpoort, Curaçao, February 26–March 2, 2018, Revised Selected Papers 22*, pages 439–457. Springer, 2018.

9 Dominic Grandjean, Lioba Heimbach, and Roger Wattenhofer. Ethereum proof-of-stake consensus layer: Participation and decentralization. *arXiv preprint arXiv:2306.10777*, 2023.

10 Lioba Heimbach and Roger Wattenhofer. Sok: Preventing transaction reordering manipulations in decentralized finance. In *4th ACM Conference on Advances in Financial Technologies (AFT), Cambridge, Massachusetts, USA*, September 2022.

11 Jens Leth Hougaard and Mohsen Pourpouneh. Farsighted miners under transaction fee mechanism eip1559. Technical report, IFRO Working Paper, 2022.

12 Gur Huberman, Jacob D Leshno, and Ciamac Moallemi. Monopoly without a monopolist: An economic analysis of the bitcoin payment system. *The Review of Economic Studies*, 88(6):3011–3040, 2021.

13 Ron Lavi, Or Sattath, and Aviv Zohar. Redesigning bitcoin's fee market. *ACM Transactions on Economics and Computation*, 10(1):1–31, 2022.

14 Stefanos Leonardos, Barnabé Monnot, Daniël Reijsbergen, Efstratios Skoulakis, and Georgios Piliouras. Dynamical analysis of the eip-1559 ethereum fee market. In *Proceedings of the 3rd ACM Conference on Advances in Financial Technologies*, AFT '21, page 114–126, New York, NY, USA, 2021. Association for Computing Machinery. `doi:10.1145/3479722.3480993`.

15 Stefanos Leonardos, Daniël Reijsbergen, Daniël Reijsbergen, Barnabé Monnot, and Georgios Piliouras. Optimality despite chaos in fee markets. *arXiv preprint arXiv:2212.07175*, 2022.

16 Yulin Liu, Yuxuan Lu, Kartik Nayak, Fan Zhang, Luyao Zhang, and Yinhong Zhao. Empirical analysis of eip-1559: Transaction fees, waiting times, and consensus security. In *Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security*, CCS '22, page 2099–2113, New York, NY, USA, 2022. Association for Computing Machinery. `doi:10.1145/3548606.3559341`.

17 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.

18 Daniël Reijsbergen, Shyam Sridhar, Barnabé Monnot, Stefanos Leonardos, Stratis Skoulakis, and Georgios Piliouras. Transaction fees on a honeymoon: Ethereum's eip-1559 one month later. In *2021 IEEE International Conference on Blockchain (Blockchain)*, pages 196–204. IEEE, 2021.

19 Tim Roughgarden. Transaction fee mechanism design for the Ethereum blockchain: An economic analysis of EIP-1559. *arXiv preprint arXiv:2012.00854*, 2020.

20 Tim Roughgarden. Transaction fee mechanism design. *ACM SIGecom Exchanges*, 19(1):52–55, 2021.

21 Christof Ferreira Torres, Ramiro Camino, et al. 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.

A   Omitted Proofs

▶ **Theorem 1.** *In expectation, it is rational for miner X to deviate from the honest strategy, if*
$$p_x > \frac{\varepsilon}{\phi \cdot b^* + (1 - \alpha)\varepsilon}.$$

Proof. We commence with the honest strategy and calculate the expected payout for consecutive turns of X as proposer by modeling the process as a Markov chain (cf. Figure 7a).

When it is miner X's first time as proposer in a consecutive turn, i.e., the previous block was not mined by X, we are in state X^h. In this state, miner X proposes a block at target size s^* and receives tips at price ε. Thus, the payout for miner X in state X^h is $P[X^h] = s^* \cdot \varepsilon$. With probability p_x, X stays in state X^h for the next block and also proposes the next block. Else, with probability $1 - p_x$, we enter an absorbing state. We calculate the expected number of times X proposes consecutive blocks and, thereby, the expected payout for X.

