HAL open science

ON PROFITABILITY OF STUBBORN MINING

Cyril Grunspan, Ricardo Pérez-Marco

To cite this version:

Cyril Grunspan, Ricardo Pérez-Marco. ON PROFITABILITY OF STUBBORN MINING. 2019. hal-02334535

HAL Id: hal-02334535

https://hal.science/hal-02334535v1

Preprint submitted on 26 Oct 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

arXiv:1808.01041v1 [cs.CR] 2 Aug 2018

ON PROFITABILITY OF STUBBORN MINING

CYRIL GRUNSPAN AND RICARDO PÉREZ-MARCO

Abstract. We compute and compare profitabilities of stubborn mining strategies that are variations of selfish mining. These are deviant mining strategies violating Bitcoin's network protocol rules. We apply the foundational set-up from our previous companion article on the profitability of selfish mining, and the new martingale techniques to get a closed-form computation for the revenue ratio, which is the correct benchmark for profitability. Catalan numbers and Catalan distributions appear in the closed-form computations. This marks the first appearance of Catalan numbers in the Mathematics of the Bitcoin protocol.

1. Introduction

In our previous article [3] we gave a rigorous foundation for the profitability analysis of alternative mining strategies in the Bitcoin network [6]. As for games with repetition, it depends on the proper analysis of the revenue and the duration over attack cycles. More precisely, we prove that the expected revenue $\mathbb{E}[R]$ and expected duration $\mathbb{E}[\tau]$ over an attack cycle give the "Revenue Ratio"

$$\Gamma = \frac{\mathbb{E}[R]}{\mathbb{E}[\tau]}$$

This is the correct benchmark for the profitability of the strategy.

This analysis was applied to the "Selfish Mining" strategy from [1]. We also introduced in [3] martingale tools that yield, by the application of Doob's Stopping Time Theorem, the expected duration of the attack cycles. With the Markov model (used in [1] and other articles in the literature) one cannot compute the expected cycle duration $\mathbb{E}[\tau]$. As we prove in [3], it is only after a difficulty adjustment that Selfish Mining or any other "block withholding strategy" can become profitable. To compute the expected time of the next difficulty adjustment it is necessary to be able to compute the expected duration of cyles $\mathbb{E}[\tau]$.

2010 *Mathematics Subject Classification.* 68M01, 60G40, 91A60.
Key words and phrases. Bitcoin, blockchain, proof-of-work, selfish mining, Catalan numbers.

In this article we apply again these new powerful techniques to some of the "Stubborn Mining" strategies presented in [7]. So far only a numerical Monte-Carlo analysis seems to be known for these strategies. For the "Lead-Stubborn Mining" (LSM) and the "Equal Fork Stubborn Mining" (EFSM) strategies we compute the "Revenue Ratios".

We fix some notations. Let $b > 0$ be the block reward, and τ_0 the average inter-block validation time for the total network (around 10 minutes for the Bitcoin network). We denote by q the relative hashing power of the attacker. Let γ be the fraction of the honest network that the attacker attracts to mine on top of his fork. For a miner that after a difficulty adjustment has a Revenue Ratio $\tilde{\Gamma}$ we define his apparent hashrate $\tilde{q}$ by

$$\tilde{q} = \frac{\tilde{\Gamma} \cdot \tau_0}{b} \ .$$

The apparent hashrate of a miner can also be defined after a difficulty adjustment as the average proportion of blocks mined by the miner in the official blockchain. We make use of

$$C(x) = \frac{1 - \sqrt{1-4x}}{2x} = \frac{2}{1 + \sqrt{1-4x}} = \sum_{n=0}^{+\infty} C_n x^n$$

which is the generating series for Catalan numbers $(C_n)_{n \geq 0}$ (see Appendix A).

Theorem 1 (Lead-Stubborn mining)**.** *The revenue ratio of the "Lead-stubborn mining" strategy is*

$$\Gamma(LSM) = \left(q - \frac{pq(p-q)(1-\gamma)}{\gamma} \cdot \frac{1 - p(1-\gamma)C((1-\gamma)pq)}{p + q(p-q)} \right) \frac{b}{\tau_0} \ .$$

After a difficulty adjustment, the apparent hashrate $\tilde{q}_{LSM}$ of the stubborn miner is

$$\tilde{q}_{LSM} = q \cdot \frac{p + pq - q^2}{p + pq - q} - \frac{pq(p-q)(1-\gamma)}{\gamma} \cdot \frac{1 - p(1-\gamma)C((1-\gamma)pq)}{p + pq - q} \ .$$

Theorem 2 (Equal Fork Stubborn mining)**.** *The revenue ratio of the "Equal Fork Stubborn mining" strategy is*

$$\Gamma(EFSM) = \left(q - \left(\frac{1-\gamma}{\gamma} \right) (p-q) \left(1 - pC((1-\gamma)pq)\right) \right) \frac{b}{\tau_0} \ .$$

After a difficulty adjustment, the apparent hashrate $\tilde{q}_{EFSM}$ of the miner is

$$\tilde{q}_{EFSM} = \frac{q}{p} - \frac{(1-\gamma)(p-q)}{\gamma p} \left(1 - pC((1-\gamma)pq)\right) \ .$$

We compare these strategies to "Honest Mining" (HM) and "Selfish Mining" (SM) and we determine in the (q, γ) parameter plane which one performs the best.

