Suppose that I control $p$ hash power. My goal is to ensure that the entire blockchain follows my rules, almost all of the time. The remaining hashpower is $1-p$. Let's suppose these are rational, free agents, that have no creed.

I declare the following rule for everyone under my command: Append to the longest chain that does not have a non-compliant block in the last 2 blocks.

Suppose that the game runs for $k$, some large number blocks. We want to figure out the expected returns for a miner who mines a C block, and compare to the miner mining a NC, whenever the two chains have the same length. Note that these two probabilities are identical, but the the miner gets to choose which one to pursue.

This game terminates when either the C chain pulls ahead by 1, or the NC chains pulls ahead by 2. Then over the remaining blocks, which are all mined by everybody, the expected value is $\varepsilon R$ per block.

Let $f$ be reward given to the miner. The expected value is

$$\int_X f d\mu$$

where $X$ is the probability space of all events.

Suppose the miner mines with NC and wins the first block putting the NC ahead by one.

From this starting point, let $E_i$ be the event that the game terminates at step $i$, and let $E_{k+1}$ be the event that the game hasn't terminated (we will let $k \to \infty$, but for now let's keep it finite.) It follows that $\sum_{i=1}^{k+1} P(E_i) = 1$

We will use the following assumption: *The free agent miners will all follow the same strategy, namely to mine the non-conformal chain.* I don't have a proof of this yet, but I would geuss that this is the most favorable strategy for the miner in question, given they want the block to be confirmed. If we can show that even under the most favorable assumptions, the miner would have been better off mining the C chain, then we conclude the miner would do such.

Now because

$$X = \bigsqcup E_i$$

we can compute

$$\int_X f = \sum \int_{E_i} f$$

That is, we can compute the expected value of the function defined by

$$\begin{aligned} f_i &= \text{total reward : if game terminates at step i} \\ f_i &= 0 \text{ otherwise} \end{aligned}$$

One can check that

$$\sum f_i = f.$$

Now we start with $E_1$

$$P(E_1) = 1 - p$$

This is the probability that a miner from the NC pool mines the next block, putting NC chain 2 ahead. This is really two situations, the probablity $(1-p-\varepsilon)$ event that sum other mines it, and then $\varepsilon$ probability that the miner in question mines the second block, for their second reward. The reward table is as follows

| Probability | Reward |
|---|---|
| $1-p-\varepsilon$ | $1R$ |
| $\varepsilon$ | $2R$ |

It follows that

$$\begin{aligned} E(f_1) &= \left((1-p)+\varepsilon\right)R + (1-p)\,(k-1)\,\varepsilon R \\ &= (1-p)\left(1+\frac{\varepsilon}{1-p}+(k-1)\,\varepsilon\right)R \end{aligned}$$

where the second term comes from the expect winnings over the next $k-1$ block multiplied by the probability that of $E_1$.

Now look at events $E_2$. This can only happen if the second block is mined by C miner, followed by another block mined by C miner. This has probability $p^2$. In this case, the reward is zero in the first two block, but, the miner still can jump back for the remainig $k-2$ so

$$E(f_2) = p^2(k-2)\varepsilon R$$

Now move to $E_3$. There is only one path - the second block must have broke for C, followed by two consecutive blocks being mined for NC. Denote this C-NC-NC. Breaking this into the possibilities for how the individual miner fared we have

| Probability | Reward |
|---|---|
| $p(1-p-\varepsilon)(1-p-\varepsilon)$ | $1R$ |
| $p\varepsilon(1-p-\varepsilon)$ | $2R$ |
| $p(1-p-\varepsilon)\varepsilon$ | $2R$ |
| $p\varepsilon\varepsilon$ | $3R$ |

Notice that by subtracting $1R$ from each Reward (and then adding back over all probabilities) we get

$$p\,(1-p)^2\sum_{i=0}^{i=2} i\binom{2}{i}(1-\frac{\varepsilon}{1-p})^i\left(\frac{\varepsilon}{(1-p)}\right)^{2-i}$$

which is

$$\sum_{i=0}^{i=2} i(1-\frac{\varepsilon}{1-p})^i\left(\frac{\varepsilon}{(1-p)}\right)^{2-i} = 2\frac{\varepsilon}{(1-p)}$$

using formulas for 2 trials of a Bernoulli random variable

So the full

$$\begin{aligned} E(f_3) &= \left(p\,(1-p)^2 + p\,(1-p)^2\,2\frac{\varepsilon}{(1-p)} + p(1-p)^2\,(k-3)\,\varepsilon\right)R \\ & p(1-p)^2\left(1+2\frac{\varepsilon}{(1-p)}+(k-3)\,\varepsilon\right)R \end{aligned}$$

We add these up, rearranging a bit (leaving out $p$ and $R$)

Now for $f_4$ there is a unique path to this, with probability $p*(1-p)*p*p$

So

$$E(f_4) = p^3(1-p)\varepsilon(k-4)R$$

Now for $f_5$. The path has to be C-NC-C-NC-NC.

