Analysis on Weak Subjectivity in Ethereum 2.0

Daejun Park[1] and Aditya Asgaonkar[2]

[1] Runtime Verification, Inc.
daejun.park@runtimeverification.com
[2] Ethereum Research
aditya.asgaonkar@ethereum.org

December 7, 2020

1 Introduction

Weak subjectivity [2] is a social-consensus-driven approach for solving the fundamental "nothing-at-stake" problem of proof-of-stake protocols. In particular, it addresses the problem in the presence of long-range forks, while the slashing mechanism handles the case of short-range forks. Specifically, the current weak subjectivity mechanism deals with the following two types of long-range attacks:[3]

- *Exploiting retired validators*: Adversaries can create and reveal a new chain branching from a certain block on the canonical chain, after 2/3 of validators who were active for the block have exited. Note that such validators can still justify and finalize conflicting blocks at earlier slots without being slashed after they have exited.
- *Exploiting diverging validator sets*: Adversaries can build a new chain until the validator set for the new chain is sufficiently different from that of the canonical chain. The larger the difference between the two validator sets, the lower the accountable safety tolerance. For example, if the intersection of the two sets is smaller than 2/3 of each set, then it is possible to have conflicting blocks to be finalized without any validators violating the slashing conditions.

The current weak subjectivity mechanism employs a social consensus layer in parallel to maintain sufficiently many checkpoints (called weak subjectivity checkpoints) so that there exist no conflicting finalized blocks that are descendants of the latest weak subjectivity checkpoint. In other words, the purpose of the latest weak subjectivity checkpoints is to *deterministically* identify the unique canonical chain even in the presence of conflicting finalized blocks caused by the long-range attacks.

[3] It is unknown whether this mechanism can deal with other types of long-range attacks, if any, in general.

2 Weak Subjectivity Period

To minimize the social consensus overhead, we want to identify the minimum cycle to update weak subjectivity checkpoints, called weak subjectivity periods. The weak subjectivity period must be smaller than the minimum time it takes for the condition of the long-range attacks to be established. Specifically, it must be smaller than each of the following:

- E_1: The minimum number of epochs for 2/3 of all active validators to exit.
- E_2: The minimum number of epochs for the original validator set to diverge into two sets so that the (accountable) safety tolerance falls below a certain threshold, $\frac{1}{3} - D$. (We call D safety decay.)

The computation of E_2 is not straightforward, especially when the balance of validators changes over epochs. For simplicity of presentation, we first analyze the weak subjectivity period in a simpler setting where all validators' balance are fixed to the same constant. Later we analyze the effect of balance top-ups separately.

2.1 Weak Subjectivity for Dynamic Validator Set with Unit Balance

In this section, we analyze the weak subjectivity period in a simpler setting where the validator set changes over epochs but their balance is equally fixed to the same amount.

Theorem 1. *Let N be the current total number of active validators. Let δ be the validator activation and exit limit per epoch.*[4] *Let D be the safety decay. Then, given $0 \leq D \leq \frac{1}{3}$, the current weak subjectivity period must be smaller than $\frac{1}{2}DN/\delta$.*

Proof. Recall that the weak subjectivity period must be smaller than $\min(\mathsf{E}_1, \mathsf{E}_2)$, where we claim that:

$$\mathsf{E}_1 = \left\lceil \frac{2N}{3\delta} \right\rceil \qquad \text{and} \qquad \mathsf{E}_2 = \left\lceil \frac{DN}{2\delta} \right\rceil$$

The conclusion immediately follows the above claim since $0 \leq D \leq \frac{1}{3}$. Now let us prove the above claim.

Let us first prove the claim for E_1. By Lemma 1, the maximum number of (originally existing) validators that can exit over the next n epochs is δn. Thus, the minimum number of epochs for $\frac{2}{3}N$ validators to exit, E_1, is $\left\lceil \frac{2}{3}N/\delta \right\rceil$.