2. Generalities.

2.1. **Profitability.** In [3], we studied the profitability of integrable repetition games which are composed of cycles, with a finite expected duration $\mathbb{E}[\tau] < +\infty$. The stopping time τ is also called a *strategy*. Mining strategies are repetition games, and sound mining strategies are integrable. For the comparison of profitability of two strategies we only need to compare the Revenue Ratio of each one (see [3])

$$\Gamma = \frac{\mathbb{E}[R]}{\mathbb{E}[\tau]}$$

where R is the revenue over a cycle.

The number of blocks $N'(t)$ and $N(t)$ validated by the attacker and honest miners respectively are independent Poisson processes (see [8] for background on Poisson Process). The attack cycle ends when the honest miners catch-up the attackers. The number of validated blocks in a cycle in the official blockchain is $N(\tau) \vee N'(\tau)$. We denote by $T_1, T_2, \ldots$ (resp. $T'_1, T'_2, \ldots$) the inter-block validation time for the honest miners (resp. attackers).

2.2. **Profitability after a difficulty adjustment.** The following Theorem describes how the Revenue Ratio changes after a difficulty adjustment (see [3]).

Theorem 2.1. *After a difficulty adjustment the new Revenue Ratio $\tilde{\Gamma}$ is given by*

$$\tilde{\Gamma} = \Gamma \cdot \delta \ ,$$

where

$$\delta = \frac{\mathbb{E}[\tau]}{\tau_0 \cdot \mathbb{E}[N(\tau) \vee N'(\tau)]} \ .$$

Proof. After a difficulty adjustment, the expected revenue is the same, $\mathbb{E}[\tilde{R}] = \mathbb{E}[R]$, but the expected duration of a cycle is $\mathbb{E}[\tilde{\tau}] = \tau_0 \cdot \mathbb{E}[N(\tau) \vee N'(\tau)]$. So, the new revenue ratio is $\tilde{\Gamma}(\tau) = \Gamma(\tau) \cdot \delta$. □

2.3. **Description of Stubborn strategies.** We describe first the Selfish Mining (SM) strategy. Let $\Delta \geq 0$ be the advance of the secret fork over the public blockchain. When the honest miners validate a block then the selfish miner does the following:

- If $\Delta = 0$, he mines normally.
- If $\Delta = 1$ then he broadcasts his block. A competition follows.
- If $\Delta = 2$ then he broadcasts his secret fork.
- If $\Delta \geq 3$ then he broadcasts blocks from his secret fork to match the length of the public blockchain.
- Except in the first two cases, he keeps working on top of his secret fork.

For the Lead-Stubborn Mining (LSM) strategy, with $\Delta \geq 2$ he proceeds as in the SM strategy for $\Delta \geq 3$ and with $\Delta = 1$ he releases all his secret fork and mines normally on top of it.

In other words, a stubborn miner following the Lead Stuborn Mining strategy (LSM) waits until the honest miners catch up with him to broadcast all of his secret fork. Then, when this happens, there is a final round. Notice that a selfish miner following SM strategy does not take the risk of being caught by the honest miners. If his advance shrinks to 1, then he broadcasts his fork.

For the Equal Fork Stubborn Mining (EFSM) strategy, everything is equal to LSM, but for $\Delta = 1$ if he finds a new block he does not reveal it.

In other words a stubborn miner following the Equal Fork Stubborn Mining strategy (EFSM) waits for the official blockchain to overcome his secret fork by one block. He only gives up when the length of the official blockchain equals the length of his secret fork plus one. In particular, the last round of the attack cycle of the strategy is always lost by such a miner. His reward comes only when blocks of the official blockchain are built by honest miners on top of one of his blocks. Indeed, the rogue miner never adds new blocks to the official blockchain. He only tries to replace old blocks mined by the honest miners with some of his blocks.

3. Lead-Stubborn Mining strategy

3.1. **Stopping time.** For each mining strategy we consider the stopping time associated with an attack cycle. Let τ_{LSM} be the stopping time of the "Lead-Stubborn" Mining strategy.

Proposition 3.1. *We have*

$$\tau_{LSM} = \tau + (T_{N(\tau)+1} \wedge T'_{N(\tau)+1}) \cdot \mathbf{1}_{T'_1 \leq T_1}$$

with

$$\tau = \inf\{t \geq T_1; N(t) = N'(t) + \mathbf{1}_{T_1 < T'_1}\} \ .$$

In other words, either $T_1 < T'_1$ and the attack cycle ends at T_1 or $T'_1 \leq T_1$ and the attack cycle ends up a final round after the honest miner catch up with the attacker.

We calculate the expected duration time of an attack cycle.

Lemma 3.2. *We have*

$$\mathbb{E}[\tau] = \frac{p}{p-q}\,\tau_0$$

and

$$\mathbb{E}[\tau_{LSM}] = \mathbb{E}[\tau] + q\tau_0 = \tau_0 + 2q\mathbb{E}[\tau] = \frac{p + pq - q^2}{p-q}\tau_0 \ .$$