As before, we only pay attention to the additional rewards - there's 1 way to get 0 with probability

$$p^2(1-p-\varepsilon)^3$$

There's $\binom{3}{1}$ ways to get one additional rewards each with probability

$$p^2(1-p-\varepsilon)^2\varepsilon$$

There's $\binom{3}{2}$ ways to get two additions rewards, etc, each with probability

$$p^2(1-p-\varepsilon)\varepsilon^2$$

etc.

So we have

$$\begin{aligned} & \sum_{i=0}^{3} i\binom{3}{i}p^2(1-p)^3(1-\frac{\varepsilon}{1-p})^{3-i}\left(\frac{\varepsilon}{1-p}\right)^i \\ = & \quad p^2(1-p)^2 3\frac{\varepsilon}{1-p} \end{aligned}$$

So

$$E(f_5) = p^2(1-p)^3\left(1+3\frac{\varepsilon}{1-p}+(k-5)\varepsilon\right)R$$

Now to summarize (not a proof yet but forumula looks like it's going somewhere obvious)

$$\begin{aligned} E(f_1) &= (1-p)\left(1+\frac{\varepsilon}{1-p}+(k-1)\,\varepsilon\right)R \\ E(f_3) &= p(1-p)^2\left(1+2\frac{\varepsilon}{(1-p)}+(k-3)\,\varepsilon\right)R \\ E(f_5) &= p^2(1-p)^3\left(1+3\frac{\varepsilon}{1-p}+(k-5)\varepsilon\right)R \end{aligned}$$

and

$$\begin{aligned} E(f_2) &= p^2(k-2)\varepsilon R \\ E(f_4) &= p^3(1-p)(k-4)\varepsilon R \end{aligned}$$

so we have

$$\begin{aligned} E(f_{2j}) &= p^{2+j}(1-p)^j(k-2j)\varepsilon R \\ E(f_{2j+1}) &= p^j(1-p)^{j+1}(1+(j+1)\frac{\varepsilon}{(1-p)}+(k-2j-1)\varepsilon)R \end{aligned}$$

Let's try to sum these up

$$\sum_{j=0}^{k/2} p^{2+j}(1-p)^j(k-2j)\varepsilon R$$

$$= \left( \begin{array}{c} p^2 k \sum_{j=0}^{k/2} (p(1-p))^j \\ -p^2 2 \sum_{j=0}^{k/2} (p(1-p))^j j \end{array} \right) \varepsilon R$$

For simplicity we will drop the R (or think of the answer in terms of units of block reward R)

For the first term recall the geometric series formula

$$\sum_{i=0}^{n} a^i = \frac{1-a^{n+1}}{1-a}.$$

For the second term, consider the function

$$b(x) = \frac{1}{(1-x)^2}$$

note that

$$\begin{aligned} b'(x) &= \frac{2}{(1-x)^3} \\ b''(x) &= \frac{3*2}{(1-x)^4} \\ b'''(x) &= \frac{4*3*2}{(1-x)^5} \end{aligned}$$

etc, so that

$$1 + 2x + \frac{3!x^2}{2!} + \frac{4!x^2}{3!} ... = \sum_{j=0}^{\infty} x^j j$$

is the Taylor series for $b$ around $x = 0$, which converges for $|x| < 1$

Thus for large $k$ we have

$$\sum_{j=0}^{k/2} (p(1-p))^j j = \left( \frac{1}{1-p(1-p)} \right)^2 - R_{k/2}(b)$$

where $R_{k/2}$ is the remainder, bounded by

$$\left( \frac{k}{2} + 1 \right) |x|^{\frac{k}{2}+1}$$

So for large $k$ we get

$$\begin{aligned}
&\sum_{j=0}^{k/2} p^{2+j}(1-p)^j(k-2j)\varepsilon \\
= \quad & p^2 k\frac{1}{1-p(1-p)}\varepsilon - p^2 k\frac{1}{1-p(1-p)}\left(p(1-p)\right)^{k/2+1}\varepsilon \\
& +\left(\frac{1}{1-p(1-p)}\right)^2\varepsilon - R_{k/2}\varepsilon
\end{aligned}$$