Now let us prove the claim for E_2. Let V_0 be the current set of all validators, that is, $|V_0| = N$. Suppose that V_0 diverges into two sets V_L and V_R over two

[4] In the current configuration, δ is `MIN_PER_EPOCH_CHURN_LIMIT` $= 4$ when $N \leq 2^{18}$, and $\left\lfloor N \cdot \texttt{CHURN\_LIMIT\_QUOTIENT}^{-1} \right\rfloor = \left\lfloor N \cdot 2^{-16} \right\rfloor$ otherwise.

chains at a later epoch. Then, by Lemma 2, the number of slashable validators $SV(V_L, V_R)$ is $\max(0, |V_L \cap V_R| - \frac{1}{3}(|V_L| + |V_R|))$, which can be rewritten as:

$$SV(V_L, V_R) = \frac{1}{3}\max(0, |V_L \cap V_R| - |V_L - V_R| - |V_R - V_L|)$$

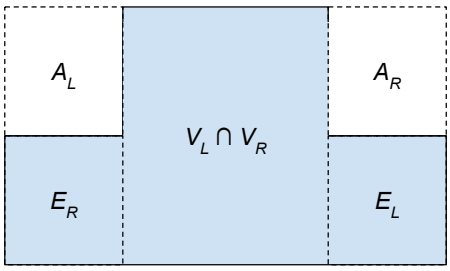
Now we need to calculate the minimum number of epochs required for:

$$\frac{SV(V_L, V_R)}{N} \leq \frac{1}{3} - D \tag{1}$$

This boils down to finding the best strategy to minimize $SV(V_L, V_R)$ as quickly as possible. Note that $SV(V_L, V_R)$ is minimized when $V_L \cap V_R$ is minimized and both $|V_L - V_R|$ and $|V_R - V_L|$ is maximized. That can be achieved by activating different new validators in different chains as much as possible, and removing different originally existing validators in different chains as much as possible. Thus, by Lemma 1, the best strategy to minimize $SV(V_L, V_R)$ as quickly as possible is that:

- Activate new validators at the maximum activation limit per epoch, where the newly activated validators of the two chains have to be disjoint, and
- Remove originally existing validators at the maximum exit limit per epoch, where the exited validators of the two chains have to be disjoint.

Let A_L and A_R be the set of new validators activated in each of the two chains, respectively, during the next n epochs following the above strategy. Similarly, let E_L and E_R be the set of existing validators removed in each of the two chains, respectively. By Lemma 1, we have $A_L = A_R = E_L = E_R = \delta n$. The following diagram illustrates how V_0 diverges to V_L and V_R over the n epochs following the best strategy, where V_0 is denoted by the blue area ($V_0 = (V_L \cap V_R) + E_L + E_R$), V_L is the partial rectangle in the left ($V_L = V + A_L - E_L$), and V_R is the partial rectangle in the right ($V_R = V + A_R - E_R$).



Now we have $SV(V_L, V_R)$ after the next n epochs following the best strategy as follows:

$$\begin{aligned} SV(V_L, V_R) &= \frac{1}{3}\max(0, |V_L \cap V_R| - |V_L - V_R| - |V_R - V_L|) \\ &= \frac{1}{3}\max(0, (|V_0| - E_L - E_R) - (A_L + E_R) - (A_R + E_L)) \\ &= \frac{1}{3}\max(0, |V_0| - 6\delta n) \end{aligned}$$

Thus, given the safety decay $0 \leq D \leq \frac{1}{3}$, the minimum number of epochs for (1), E_2, is $\lceil \frac{1}{2} DN/\delta \rceil$. □

Lemma 1 (Maximum Validator Exits). *Let N be the current total number of active validators. Let $\delta(x)$ be the validator activation and exit limit per epoch for the given set of all active validators whose size is x. Suppose that $\delta(x)$ is proportional to x with a fixed lower bound, that is, that there exist constants δ_0, x_0, and $\rho < 1$, such that $\delta(x) = \max(\delta_0, \lfloor \rho x \rfloor)$ where $\delta_0 = \rho x_0$. Also suppose that ρ is sufficiently small compared to δ_0, that is, $\rho\delta_0 < 1$. Then, the maximum number of validators that can exit over the next n epochs is $n\delta(N)$.*

Proof. First let us show that there exists a strategy that can remove $n\delta(N)$ validators over n epochs. The strategy is simply to remove validators at the maximum exit limit while activating new validators at the maximum activation limit. Since the maximum activation limit is equal to the maximum exit limit, the total number of validators does not change over the n epochs under the strategy, thus $\delta(N)$ validators can exit for every epoch over the n epochs.