Proof. By the strong Markov property (see [8]), we have

$$\begin{aligned}\mathbb{E}[\tau] &= \mathbb{E}[\tau|T_1 < T_1'] \cdot \mathbb{P}[T_1 < T_1'] + \mathbb{E}[\tau|T_1 > T_1'] \cdot \mathbb{P}[T_1 > T_1'] \\ &= \mathbb{E}[T_1|T_1 < T_1'] \cdot \mathbb{P}[T_1 < T_1'] + \mathbb{E}[T_1' + \tilde{\tau}|T_1 > T_1'] \cdot \mathbb{P}[T_1 > T_1'] \\ &= \mathbb{E}[T_1 \wedge T_1'] + \mathbb{E}[\tilde{\tau}] \cdot q\end{aligned}$$

where $\tilde{\tau} = \inf\{t; \tilde{N}(t) = \tilde{N}'(t) + 1\}$ with $\tilde{N}(t) = N(t + T_1') - N(T_1')$ and $\tilde{N}'(t) = N'(t + T_1') - N'(T_1')$. Both $\tilde{N}$ and $\tilde{N}'$ are Poisson processes with parameters α and α'. Thus, from Appendix 2 on Poisson Games, we have

$$\mathbb{E}[\tilde{\tau}] = \frac{\tau_0}{p - q}$$

and

$$\mathbb{E}[\tau] = \tau_0 + \frac{q}{p - q}\tau_0 = \frac{p}{p - q}\tau_0$$

If $T_1 < T_1'$ then $\tau_{LSM} = \tau = T_1$. Otherwise (and this event occurs with probability q), once the honest miners catch-up with the attacker at τ-time, there is a final round. □

Note that if $T_1 < T_1'$ then, $\tau_{LSM} = \tau = T_1$ and $N'(\tau) = 0$. Otherwise, $T_1' \leq T_1$ and $N'(\tau) > 0$.

3.2. **Revenue Ratio.** For $n \geq 0$, we denote by C_n the n-th Catalan number

$$C_n = \frac{1}{2n+1}\binom{2n}{n} = \frac{(2n)!}{n!(n+1)!} \ .$$

We present in Appendix A the combinatorial properties of Catalan numbers used in this article, and the definition of (p, q)-Catalan distributions.

The link between Catalan numbers and Bitcoin appears in the following lemma.

Lemma 3.3. *The random variable $N'(\tau)$ follows the second type (p, q)-Catalan distribution, more precisely, we have $\mathbb{P}[N'(\tau) = 0] = p$ and for $n \in \mathbb{N}^*$,*

$$\mathbb{P}[N'(\tau) = n] = C_{n-1}(pq)^n \ .$$

Proof. For $n = 0$ we have $\mathbb{P}[N'(\tau) = 0] = p$. Consider $n \geq 1$. The event $\{N'(\tau) = n\}$ is the disjoint union of sub-events of the form $\{\Sigma_1 < \ldots < \Sigma_{2n+1}\}$ where for each Σ_i there is j such that $\Sigma_i \in \{S_j, S_j'\}$, and $\Sigma_{2n+1} = S_{n+1}'$. The sequence of points with coordinates $(N(\Sigma_i), N'(\Sigma_i))$ form a path starting at $(0, 0)$ and ending at (n, n) which stays strictly above the first bisector $\{x = y\}$ in the Euclidean plane. For example, with $n = 3$,

$$\{N'(\tau) = 3\} = \{S_1' < S_2' < S_3' < S_1 < S_2 < S_3 < S_4'\} \cup \{S_1' < S_2' < S_1 < S_3' < S_2 < S_3 < S_4'\}$$

The number of such paths is C_{n-1} (see Proposition A.4 in Appendix A). Moreover, the number of S_j (resp. S'_j) in the sequence of $(\Sigma_k)_{1\leq k\leq 2n}$ is equal to n. □

From the expected value computation of a second type (p,q)-Catalan random variable (Proposition A.2 in Appendix A) we get:

Corollary 3.4. *We have*

$$\mathbb{E}[N'(\tau)] = \frac{pq}{p-q} \ .$$

Remark 3.5. An alternative probabilistic proof follows from Doob's Theorem: We have $\mathbb{E}[N'(\tau\wedge t)] = \alpha'\mathbb{E}[\tau\wedge t]$ for $t>0$ and we let $t\to+\infty$ (similar to Theorem C.1 in Appendix C).

Next we compute the expected revenue per attack cycle.

Proposition 3.6. *Let R_{LSM} be the revenue over an attack cycle. We have*

$$\mathbb{E}[R_{LSM}] = \left(\frac{p}{p-q}+q\right)qb - f(\gamma)b$$

with

$$f(\gamma) = \frac{pq(1-\gamma)}{\gamma}\cdot(1-p(1-\gamma)C((1-\gamma)pq)) \ .$$

Proof. Note that for $n>0$,

$$\begin{aligned}
\mathbb{P}[R_{LSM}=nb|N'(\tau)=n] &= \gamma p\\
\mathbb{P}[R_{LSM}=(n+1)b|N'(\tau)=n] &= q\\
\mathbb{P}[R_{LSM}<nb|N'(\tau)=n] &= (1-\gamma)p
\end{aligned}$$

Moreover, during an attack cycle, each time the honest miners find a block (except for the first block which is mined on a common root), there is a probability γ that it is found by a miner mining on top of the attacker's fork. If $N'(\tau)=n$ and $R(\tau_{LSM})<nb$, this can happen at most $n-1$ times over an attack cycle. So, by Lemma B.1 from Appendix B,

$$\mathbb{E}[R_{LSM}|(N'(\tau)=n)\wedge(R_{LSM}<nb)] = \left(n-\frac{1-(1-\gamma)^n}{\gamma}\right)b$$