By linearity of expectation, it follows that the expected payout for a sequence of consecutive turns by X as a proposer following the honest strategy is given by

$$\mathbb{E}[R[X^h]] = p_x \cdot \mathbb{E}[R[X^h]] + P[X^h] \iff \mathbb{E}[R[X^h]] = \frac{s^* \cdot \varepsilon}{1 - p_x}, \tag{A.1}$$

as miner X is awarded $P[X^h]$ every time she proposes a block. Note that the expected reward of the honest strategy i.e., $\mathbb{E}[R[H]]$ corresponds to the expected payout of the Markov process starting in state X^h, i.e., $\mathbb{E}[R[H]] = \mathbb{E}[R[X^h]]$.

(a) honest strategy of miner X **(b)** attacking strategy of miner X
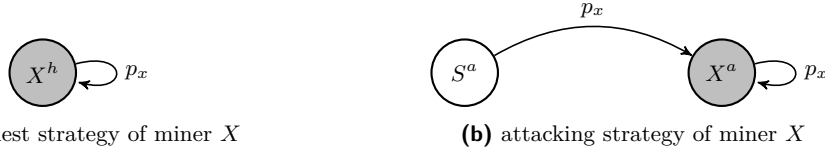
Figure 7 The honest strategy (cf. Figure 7a) and deviation from the honest strategy (cf. Figure 7b) modeled with discrete Markov chains. All states with a nonzero payout for miner X are highlighted in gray. We transition between states with every block. Note that for all remaining probabilities, the Markov process enters an absorbing state, and X's consecutive turns as a proposer finish.

We now examine the reward for miner X if she chooses to carry out the attack, modeling the deviation from the honest strategy as a Markov chain (as shown in Figure 7b). The starting point for the attack, denoted as state S^a, is when miner X submits an empty block. Therefore, the payout in state S^a is 0. But with a probability of p_x, miner X is chosen as the proposer for the next block and enters state X^a, where she submits a block at the target size. The payout in state X^a can be calculated by $P[X^a] = s^* (\phi \cdot b^* + (1 - \alpha)\varepsilon)$.

If X is not selected to propose another block in a row, we enter an absorbing state and the attack stops. We make the approximation that whenever X's consecutive turns as proposer finish, honest miners will bring the base fee back to b^* before X gets to mine another block. It is possible that the honest miners do not bring the base fee back up to the target before X is selected to propose again, and X could continue the attack at a lower cost. However, assuming that the base fee had fully recovered simplifies the analysis and only makes our results stronger, as it reduces X's attack rewards. From state X^a we remain in state X^a for the next block with probability p_x, i.e., X is selected to propose another block or enter the absorbing state with probability $1 - p_x$.

The expected payout for miner X in state S^a is given by $\mathbb{E}[R[S^a]] = p_x \cdot \mathbb{E}[R[X^a]] + P[S^a]$, where $\mathbb{E}[R[X^a]]$ is the expected payout starting from state X^a and we have $\mathbb{E}[R[X^a]] =$

$p_x \cdot \mathbb{E}[R[X^a]] + P[X^a]$. It follows that the expected payout of the attack is

$$\mathbb{E}[R[A]] = \mathbb{E}[S^a] = \frac{p_x \cdot s^* \left(\phi \cdot b^* + (1-\alpha)\varepsilon\right)}{(1-p_x)}. \tag{A.2}$$

Note that the payout of the attack is given by the expected payout starting from state S^a, as the attack commences in said state.