Now, let us show that the simple strategy is optimal. Let us first consider the case of $N \geq k_0$, that is, $\delta(N) = \rho N$. Assume that there is a strategy S that can remove more validators than the simple strategy. Then, first, it is clear that the strategy S has to activate new validators at the max per-epoch limit to maximize the size of the validator set, which in turn maximizes the max exit limit. Now, assume that the strategy S removes existing validators at the maximum rate *only* after the first k epochs.[5] Then, after the first k epochs, the total number of active validators becomes $N' = N(1+\rho)^k$, and thus the total number of removed validators is $N'\rho(n-k)$, that is, $N(1+\rho)^k\rho(n-k)$. However, we have:[6]

$$N(1+\rho)^k\rho(n-k) \leq N\rho n \quad \text{for } 0 \leq k \leq n \leq \rho^{-1} \text{ and } 0 \leq \rho < 1 \tag{2}$$

where the equality holds only when $k = 0$. This means that the strategy S is not better than the simple strategy, which is a contradiction, thus we conclude that the simple strategy is optimal. Note that we consider only $n \leq \rho^{-1}$, since all of the originally existing validators are able to completely exit over the ρ^{-1} epochs in the simple strategy.

Similarly in case of $N < k_0$, it is optimal to keep removing validators at the max limit from the first epoch, because the max exit limit is fixed to the constant δ_0 while N is sufficiently small such that $N + \delta_0 \leq k_0$. Even in the case of $N + \delta_0 > k_0$, we have $\delta(N + \delta_0) = \delta_0$ because $\rho\delta_0 < 1$. Note that once the total validator set size becomes greater than k_0, it follows the above reasoning. □

[5] Note that the strategy S is better than any other strategy that removes validators at the maximum rate only for some $(n-k)$ epochs rather than the last $(n-k)$ epochs. That is because the total number of validators at each of those epochs is not greater than that of the last epochs. Similarly, the strategy S is better than any strategy that removes fewer validators than the exit limit but over multiple epochs. For example, removing validators up to only 50% of the exit limit over two epochs is worse than removing validators at the max rate at only the last epoch, because the exit limit for the latter is higher than that of the former.

[6] Graphically, `https://www.desmos.com/calculator/ooq6jzbpzb`

Remark 1. Let us give an approximate (but intuitive) analysis of (2). When $0 \leq \rho \ll 1$, (which is the case in the current configuration for a large N), $(1+\rho)^k$ can be approximated to $(1+k\rho)$. Thus, $N(1+\rho)^k\rho(n-k) - N\rho n$ can be approximated to: $N(1+k\rho)\rho(n-k) - N\rho n$, which can be simplified to: $N\rho k(-\rho k + \rho n - 1)$, whose maximum is 0 at $k = 0$ (for $0 \leq k \leq n \leq \rho^{-1}$ and $0 \leq \rho \ll 1$), since it is a $\cap$-shaped parabola (opens downward) with the roots 0 and $n - \rho^{-1} \leq 0$.

Remark 2. There could exist a better strategy if the activation limit is bigger than the exit limit. For simplicity, suppose that the activation limit is bigger than the exit limit for the first k epochs in the strategy S. Then, the number of removed validators under the strategy S is $N(1+a)^k e(n-k)$, where a is the activation limit, and e is the exit limit, while that of the simple strategy is Nen. Then there exists a non-zero k such that $N(1+a)^k e(n-k) > Nen$.[7] This means that it could be better to add new validators without removing any existing ones for a while to quickly increase the total number of validators, and then remove validators later at a much higher exit rate. However, such a strategy does not outperform the simple strategy when the activation limit is not greater than the exit limit.

Below we analyze the number of slashable validators in case that conflicting blocks are justified (and later finalized) under two different validator sets (on two different chains) at the same epoch.

Lemma 2 (Minimum Slashable Validators). *Let X and Y be the two different validator sets on two different chains at the same epoch. Then, for the given X and Y, the minimum number of validators*[8] *who must have violated the slashing conditions in order for two conflicting blocks to be justified, written as $SV(X,Y)$, is given as follows:*[9]

$$SV(X,Y) = \max(0, |X \cap Y| - \frac{|X| + |Y|}{3})$$

Proof. To justify (and later finalize) conflicting blocks, say A and B, we need votes from 2/3 of both X and Y, say 2/3 of X voting for A, and 2/3 of Y voting for B. Now the question is what is the minimum number of validators (who belong to both X and Y) who must have voted for both A and B. To minimize the number of double-voted validators, all the validators in $X - Y$ and $Y - X$ must vote for A and B, respectively.

Here we have two cases. If either $|X - Y|$ or $|Y - X|$ is greater than or equal to 2/3 of X and Y, respectively, then no one needs to double-vote, that is, the safety tolerance is zero, which agrees on the above formula.