Therefore, by conditioning on τ and using Lemma 3.2, Lemma 3.3 and Corollary 3.4, we get

$$\begin{aligned}
\frac{\mathbb{E}[R_{LSM}]}{b} &= \sum_{n>0} \mathbb{E}\left[\frac{R_{LSM}}{b}\middle| N'(\tau)=n\right]\cdot \mathbb{P}[N'(\tau)=n] \\
&= \sum_{n>0}\left(\left(n-\frac{1-(1-\gamma)^n}{\gamma}\right)\cdot(1-\gamma)p+n\gamma p+(n+1)q\right)\cdot \mathbb{P}[N'(\tau)=n] \\
&= \sum_{n>0}\left(n+1-\frac{p}{\gamma}+\frac{(1-\gamma)^{n+1}}{\gamma}p\right)\cdot \mathbb{P}[N'(\tau)=n] \\
&= \mathbb{E}[N'(\tau)]+\left(1-\frac{p}{\gamma}\right)(1-\mathbb{P}[N'(\tau)=0])+\frac{p^2q(1-\gamma)^2}{\gamma}\sum_{n>0}C_{n-1}((1-\gamma)pq)^{n-1} \\
&= \frac{pq}{p-q}+\left(1-\frac{p}{\gamma}\right)q+\frac{p^2q(1-\gamma)^2}{\gamma}C((1-\gamma)pq) \\
&= \left(\frac{p}{p-q}+q\right)q-\frac{pq}{\gamma}\cdot(1-\gamma-p(1-\gamma)^2C((1-\gamma)pq))
\end{aligned}$$

□

Theorem 3.7. *The revenue ratio of the Lead Stubborn strategy is*

$$\Gamma(LSM)=\left(q-\frac{pq(p-q)(1-\gamma)}{\gamma}\cdot\frac{1-p(1-\gamma)C((1-\gamma)pq)}{p+q(p-q)}\right)\frac{b}{\tau_0}\ .$$

Proof. We have $\Gamma(LSM)=\mathbb{E}[R_{LSM}]/\mathbb{E}[\tau_{LSM}]$. Use Lemma 3.2 and Proposition 3.6. □

3.3. **Difficulty adjustment.** We compute now the revenue ratio and the apparent hashrate after a difficulty adjustment.

Lemma 3.8. *We have* $\mathbb{E}[N(\tau_{LSM})\vee N'(\tau_{LSM})]=\frac{\mathbb{E}[\tau_{LSM}]}{2\tau_0}+\frac{1}{2}$.

Proof. For all $t>0$, $\tau_{LSM}\wedge t$ is a bounded stopping time. So, by proceeding as in Appendix 2, Doob's theorem yields $\mathbb{E}[N(\tau_{LSM}\wedge t)]=\alpha\mathbb{E}[\tau_{LSM}\wedge t]$ and $\mathbb{E}[N'(\tau_{LSM}\wedge t)]=\alpha'\mathbb{E}[\tau_{LSM}\wedge t]$. Taking limits when $t\to 0$, the monotone convergence theorem yields $\mathbb{E}[N(\tau_{LSM})]=\alpha\mathbb{E}[\tau_{LSM}]$ and $\mathbb{E}[N'(\tau_{LSM})]=\alpha'\mathbb{E}[\tau_{LSM}]$. Now, we observe that at the end of an attack cycle, we have necessarily $|N(\tau_{LSM})-N'(\tau_{LSM})|=1$. So, $N(\tau_{LSM})\vee N'(\tau_{LSM})=\frac{N(\tau_{LSM})+N'(\tau_{LSM})+1}{2}$. Hence we get the result by taking expected values on both sides of the last equality. □

The following proposition is now a consequence of Theorem 2.1.

Proposition 3.9. *The parameter δ_{LSM} updating the difficulty is*

$$\delta_{LSM} = \frac{p + pq - q^2}{p + pq - q} > 1 \ .$$

Proof. Using Theorem 2.1, Lemma 3.2 and Lemma 3.8, we have

$$\delta_{LSM} = \frac{\frac{\mathbb{E}[\tau_{LSM}]}{\tau_0}}{\frac{1}{2}\left(\frac{\mathbb{E}[\tau_{LSM}]}{\tau_0} + 1\right)} = \frac{\frac{\mathbb{E}[\tau]}{\tau_0} + q}{1 + q\frac{\mathbb{E}[\tau]}{\tau_0}} = \frac{\frac{p}{p-q} + q}{1 + \frac{pq}{p-q}} = \frac{p + pq - q^2}{p + pq - q} > 1$$

□

3.4. **Apparent hashrate after a difficulty adjustment.** From Proposition 3.9, we can deduce the hashrate of the strategy on the long term.

Corollary 3.10. *After a difficulty adjustment, the apparent hashrate $\tilde{q}_{LSM}$ is*

$$\tilde{q}_{LSM} = q \cdot \frac{p + pq - q^2}{p + pq - q} - \frac{pq(p-q)(1-\gamma)}{\gamma} \cdot \frac{1 - p(1-\gamma)C((1-\gamma)pq)}{p + pq - q}$$

Proof. By Theorem 2.1, we have:

$$\begin{aligned}
&\tilde{\Gamma}(LSM) = \Gamma(LSM)\delta_{LSM} \\
&= \left(q - \frac{pq(p-q)(1-\gamma)}{\gamma} \cdot \frac{1 - p(1-\gamma)C((1-\gamma)pq)}{p + q(p-q)}\right)\frac{b}{\tau_0} \cdot \frac{p + pq - q^2}{p + pq - q} \\
&= \left(q \cdot \frac{p + pq - q^2}{p + pq - q} - \frac{pq(p-q)(1-\gamma)}{\gamma} \cdot \frac{1 - p(1-\gamma)C((1-\gamma)pq)}{p + pq - q}\right)\frac{b}{\tau_0}
\end{aligned}$$

□

4. Equal Fork Stubborn Mining strategy

4.1. **Stopping time.** Let τ_{EFSM} be the stopping time of an attack cycle for the Equal Fork Stubborn Mining strategy.