We conclude that it is rational for miner X to deviate from the honest strategy if $\mathbb{E}[R[A]] > \mathbb{E}[R[H]]$. It follows that X attacks when

$$p_x > \tfrac{\varepsilon}{\phi \cdot b^* + (1-\alpha)\varepsilon}.$$

$\blacktriangleleft$

$\blacktriangleright$ **Theorem 2.** *In expectation, it is rational for a miner Y to deviate from the honest strategy and join X in keeping the base fee low, if*

$$p_y > \frac{\Delta((1-\alpha)\varepsilon + \phi \cdot b^*)}{(1-\alpha)\Delta \cdot \varepsilon + (1+\Delta)\phi \cdot b^* - \alpha \cdot \varepsilon}.$$

Proof. We commence with the honest strategy of miner Y which we model as a Markov chain in Figure 8a. Miner Y starts in state Y_X^h and is tasked with proposing a block after miner X at an artificially lowered base fee b, where $b = (1-\phi)b^* < b^*$. The payout for proposing a block at twice the target size s^* is given by

$$P[Y_X^h] = s^* \left(\phi \cdot b^* + (1-\alpha)\varepsilon\right)(1+\Delta). \tag{A.3}$$

From state Y_X^h, we move to state Y_Y^h, where Y proposes a target size block, with probability p_y. The payout for miner Y in state Y_Y^h is given as

$$P[Y_Y^h] = s^* \cdot \varepsilon, \tag{A.4}$$

and we remain in this state for a subsequent block with probability p_y.

In both state Y_X^h and state Y_Y^h, the probability of moving to state X^h is p_x. When we re-enter state X^h for the first time, X will propose an empty block to lower the base fee again. Then, with a probability p_x, we remain in state X for the next block, where miner X will propose target size blocks until her consecutive turn as a proposer is interrupted. Regardless, the payout for miner Y whenever we are in state X^h is zero. We move to state Y_Y^h with probability p_y from state X^h and enter an absorbing state with probability $1 - p_y - p_x$ from all states.

Next, we calculate the expected payout of the honest strategy and start with the expected reward for miner Y starting from state Y_X^h $\mathbb{E}[R[Y_X^h]] = p_x \cdot \mathbb{E}[R[X^h]] + p_y \cdot \mathbb{E}[R[Y_Y^h]] + P[Y_X^h]$ where $\mathbb{E}[R[X^h]] = p_x \cdot \mathbb{E}[R[X^h]] + p_y \cdot \mathbb{E}[R[Y_X^h]]$, and $\mathbb{E}[R[Y_Y^h]] = p_x \cdot \mathbb{E}[R[X^h]] + p_y \cdot \mathbb{E}[R[Y_Y^h]] + P[Y_Y^h]$. By solving the system of linear equations, we find that the expected reward from the honest strategy is given by

$$\mathbb{E}[R[H]] = \mathbb{E}[R[Y_X^h]] \quad = \tfrac{(1-p_x)(((1-p_y)(1+\Delta)\phi \cdot b^*) + \varepsilon(1+\Delta - \alpha(1+\Delta)(1-p_y) - \Delta p_y))s^*}{(1-p_x-p_y)}. \tag{A.5}$$

We now consider the deviation from the honest strategy, whereby miner Y also keeps the base fee artificially low, and model the strategy with a Markov chain in Figure 8b. State Y^a is the starting state of the deviating strategy in which Y proposes a block with target size s^* at an artificially lowered base fee $(1-\phi)b^*$. Thus, the state's payout is

$$P[Y^a] = s^* \left(\phi \cdot b^* + (1-\alpha)\varepsilon\right). \tag{A.6}$$

For the next block, we stay in state Y^a with probability p_y, move to state X^a with probability p_x, or else move to an absorbing state, i.e., the consecutive turns of X and Y as proposers end and the attack finishes. In state X^a, proposer X will also propose a target size block, but the payout for miner Y is zero as we assume no collaboration between the two. The transition probabilities from state X^a are identical to those from state Y^a: we