[7] Graphically, `https://www.desmos.com/calculator/1vmjuxciiy`

[8] Here we assume the unit stake balance model for simplicity.

[9] Note that in the ideal case of $X = Y$, we have $SV(X,Y) = \frac{1}{3} \cdot |X|$, which agrees on the ideal safety tolerance.

So, now let us assume that both $|X-Y|$ and $|Y-X|$ is smaller than $2/3$ of X and Y, respectively. Then, there should exist some validators in $X \cap Y$ who voted for A, and there should also exist some (possibly different) validators in $X \cap Y$ who voted for B. Here, the (minimum) size of the former and the latter must be $\frac{2}{3}|X| - |X-Y|$, and $\frac{2}{3}|Y| - |Y-X|$, respectively. Then, by the pigeonhole principle, the following number of validators must have double-voted:

$$\begin{aligned}
&\max(0, (\frac{2}{3}|X| - |X-Y|) + (\frac{2}{3}|Y| - |Y-X|) - |X \cap Y|) \\
=&\max(0, (\frac{2}{3}|X| - (|X| - |X \cap Y|)) + (\frac{2}{3}|Y| - (|Y| - |X \cap Y|)) - |X \cap Y|) \\
=&\max(0, |X \cap Y| - \frac{1}{3}(|X| + |Y|))
\end{aligned}$$

□

2.2 Weak Subjectivity for Static Validator Set with Balance Top-ups

Validators whose balance falls below a certain threshold (16 ETH in the current configuration) are subject to removal, and they are allowed to top up their balance to avoid that. Since the balance top-up essentially has a similar effect to adding new validators, it affects the weak subjectivity period.

In this section, we analyze the effect of balance top-ups on the weak subjectivity period. To highlight its sole effect, here we assume the static validator set, where no additions or removals of validators are allowed.

Lemma 3 (Potential Minimum Safety Tolerance). *Assume the static validator set. Let N be the total number of validators. Let T be the maximum effective balance per validator. Let t be the average effective balance of all active validators at a certain point, that is, the total effective balance at that point is tN. Let Δ be the number of validators allowed to top-up their balance per epoch.*[10] *Then, there exists an adversarial scenario where after a certain number of epochs, the (accountable) safety tolerance can be reduced to the following:*[11]

$$\frac{2t - T}{4t - T} \quad (\textit{for } \frac{T}{2} \leq t \leq T) \tag{3}$$

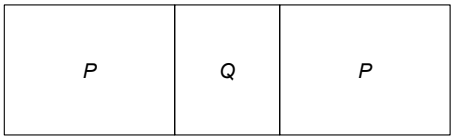
The number of epochs required for (3) *can be as small as:*

$$\frac{tN}{\Delta(4t - T)} \quad (\textit{for } \frac{T}{2} \leq t < T) \tag{4}$$

Proof. Let us present such an adversarial scenario. Suppose that the validator set is split into the following three subsets, where P and Q denote the size of the corresponding subset, that is, $N = P + Q + P$. For simplicity, we assume that the average balance of every subset is t.

[10] In the current configuration, $\Delta = \texttt{MAX\_DEPOSITS} \times \texttt{SLOTS\_PER\_EPOCH} = 512$.

[11] In the ideal case of $t = T$, this agrees on the ideal safety tolerance $1/3$.



Now, suppose that two chains X and Y are created starting from this, where on chain X (and Y), the P validators in the left (and the right, respectively) top-up their balance to the (effective) maximum T. Then suppose that the left P validators in chain X vote for a block A, the right P in chain Y vote for a conflicting block B, and the Q validators in the middle double-vote for both A and B.

Now, let us calculate the minimum Q required for the conflicting blocks A and B to be justified (and later finalized). First, we have that the total amount of stake in chain X or Y is $TP + tQ + tP$, and the weight of votes for block A or B is $TP + tQ$. Then, $TP + tQ$ must be at least $\frac{2}{3}(TP + tQ + tP)$ to justify block A or B. Since $N = 2P + Q$, the minimum Q is $N(2t - T)/(4t - T)$, and the maximum P is $tN/(4t - T)$. Thus we can conclude since the safety tolerance is the minimum Q/N, and the number of epochs required to top-up the balance of P validators is P/Δ. □

Remark 3. The minimum of (3) is 0 at $t = T/2$, and its maximum is $N/3$ at $t = T$. In other words, when the average balance of validators is a half of T (which is 16 ETH in the current configuration), the safety tolerance can be reduced to 0, meaning that conflicting blocks can be finalized without any validators needing to violate the slashing conditions. In the ideal case that the balance of every validator is T, (3) agrees on the ideal safety tolerance $1/3$. Moreover, notably, when $t = \frac{3}{4}T$ (which is 24 ETH in the current configuration), the safety tolerance can be reduced to $\frac{1}{4}$ (i.e., $\sim$8% safety decay).