Proposition 4.1. *We have*

$$\tau_{EFSM} = \inf\{t \geq 0; N(t) = N'(t) + 1\} \ .$$

Proof. If $T_1 < T_1'$, then we have $\tau_{EFSM} = \tau = T_1$. Otherwise, we wait for the honest miners to catch up with the stubborn miner and win the last round. □

From Theorem C.1 in Appendix C we get:

Lemma 4.2. *We have*

$$\mathbb{E}[\tau_{EFSM}] = \frac{\tau_0}{p-q} \ .$$

4.2. **Revenue ratio.** We denote by R_{EFSM} the revenue of the stubborn miner after an attack cycle.

Lemma 4.3. *The random variable $N'(\tau_{EFSM})$ is a (p,q)-Catalan distribution, i.e. for $n \geq 0$, we have*

$$\mathbb{P}[N'(\tau_{EFSM}) = n] = C_n \, p(pq)^n \ .$$

Proof. The event $\{N'(\tau_{EFSM}) = n\}$ can be decomposed as a disjoint union of sub-events of the form $\{\Sigma_1 < \ldots < \Sigma_{2n+1} < \Sigma_{2n+2}\}$ where for each Σ_i there is j such that $\Sigma_i \in \{S_j, S'_j\}$, and $\Sigma_{2n+1} = S_{n+1}$ and $\Sigma_{2n+2} = S'_{n+1}$. The sequence of points with coordinates $(N(\Sigma_i), N'(\Sigma_i))$ for $i \in \{1, \ldots, 2n+1\}$ form a path starting at $(0,0)$ and ending at $(n+1, n)$ which never crosses the first diagonal $\{x = y\}$ in the Euclidean plane before reaching the point $(n+1, n)$. For example,

$$\{N'(\xi) = 2\} = \{S'_1 < S'_2 < S_1 < S_2 < S_3 < S'_3\} \cup \{S'_1 < S_1 < S'_2 < S_2 < S_3 < S'_3\}$$

The number of such paths is C_n (see Proposition A.4 from Appendix A). Moreover, the number of S_j (resp. S'_j) in the sequence of $(\Sigma_k)_{1 \leq k \leq 2n+1}$ is equal to $n+1$ (resp. n). Hence we get the result. □

From the expected value computation of a (p,q)-Catalan random variable (Proposition A.2 in Appendix A) we get:

Corollary 4.4. *We have*

$$\mathbb{E}[N'(\tau_{EFSM})] = \frac{q}{p-q} \ .$$

We compute now $\mathbb{E}[R_{EFSM}]$.

Proposition 4.5. *We have*

$$\mathbb{E}[R_{EFSM}] = \frac{q}{p-q} b - g(\gamma) b$$

with

$$g(\gamma) = \frac{1-\gamma}{\gamma} \left(1 - pC((1-\gamma)pq)\right) \ .$$

Proof. By definition of τ_{EFSM}, if we know that $N'(\tau_{EFSM}) = n$, then the honest miners have mined $n+1$ blocks at time $t = \tau_{EFSM}$. Except for the first block, the

probability that the block validated by the honest miners is mined on a fork created by the stubborn miner is γ. So, by Lemma B.1 from Appendix B,

$$\mathbb{E}\left[\frac{R_{EFSM}}{b}\middle| N'(\tau_{EFSM})=n\right]=n+1-\frac{1-(1-\gamma)^{n+1}}{\gamma}\ .$$

Therefore, we have

$$\begin{aligned}
&\frac{\mathbb{E}[R_{EFSM}]}{b}=\sum_{n\geq 0}\mathbb{E}\left[\frac{R_{EFSM}}{b}\middle| N'(\tau_{EFSM})=n\right]\cdot\mathbb{P}[N'(\tau_{EFSM})=n]\\
&=\sum_{n\geq 0}\left(n+1-\frac{1-(1-\gamma)^{n+1}}{\gamma}\right)\cdot\mathbb{P}[N'(\tau_{EFSM})=n]\\
&=\sum_{n\geq 0}n\,\mathbb{P}[N'(\tau_{EFSM})=n]+1-\frac{\gamma-1}{\gamma}\sum_{n\geq 0}\mathbb{P}[N'(\tau_{EFSM})=n]+\frac{1-\gamma}{\gamma}p\sum_{n\geq 0}((1-\gamma)pq)^nC_n\\
&=\mathbb{E}[N'(\tau_{EFSM})]-\frac{1-\gamma}{\gamma}+\frac{1-\gamma}{\gamma}pC((1-\gamma)pq)\\
&=\frac{q}{p-q}-\frac{1-\gamma}{\gamma}\left(1-pC((1-\gamma)pq)\right)\ .
\end{aligned}$$

□

Theorem 4.6. *The revenue ratio of the "Equal Fork Stubborn mining" strategy is*

$$\Gamma(EFSM)=\left(q-\frac{1-\gamma}{\gamma}(p-q)\left(1-pC((1-\gamma)pq)\right)\right)\frac{b}{\tau_0}\ .$$

Proof. Use Lemma 4.2 and Proposition 4.5. □

4.3. **Difficulty adjustment.**

Proposition 4.7. *The parameter δ_{EFSM} updating the difficulty is*

$$\delta_{EFSM}=\frac{1}{p}>1\ .$$

Proof. At the end of an attack cycle, the number of new blocks in the official blockchain is $N(\tau_{EFSM})$ and $\mathbb{E}[N(\tau_{EFSM})]=\alpha\mathbb{E}[\tau_{EFSM}]$. Therefore we have

$$\delta_{EFSM}=\frac{\frac{\mathbb{E}[\tau_{EFSM}]}{\tau_0}}{\mathbb{E}[N(\tau_{EFSM})]}=\frac{1}{p}\ .$$

□

4.4. **Apparent hashrate after a difficulty adjustment.** It's now easy to get the long term apparent hashrate of the EFSM strategy.