(a) honest strategy of miner Y

(b) attacking strategy of miner Y
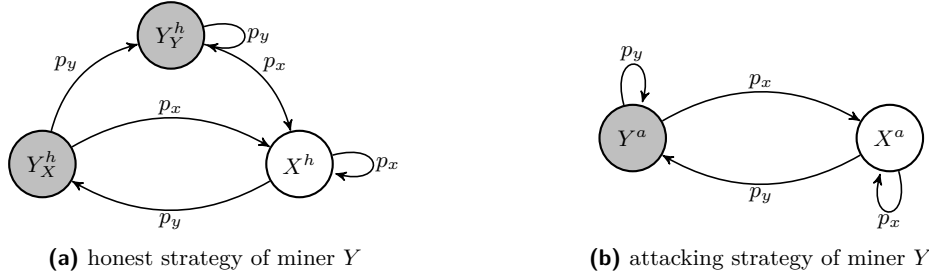
Figure 8 The honest strategy (cf. Figure 8a) and the deviation from the honest strategy, i.e., keep the base fee artificially low, (cf. Figure 8b) modeled with discrete Markov chains for miner Y. We transition between states with every block. All states with a nonzero payout for miner Y are highlighted in gray. Note that for all remaining probabilities, the Markov process enters an absorbing state.

move to state Y^a with probability p_y, stay in state X^a with probability p_x and enter an absorbing state with probability $1 - p_y - p_x$ for the next block. Thus, the expected reward for miner Y starting in state Y^a is given by $\mathbb{E}[R[Y^a]] = p_x \cdot \mathbb{E}[R[X^a]] + p_y \cdot \mathbb{E}[R[Y^a]] + P[Y^a]$, where $\mathbb{E}[R[X^a]]$ is the expected payout for miner Y starting from state X^a and we have $\mathbb{E}[R[X^a]] = p_x \cdot \mathbb{E}[R[X^a]] + p_y \cdot \mathbb{E}[R[Y^a]]$. We follow that the expected payout of the deviating strategy for miner Y is given by

$$\mathbb{E}[R[A]] = \mathbb{E}[R[Y^a]] = \frac{(1 - p_x)s^*(\phi \cdot b^* + (1 - \alpha)\varepsilon)}{(1 - p_x - p_y)}. \tag{A.7}$$

We conclude it is rational for miner Y to deviate from the honest strategy when $\mathbb{E}[R[A]] - \mathbb{E}[R[H]] > 0$. It follows that Y attacks when

$$p_y > \frac{\Delta((1-\alpha)\varepsilon + \phi \cdot b^*)}{(1-\alpha)\Delta \cdot \varepsilon + (1+\Delta)\phi \cdot b^* - \alpha \cdot \varepsilon}.$$

◀

▶ **Theorem 3.** *In expectation, it is rational for a miner Y to deviate from the honest strategy and lower the base fee, if*

$$p_y > \frac{\varepsilon(1 - p_x)}{(1 - \alpha)\varepsilon + \phi \cdot b^*(1 - p_x) + (1 + \Delta)\varepsilon\alpha p_x - \Delta p_x(\varepsilon + \phi \cdot b^*)}.$$

Proof. We, again, model the honest strategy for miner Y as a Markov chain (cf. Figure 9a). The honest strategy starts in state Y^h, where miner Y proposes a block at target size s^* and receives a payout of

$$P[Y^h] = s^* \cdot \varepsilon. \tag{A.8}$$

The transition probability to state X^h is p_x and to state Y^h, i.e., Y proposes consecutive blocks, is p_y. Otherwise, the consecutive turn of X and Y as miners stops and we enter an absorbing state. We note that the states X^h, Y^h_X and Y^h_Y correspond exactly to the eponymous states in Theorem 2 (cf. Figure 8a). Thus, the actions of the miners, the payout for miner Y, and the transition probabilities are as previously described in the proof of Theorem 2.