2.3 Weak Subjectivity for Dynamic Validator Set with Balance Top-ups

In this section, we put together the results of the previous sections, analyzing the weak subjectivity period for the dynamic validator set with allowing validators to top-up their balance.

Theorem 2 (Weak Subjectivity Period). *Let N be the total number of validators. Let T be the maximum effective balance per validator. Let t be the average effective balance. Let Δ be the per-epoch limit of balance top-ups. Let δ be the per-epoch limit of validator activations and exits. Let D be the safety decay. Assume that $\Delta \gg \delta$. Then, given $0 \leq D \leq \frac{1}{3}$, and $\frac{T}{2} \leq t \leq T$, the current weak subjectivity period must be smaller than:*

- *Case $(\frac{1}{3} - D) < (2t - T)/(4t - T)$:*

$$\max\left(\frac{N}{\delta} \cdot \frac{(\frac{1}{3} + 2D)t - (\frac{1}{3} + \frac{D}{2})T}{2t + T}, \quad \frac{N}{\Delta} \cdot \frac{(\frac{1}{3} - D)t + (\frac{2}{3} + D)T}{2t + T}\right)$$

– *Case* $(\frac{1}{3} - D) \geq (2t - T)/(4t - T)$*:*

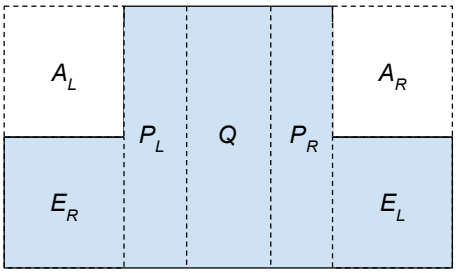
$$\frac{3N}{2\Delta} \cdot \frac{tD}{T - t}$$

Proof. Let us present an adversarial scenario of exploiting diverging validator sets.

Case 1: Let us first consider the case of $\frac{1}{3} - D < (2t - T)/(4t - T)$ for the given D and t. By Lemma 3, the given safety decay cannot be achieved by the sole effect of balance top-ups. Thus both validator activations/exits and balance top-ups are needed to obtain sufficiently diverging validator sets. Suppose that the original validator set evolves over a certain number of epochs as in the following diagram, where A_L and A_R denote the number of newly activated validators on each of two chains respectively, E_L and E_R denote the number of removed validators on each of two chains respectively, and P_L (as well as E_R) and P_R (as well as E_L) denote the number of validators who topped up their balance to T on each of two chains respectively during the period. Also suppose that the two diverging validator sets evolve following the strategies presented in (the proof of) Theorem 1 and Lemma 3. Thus we have:

$$A_L = A_R = E_L = E_R \quad \text{and} \quad P_L = P_R \tag{5}$$

and we omit the subscript when not needed. The original validator set is denoted by the gray area, that is, $N = E_R + P_L + Q + P_R + E_L$. For simplicity, we assume that the original average balance of each subset of validators (i.e., each of E_*, P_*, and Q) is uniformly t. At this point, the total amount of stake for each of the two chains $\mathcal{S} = AT + ET + PT + Qt + Pt$, while the original total amount of stake $\mathcal{S}_0 = Nt = (2E + 2P + Q)t$.



Now, suppose that the validators denoted by A_L, E_R, P_L, and Q vote for a block on one of the two chains, and the validators denoted by A_R, E_L, P_R, and Q vote for another block on the other chain at the same epoch. Note that the Q validators double-vote for both of the conflicting blocks. To achieve the safety decay D, we need to have:

$$Qt = (\frac{1}{3} - D) \cdot \mathcal{S}_0 \tag{6}$$

Also, to justify (and later finalize) each of the conflicting blocks, we need to have:

$$AT + ET + PT + Qt = \frac{2}{3}\mathcal{S} \tag{7}$$

Now, by (5), (6), and (7), we have:

$$E = N\frac{(\frac{1}{3} + 2D)t - (\frac{1}{3} + \frac{D}{2})T}{2t + T}$$
$$P = N\frac{(\frac{1}{3} - D)t + (\frac{2}{3} + D)T}{2t + T}$$

and the number of epochs required for that is $\max(E/\delta, P/\Delta)$.