Corollary 4.8. *After a difficulty adjustment, the apparent hashrate is*

$$\tilde{q}_{EFSM} = \frac{q}{p} - \frac{(1-\gamma)(p-q)}{\gamma p}\left(1 - p\,C((1-\gamma)pq)\right) \ .$$

Proof. Use Theorem 2.1 and Proposition 4.7. □

5. Comparison of strategies.

For different values of the parameters q and γ, we can compare the profitability of the different strategies after a difficulty adjustment by comparing the revenue ratios after a difficulty adjustment, or, equivalently, their apparent hashrate.

We consider the four strategies:

- Honest Mining (HM).
- Selfish Mining (SM).
- Lead-Stubborn Mining (LSM).
- Equal Fork Stubborn mining (EFSM).

We color the region $(q, \gamma) \in [0, 0.5] \times [0, 1]$ according to which strategy is more profitable, and we obtain Figure 1.

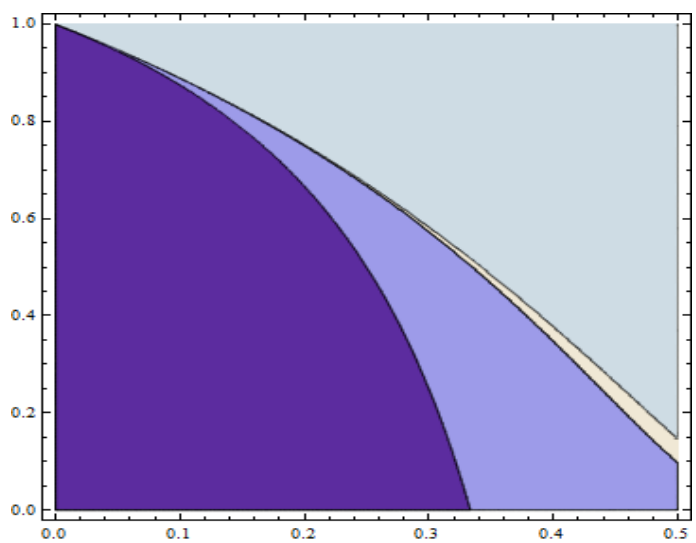

Figure 1. Profitability. From left to right: HM, SM, LSM, EFSM.

From left to right, the best strategy is successively HM, SM, LSM and EFSM. Note that, the LSM strategy is superior only on a thin domain.

A similar picture is numerically computed by Montecarlo simulations in [7] where only a numerical study is carried out. See Figure 2 (disregard regions R4 to R7 that correspond to "Trail T_j Stubborn Mining" strategies that are studied in [4]).

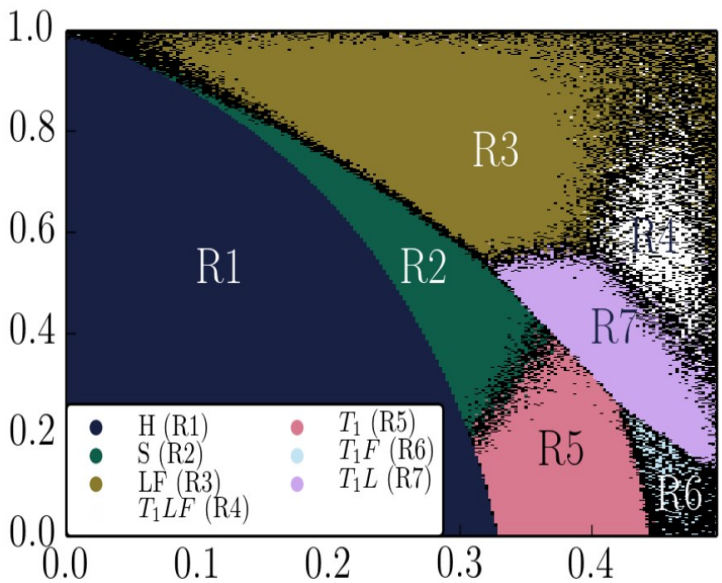

FIGURE 2. Figure from [7].

Regions R1, R2 and R3 correspond respectively to HM, SM and EFSM. We notice the absence of the region corresponding to LSM in between R2 and R3. Apparently the numerical methods did not detect it (also it is noticeable how in their figure the boundary between R2 and R3 is more blurred than other boundaries, as for example the one separating R1 and R2).

APPENDIX A. CATALAN DISTRIBUTIONS.