The strategy's payout is the expected reward starting from state Y^h, which is

$$\mathbb{E}[R[Y^h]] = p_x \cdot \mathbb{E}[R[X^h]] + p_y \cdot \mathbb{E}[R[Y^h] + P[Y^h]. \tag{A.9}$$

As previously in Theorem 2, we have

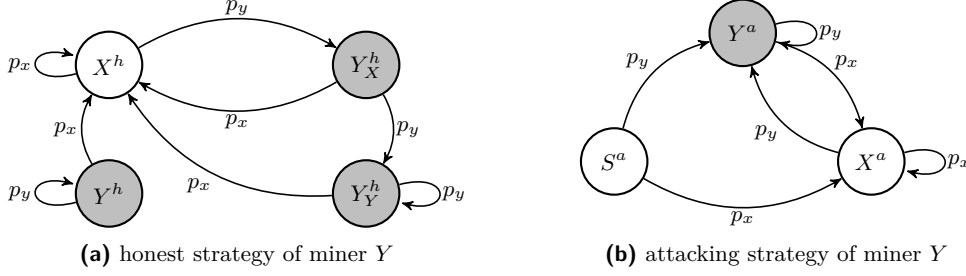$$\mathbb{E}[R[Y^h_X]] = p_x \cdot \mathbb{E}[R[X^h]] + p_y \cdot \mathbb{E}[R[Y^h_Y]] + P[Y^h_X], \tag{A.10}$$

$$\mathbb{E}[R[X^h]] = p_x \cdot \mathbb{E}[R[X^h]] + p_y \cdot \mathbb{E}[R[Y^h_X]], \tag{A.11}$$

$$\mathbb{E}[R[Y^h_Y]] = p_x \cdot \mathbb{E}[R[X^h]] + p_y \cdot \mathbb{E}[R[Y^h_Y]] + P[Y^h_Y]. \tag{A.12}$$

Solving the system of linear equation, we conclude that the expected payout of the honest strategy is

$$\mathbb{E}[R[H]] = \mathbb{E}[R[Y^h]] = \frac{(((1+\Delta)b^*\phi p_x p_y) + \varepsilon(1 - p_x + ((1-\alpha)\Delta - \alpha)p_x p_y))s^*}{1 - p_x - p_y}. \quad \text{(A.13)}$$

(a) honest strategy of miner Y **(b)** attacking strategy of miner Y

■ **Figure 9** The honest strategy (cf. Figure 9a) and the deviation from the honest strategy (cf. Figure 9b) modeled with discrete Markov chains. We transition between states with every block. All states with a nonzero payout for miner Y are highlighted in gray. Note that for all remaining probabilities, the Markov process enters an absorbing state and the consecutive turn of X and Y as proposers finishes.

We proceed with the attack strategy, which we model in Figure 9b. Miner Y starts in state S^a and mines an empty block and, therefore, receives no rewards. With probability p_y we move to state Y^a for the next block, i.e., Y proposes a target size (s^*) block, with probability p_x we move to state X^a, i.e., X proposes a target size (s^*) block, and with probability $1 - p_x - p_y$ we move to an absorbing state, i.e., the attack ends. Notice that the states X^a and Y^a are identical to those described in the proof of Theorem 3. Thus, the expected returns starting in the respective states are as follows

$$\mathbb{E}[R[S^a]] = p_x \cdot \mathbb{E}[R[X^a]] + p_y \cdot \mathbb{E}[R[Y^a]], \quad \text{(A.14)}$$

$$\mathbb{E}[R[Y^a]] = p_x \cdot \mathbb{E}[R[X^a]] + p_y \cdot \mathbb{E}[R[Y^a]] + P[Y^a], \quad \text{(A.15)}$$

$$\mathbb{E}[R[X^a]] = p_x \cdot \mathbb{E}[R[X^a]] + p_y \cdot \mathbb{E}[R[Y^a]], \quad \text{(A.16)}$$

and find that the expected reward of the attack strategy is

$$\mathbb{E}[R[A]] = \mathbb{E}[R[S^a]] = \frac{(\phi \cdot b^* + (1-\alpha)\varepsilon)p_y s^*}{1 - p_x - p_y}. \quad \text{(A.17)}$$

To conclude, it is rational behavior for Y to deviate from the honest strategy, when $\mathbb{E}[R[A]] - \mathbb{E}[R[H]] > 0$, which holds when

$$p_y > \frac{\varepsilon(1 - p_x)}{(1-\alpha)\varepsilon + \phi \cdot b^*(1 - p_x) + (1+\Delta)\varepsilon\alpha p_x - \Delta p_x(\varepsilon + \phi \cdot b^*)}.$$

◀

B   Delay Incurred by the Mitigation of Section 6

Suppose we are in a steady state with a base fee b^*, when after block height τ a fixed change in demand occurs for which the new steady state will be achieved with the new base fee $\beta \cdot b^*$. We denote by T the number of consecutively full blocks it takes to reach the new base fee $\beta \cdot b^*$.