Case 2: Now let us consider the remaining case of $\frac{1}{3} - D \geq (2t - T)/(4t - T)$. By Lemma 3, the given safety decay can be achieved by the sole effect of balance top-ups. Moreover, $\Delta \gg \delta$, it takes a smaller number of epochs to rely only on the balance top-ups. In this case, we have no validator activations or exits, i.e., $A = E = 0$, and only a subset of P validators top-up their balance, i.e., there exist P_1 and P_2 such that $P = P_1 + P_2$, where only P_1 validators top-up their balance. Then, the total stake of each chain $\mathcal{S}' = P_1T + P_2t + Qt + Pt$, and we need to have $P_1T + P_2t + Qt = \frac{2}{3}\mathcal{S}'$. Since $N = 2P + Q$, $\mathcal{S}_0 = Nt$, and (6), we have:

$$P_1 = \frac{3}{2} \cdot \frac{tDN}{T - t}$$

and the number of epochs required is P_1/Δ. □

3 Discussion

Compared to the limit of the validator activations and exits, the per-epoch limit of balance top-ups is quite high. In the current configuration, while the activation/exit limit is $\max(4, N \cdot 2^{-16})$, the balance top-up limit is 512 regardless of the total number of validators N. The high limit of balance top-ups causes the weak subjectivity period to quickly decrease as the average balance decreases. As shown in Table 1, the weak subjectivity period becomes too small to be practically maintained even for a reasonably large validator set, if the average balance falls below a certain threshold, say 24 ETH. This shows that the current limit of balance top-ups is too high, and it is recommended to significantly decrease the top-up limit. Note that it is *not* recommended to merely increase the validator ejection balance (currently 16 ETH). Although it could help to increase the lower bound of the average balance, it is risky because a higher ejection balance could be abused by adversaries to make it easier to forcibly eject honest validators.

On the other hands, as a workaround for the small weak subjectivity period problem, a relaxed notion of weak subjectivity has been proposed by others. The

[13] The entries of the first low for $t = 32$ agree with the table presented in [1]. The off-by-one difference is due to the use of the floor function instead of ceiling.

Table 1. Weak subjectivity period (in number of epochs) for dynamic validator set with balance top-ups,[13] where the safety decay $D = 10\%$, the maximum effective balance $T = 32$ ETH, the balance top-up limit $\Delta = 512$, and the activation/exit limit $\delta = 4$.

		Validator Set Size (N)						
		262,144	131,072	65,536	32,768	16,384	8,192	4,096
Average Balance (t)	32	3,276	1,638	819	409	204	102	51
	30	2,659	1,329	664	332	166	83	41
	28	1,985	992	496	248	124	62	31
	26	1,248	624	312	156	78	39	19
	24	436	218	109	54	27	13	6
	22	169	84	42	21	11	5	3
	20	128	64	32	16	8	4	2
	18	99	49	25	12	6	3	2
	16	77	38	19	10	5	2	1

relaxed notion does *not* aim to prevent conflicting finalized blocks from being descendants of the latest weak subjectivity checkpoint, thus does *not* guarantee the unique finalized block to be identified by the fork choice rule. Instead, it relies on the withdrawal delay (currently 256 epochs $\approx$ 27 hours), so that arguably the slashable validators can be indeed slashed. However, the implication of the relaxed notion of weak subjectivity has not been thoroughly analyzed, thus it is not recommended to adopt it until more detailed analysis has been made.

Limitation. It is not yet known whether the period given in Theorem 2 is indeed the lower bound (i.e., E_2) or not. This means that the weak subjectivity period in the setting of dynamic validator set with balance top-ups might need to be much smaller. However, when N is not large enough, the given weak subjectivity period is already very small, thus does not affect the practical implication.

On the other hands, if the balance top-up limit Δ is reduced to close to the activation/exit limit δ, then Theorem 2 is no longer directly applicable as it assumes $\Delta \gg \delta$.

References

1. Aditya Asgaonkar: Weak Subjectivity in Eth2.0. `https://notes.ethereum.org/@adiasg/weak-subjectvity-eth2`
2. Vitalik Buterin: Proof of Stake: How I Learned to Love Weak Subjectivity. `https://blog.ethereum.org/2014/11/25/proof-stake-learned-love-weak-subjectivity/`