For $n \geq 0$, we denote by C_n the n-th Catalan number

$$C_n = \frac{1}{2n+1}\binom{2n}{n} = \frac{(2n)!}{n!(n+1)!} \ .$$

We refer to [5] for background and combinatorial properties of Catalan numbers. Their generating series is

$$C(x) = \sum_{n=0}^{+\infty} C_n x^n = \frac{1-\sqrt{1-4x}}{2x} = \frac{2}{1+\sqrt{1-4x}} \ .$$

Observe that $\sqrt{1-4pq} = p-q$ and $C(pq) = 1/p$, which justifies the following definition:

Definition A.1 (Catalan distributions)**.** *Let $1/2 < p < 1$ and $q = 1-p$. A random variable X taking values in $\mathbb{N}$ is a (p,q)-Catalan random variable if it follows the (p,q)-Catalan distribution, that is, for $n \geq 0$*

$$\mathbb{P}[X = n] = C_n p (pq)^n \ .$$

The second type (p,q)-Catalan distribution is defined by $\mathbb{P}[X=0] = p$ and for $n \geq 1$,

$$\mathbb{P}[X = n] = C_{n-1} (pq)^n \ .$$

Proposition A.2. *The expected value of a (p,q)-Catalan random variable X is*

$$\mathbb{E}[X] = \frac{q}{p-q}$$

If X is a second type (p,q)-Catalan random variable then

$$\mathbb{E}[X] = \frac{pq}{p-q} \ .$$

Lemma A.3. *We have*

$$\frac{d}{dx}\left(xC(x)\right)\Big|_{x=pq} = \frac{1}{p-q} \ .$$

Proof of the Lemma. We have $\frac{d}{dx}\left(xC(x)\right) = (1-4x)^{-1/2}$ and the result follows. □

Proof of the Proposition. For the (p,q)-Catalan random variable we have

$$\begin{aligned}
\mathbb{E}[X] &= \sum_{n=0}^{+\infty} n.\mathbb{P}[X=n] = \sum_{n=0}^{+\infty} n.C_n p (pq)^n \\
&= p \sum_{n=0}^{+\infty} C_n (n+1)(pq)^n - p \sum_{n=0}^{+\infty} C_n (pq)^n \\
&= p \frac{d}{dx}\left(xC(x)\right)\Big|_{x=pq} - p\, C(pq) \\
&= \frac{p}{p-q} - 1 = \frac{q}{p-q} \ .
\end{aligned}$$

For the second type (p,q)-Catalan random variable we have

$$\begin{aligned}\mathbb{E}[X] &= \sum_{n=0}^{+\infty} n.\mathbb{P}[X=n] = \sum_{n=1}^{+\infty} n.C_{n-1}(pq)^n \\ &= \sum_{m=0}^{+\infty} C_m(m+1)(pq)^m = pq\,\frac{d}{dx}\left(xC(x)\right)\Big|_{x=pq} \\ &= \frac{pq}{p-q} \quad .\end{aligned}$$

□

The main geometric combinatorial property of Catalan numbers used in this article is the following well known enumeration (see section 9 of [5] p.259)

Proposition A.4. *The number of paths in $\mathbb{N}^2$ going up and right at each step that start at $(0,0)$ and end-up at $(n+1,n+1)$ without touching the first diagonal is C_n.*

Appendix B. Biased coin tossing.

The following lemma is useful for both LSM and EFSM strategies.

Lemma B.1. *Let $0<\gamma<1$ and $n\in\mathbb{N}$. Let $\omega\in\{0,1\}^n$ denotes the outcome of tossing a biased coin n times with $\mathbb{P}[\omega_i=1]=\gamma$ for $i\in\{1,\dots n\}$. Let $Z(\omega)=\sup\{i\in\{1,\dots n\};\omega_i=1\}\cup\{0\}$. Then we have $\mathbb{E}[Z]=n+1-\frac{1-(1-\gamma)^{n+1}}{\gamma}$.*

Proof. Note that $n+1-Z=\inf\{i\in\{1,\dots,n\};\omega_{n+1-i}=1\}\cup\{n+1\}=\tilde{Z}\wedge(n+1)$ where $\tilde{Z}$ is the number of trials (stopping time) before getting "Heads" when a coin is flipped repeatedly with a probability γ of getting "Heads" each time. We have

$$\begin{aligned}\mathbb{E}[\tilde{Z}\wedge(n+1)] &= \sum_{i=1}^{n} i\mathbb{P}[\tilde{Z}=i]+(n+1)\mathbb{P}[\tilde{Z}>n] \\ &= \sum_{i=1}^{n} i(\mathbb{P}[\tilde{Z}>i-1]-\mathbb{P}[\tilde{Z}>i])+(n+1)\mathbb{P}[\tilde{Z}>n] \\ &= \sum_{i=0}^{n}\mathbb{P}[\tilde{Z}>i]=\sum_{i=0}^{n}(1-\gamma)^i=\frac{1-(1-\gamma)^{n+1}}{\gamma}\end{aligned}$$

□

Appendix C. Poisson Games.

