Prisoner puzzle: (badly worded description of the puzzle follows)

Prisoners have a hat color of either black or white. The prisoners are lined up such that they can only see the hats of prisoners which are in front of them. The prisoners have to answer what their hat color is. The last prisoner has to give the first answer and then the prisoner which is directly in front of him has to answer and so on. If more than one prisoner gives an incorrect answer all prisoners die (losing situation). If they get all hats right except at most one they get free (winning situation). Devise a strategy for the prisoners such that they can get free for every possible hat distribution.

The problem can be abstracted in the following way. The sequence of prisoners with either black or white hat color represents a finite sequence of bits $b_0\ b_1\ \cdots b_n$ with $b_i \in \mathbb{F}_2$ for some $n \geq 0$ in which a white hat is represented by a 0 and a black hat is represented by a 1.

Each prisoner has to give an answer which is either black or white. We denote the answer of the j-th prisoner by $A_j \in \mathbb{F}_2$ for $0 \leq j \leq n$. To give a winning strategy for every possible sequence of bits, we define the k-th parity for $0 \leq k \leq n$ by

$$p_k := \sum_{i=0}^{k} b_i \in \mathbb{F}_2.$$

The strategy is the following: The last prisoner can see every person in front of him and so he can calculate the parity of the corresponding sequence of the prisoners in front of him. In the notation this means the last prisoner gives the answer $A_n := p_n$. Now the second last prisoner can hear the answer of the last prisoner, so he knows A_n and he can see every hat of the $n-2$ prisoners in front of him, which tells him what p_{n-2} is. By definition of the parity it follows that $p_{n-2}+p_{n+1}$ is equal to b_{n-1} in $\mathbb{F}_2$. So define $A_{n-1} := p_{n-2}+p_{n-1} = p_{n-2}+A_n$.

Now every prisoner can hear the answers from the prisoners behind him. In general the k-th prisoner gives the following answer:

$$A_k := p_{k-1} + \sum_{i=k+1}^{n} A_i$$

for $0 \leq k \leq n$, which is also defined for $k = 0$ by setting $p_{-1} := 0$. In this strategy only the answer of the last prisoner A_n is possibly wrong.

achaic