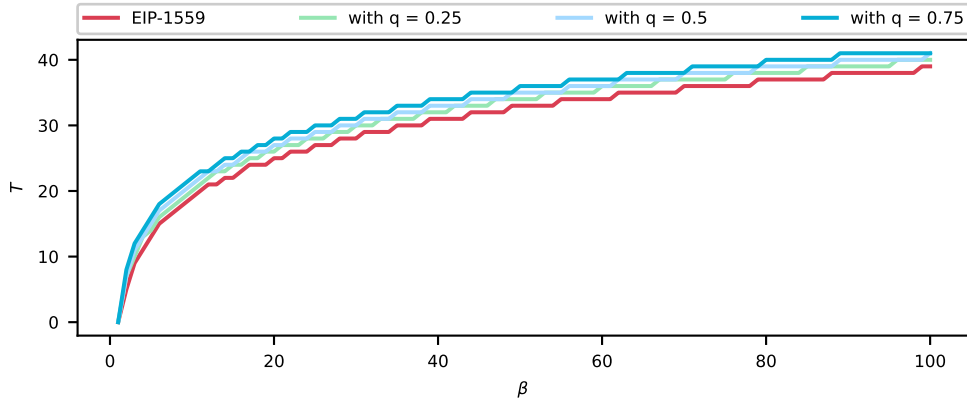After k consecutively full blocks, according to Eq. 6.1

$$
\begin{aligned}
s_{avg}[\tau + k] &= (1 - q)2s^* + q \cdot s_{avg}[\tau + k - 1] \\
&= \left(2(1 - q)(q^0 + q^1 + \ldots + q^{k-1}) + q^k\right) s^* \\
&= \left(2(1 - q)\left(\frac{1 - q^k}{1 - q}\right) + q^k\right) s^* \\
&= (2 - q^k)s^*.
\end{aligned}
$$

Plugging the above into Eq. 6.2 yields

$$
\begin{aligned}
b[\tau + k] &= b_{avg}[\tau + k - 1] \cdot \left(1 + \phi \cdot \frac{(2 - q^k)s^* - s^*}{s^*}\right) = b_{avg}[\tau + k - 1] \cdot \left(1 + \phi(1 - q^k)\right) \\
&= b[\tau] \cdot \left(1 + \phi(1 - q^1)\right)\left(1 + \phi(1 - q^2)\right) \cdots \left(1 + \phi(1 - q^k)\right) \\
&= b^* \cdot \prod_{i=1}^{k}(1 + \phi(1 - q^i)).
\end{aligned}
$$

Therefore, T is the smallest integer that satisfies

$$
b[\tau + T] = b^* \cdot \prod_{i=k}^{T}(1 + \phi(1 - q^k)) \geq \beta \cdot b^* \quad \Longleftrightarrow \quad \prod_{k=1}^{T}(1 + \phi(1 - q^k)) \geq \beta.
$$

Figure 10 The number of consecutive full blocks T required to increase the base fee by factor β for EIP-1559 and the proposed mitigation for $q \in \{1/4, 1/2, 3/4\}$.

Figure 10 plots $T(\beta)$ for both EIP-1559 and our mitigation. We set $\phi = 1/8$ (current Ethereum) and use $q \in \{1/4, 1/2, 3/4\}$. All plots follow a logarithmic trend and the response times to dramatic changes in demand are only mildly affected by the proposed mitigation (even for β factors as large as 100).