Theorem C.1. *Let $N(t)$ (resp. $N'(t)$) be a Poisson process with parameter α (resp. α'). Let τ be the stopping time defined by $\tau = \inf\{t \in \mathbb{R}_+; N(t) = N'(t) + 1\}$. If $\alpha > \alpha'$ then $\tau \in L^1, N(\tau) \in L^1, \mathbb{E}[\tau] = \frac{1}{\alpha-\alpha'}$ and $\mathbb{E}[N(\tau)] = \frac{\alpha}{\alpha-\alpha'}$.*

Proof. The proof is similar to the proof of Theorem 4.4 from [2]. For $t \geq 0$, the stopping time $\tau \wedge t$ is bounded. So, by applying Doob's Theorem ([9]) to the martingales $N(t) - \alpha t$ and $N'(t) - \alpha' t$, we get

$$\begin{aligned}
\alpha\mathbb{E}[\tau \wedge t] &= \mathbb{E}[N(\tau \wedge t)] \\
&= \mathbb{E}[N(\tau \wedge t)|\tau \leq t]\mathbb{P}[\tau \leq t] + \mathbb{E}[N(\tau \wedge t)|\tau > t]\mathbb{P}[\tau > t] \\
&= \mathbb{E}[N(\tau)|\tau \leq t]\mathbb{P}[\tau \leq t] + \mathbb{E}[N(t)|\tau > t]\mathbb{P}[\tau > t] \\
&= \mathbb{E}[N'(\tau) + 1|\tau \leq t]\mathbb{P}[\tau \leq t] + \mathbb{E}[N(t)]\mathbb{P}[\tau > t] \\
&= \mathbb{E}[N'(\tau)|\tau \leq t]\mathbb{P}[\tau \leq t] + \alpha t\mathbb{P}[\tau > t] + \mathbb{P}[\tau \leq t] \\
&= \mathbb{E}[N'(\tau \wedge t)|\tau \leq t]\mathbb{P}[\tau \leq t] + \mathbb{E}[N'(\tau \wedge t)|\tau > t]\mathbb{P}[\tau > t] \\
&- \mathbb{E}[N'(\tau \wedge t)|\tau > t]\mathbb{P}[\tau > t] + \alpha t\mathbb{P}[\tau > t] + \mathbb{P}[\tau \leq t] \\
&= \mathbb{E}[N'(\tau \wedge t)] - \mathbb{E}[N'(t)|\tau > t]\mathbb{P}[\tau > t] + \alpha t\mathbb{P}[\tau > t] + \mathbb{P}[\tau \leq t] \\
&= \alpha'\mathbb{E}[\tau \wedge t] - \alpha' t\mathbb{P}[\tau > t] + \alpha t\mathbb{P}[\tau > t] + \mathbb{P}[\tau \leq t]
\end{aligned}$$

So we have $(\alpha - \alpha')(\mathbb{E}[\tau \wedge t] - t\mathbb{P}[\tau > t]) = \mathbb{P}[\tau \leq t]$ and

$$(\alpha - \alpha')\mathbb{E}[\tau \wedge t|\tau \leq t]\mathbb{P}[\tau \leq t] = (\alpha - \alpha')(\mathbb{E}[\tau \wedge t] - t\mathbb{P}[\tau > t]) = \mathbb{P}[\tau \leq t] \ .$$

Therefore,

$$\mathbb{E}[\tau \mathbf{1}_{\tau \leq t}] = \frac{\mathbb{P}[\tau \leq t]}{\alpha - \alpha'} \ .$$

Making $t \to \infty$, by monotone convergence we have $\tau \in L^1$ and $\mathbb{E}[\tau] = \frac{1}{\alpha-\alpha'}$. Moreover, using Doob's Theorem again, we have for $t > 0$,

$$\begin{aligned}
\mathbb{E}[N(\tau)\mathbf{1}_{\tau \leq t}] &= \mathbb{E}[N(\tau \wedge t)] - \mathbb{E}[N(t)|\tau > t]\mathbb{P}[\tau > t] \\
&= \mathbb{E}[\tau \wedge t] - \alpha t\mathbb{P}[\tau > t] \\
&= \alpha\mathbb{E}[\tau \mathbf{1}_{\tau \leq t}]
\end{aligned}$$

So, using the monotone convergence theorem again, we get $\mathbb{E}[N(\tau)] = \frac{\alpha}{\alpha-\alpha'}$ □

References

[1] I. Eyal, E. G. Sirer. *Majority is not enough: bitcoin mining is vulnerable*, Int. Conf. Financial Cryptography and Data Security, Springer, p.436-454, 2014.
[2] C. Grunspan and R. Pérez-Marco. *Double spend races*, arXiv:1702.02867, 2017.
[3] C. Grunspan and R. Pérez-Marco. *On profitability of Selfish Mining*, arXiv:1805.08281, 2018.
[4] C. Grunspan and R. Pérez-Marco. *On profitability of Trailing Mining*, ArXiv, 2018.
[5] T. Koshy. *Catalan Numbers with Applications*, Oxford University Press, 2008.
[6] S. Nakamoto. *Bitcoin: a peer-to-peer electronic cash system*, Bitcoin.org, 2008.
[7] K. Nayak, S. Kumar, A. Miller, E. Shi. *Stubborn Mining: Generalizing Selfish Mining and Combining with an Eclipse Attack*, 2016 IEEE Europ. Symp. on Security and Privacy, 2016.
[8] S. Ross. *Introduction to Probability Models*, 10th Edition, Academic Press Inc, 2012.
[9] M. Williams. *Probability with Martingales*, 5th Edition, Cambridge University Press, 1991.

Cyril Grunspan
Léonard de Vinci Pôle Univ, Research Center, Labex Réfi
Paris, France,

E-mail address: cyril.grunspan@devinci.fr

Ricardo Pérez-Marco
CNRS, IMJ-PRG, Labex Réfi
Paris, France

E-mail address: ricardo.perez.marco@gmail.com

Author's Bitcoin Beer Address (ABBA)[1]:
1KrqVxqQFyUY9WuWcR5EHGVvhCS841LPLn

[1]Send some anonymous and moderate satoshis to support our research at the pub.