And

$$\begin{aligned}
\sum_{j=0}^{j=k/2} E(f_{2j+1}) &= \sum_{j=0}^{j=k/2}\left(p^j(1-p)^{j+1}(1+(j+1)\frac{\varepsilon}{(1-p)}+(k-2j-1)\varepsilon)\right) \\
&= (1-p)\sum_{j=0}^{j=k/2} p^j(1-p)^j \\
&\quad + \sum_{j=0}^{j=k/2} p^j(1-p)^j(j+1)\varepsilon \\
&\quad +(1-p)\,(k-1)\,\varepsilon\sum_{j=0}^{j=k/2} p^j(1-p)^j \\
&\quad -(1-p)2\varepsilon\sum_{j=0}^{j=k/2} p^j(1-p)^j j
\end{aligned}$$

Evaluating each in turn

$$\begin{aligned}
(1-p)\sum_{j=0}^{j=k/2} p^j(1-p)^j &= (1-p)\frac{1}{1-p+p^2} - (1-p)\frac{p^{k/2}(1-p)^{k/2}}{1-p+p^2} \\
\sum_{j=0}^{j=k/2} p^j(1-p)^j(j+1)\varepsilon &= \varepsilon\left(\frac{1}{1-p+p^2} - \frac{p^{k/2}(1-p)^{k/2}}{1-p+p^2}\right) \\
&\quad +\varepsilon\left(\frac{1}{\left(1-p+p^2\right)^2} - R_{k/2}\right) \\
(1-p)\,(k-1)\,\varepsilon\sum_{j=0}^{j=k/2} p^j(1-p)^j &= (1-p)\,(k-1)\,\varepsilon\frac{1}{1-p+p^2} - (1-p)\,(k-1)\,\varepsilon\frac{p^{k/2}(1-p)^{k/2}}{1-p+p^2} \\
-(1-p)2\varepsilon\sum_{j=0}^{j=k/2} p^j(1-p)^j j &= -(1-p)2\varepsilon\left(\frac{1}{\left(1-p+p^2\right)^2} - R_{k/2}\right)
\end{aligned}$$

Now each of the above has been separated into term that decays geometrically with $k$ and those who don't. So we sum up the non decaying terms and ignore the decaying term

$$
\begin{aligned}
= \quad & p^2 k \frac{1}{1-p(1-p)} \varepsilon + (1-p) k \varepsilon \frac{1}{1-p+p^2} \\
& + \left( \frac{1}{1-p(1-p)} \right)^2 \varepsilon + \varepsilon \frac{1}{(1-p+p^2)^2} - (1-p) 2 \varepsilon \frac{1}{(1-p+p^2)^2} \\
& +(1-p) \frac{1}{1-p+p^2}
\end{aligned}
$$

$$
\begin{aligned}
= \quad & (1-p) \frac{1}{1-p+p^2} + k\varepsilon \\
& + \frac{2p}{(1-p(1-p))^2} \varepsilon
\end{aligned}
$$

These three terms can be interpreted: The first term which can be written

$$\frac{(1-p)+p^2-p^2}{1-p+p^2} = 1 - \frac{p^2}{1-p+p^2} < 1$$

is the probability of keeping the block earned in round 0. The $k\varepsilon$ is the "steady-state" term that describes return to expected average as the game ends almost surely. Finally,

$$\frac{2p}{(1-p(1-p))^2} \varepsilon$$

describes the additional expectation gained by mining on a chain away from the C miners.

Now if the miner had directly mined with $C$ chain the expectation, after $k$ blocks would be

$$1 + k\varepsilon.$$

So the thing we have to compare is

$$\frac{2p}{(1-p(1-p))^2} \varepsilon - - \frac{p^2}{1-p+p^2}$$

If

$$\frac{2p}{(1-p(1-p))^2} \varepsilon \geq \frac{p^2}{1-p+p^2}$$

it follows that the miner will be best off mining the NC chain (assuming the other miners do as well - we haven't analyzed this yet.)

In other words.

$$\varepsilon \geq \frac{p\,(1 - p(1-p))}{2}$$

This certainly fails as $\varepsilon \to 0$.

(I'm not 100% - I did this computation quickly, but the results smell right.)

Other considerations that make this a tough problem in general: To model the game correctly, you have a finite number of free agent miners, each with hash rate $\varepsilon_i$ such that $\sum \varepsilon_i = 1 - p$. The game then actually isn't a finite state Markov Decision Process, it's infinite state, because each round will give a new miner a total number of blocks they may have mined on one or either of the given chain. So while the difference in the chain length is the most important state variable to keep track of, the amount of investment in each of the chains by each of the miners will determine whether they choose the C or NC chain.

It's also fun to think about the initial reward that caused the NC block.

Consider the situation where you have 4 miners, each with 10%, 20%, 30% and 40% of the hashrate. Putting aside non-compliance, if someone offers a fee that is 3 times the block reward, what happens?