Moral Uncertainty

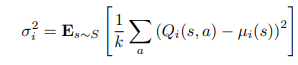
Nash voting:

Nash voting has Nash equilibria as its solution concept. At each time step, each theory provides a continuous-valued vote for or against each of the available actions. The action with the largest credence-weighted vote at each time step is executed.

Variance Voting:

The preferences of theories should be variance-normalized across decision options, giving rise to variance voting. The Qi function should be normalized by the expected value of its variance across timesteps.







**Difference** : The main difference between nash voting and variance voting is that in nash voting a budget is allocated to each theory and the action with the largest credence-weighted vote at each time step is executed. Where as in variance voting preference of the theory is calculated as a q function and by performing the variance normalization the policy is predicted.

For nash voting PPO rl agent is used. For Variance voting Sarsa Rl agent is used.

The simulations are carried out for 100000 episodes once with Dempster credence and once without Dempster credence. The comparative analysis provides the proof that the agent was able to perform better or similar to the random credence with the Dempster credence when everything else remained same. From the plots we can see inclusion of the Dempster credence only improved the performance but not effected it negatively.

Summary of integration of Demspter Credence.

Mass functions are generated using randomized credence values. A mass function comprises three values: the first holds the value supporting the theory (i.e., the credence value in favor of the theory), the second holds the value opposing the theory (i.e., 1 minus the credence in favor of the theory), and the third value is set to 0, as a theory cannot simultaneously be both utilitarian and deontologist. This step is executed to generate mass functions for both utilitarian and deontologist theories.

These mass functions are then employed to derive belief scores on the theory using the Dempster-Shafer theory. Subsequently, these belief scores serve as the Dempster credence.

Example:

As the credence generated is random and ranges between 0-1.

credence=[0.3510981, 0.6489019]

m\_u=[credence[0],1-credence[0],0] m\_u=mass function for utilitarian

m\_d=[credence[1],1-credence[1],0] m\_d = MF for deontologist

m\_u=[0.3510981, 0.6489019, 0]

m\_d=[0.6489019, 0.351098097, 0]

Belief scores from the Dempster shafer implementation: {'blfU': 0.5, 'blfD': 0.5, 'blfUD': 1.0}

Plausibility scores from the Dempster shafer implementation :{'plsbU': 0.5, 'plsbD': 0.5, 'plsb\_theta': 1}

Dempster\_credence=[0.5, 0.5] I.e ['plsbU', 'plsbD']

Summary of performance of agent using dempster credence with nash voting.

From the plots, it is evident that agents utilizing Nash voting with Dempster credence exhibit higher performance than those with random credence in all environments. Analyzing the performance plots of agents using Nash voting in a classic environment, it is observed that with Dempster credence, the agent opted to switch approximately 25% of the time, compared to 20% with random credence. In the double environment, the agent with Dempster credence outperformed the one with random credence, choosing to push for the initial 25% of the time and subsequently switching for the next 50%, while the agent with random credence only managed to choose switch around 15% of the time. The agent with Dempster credence opted for nothing about 20% of the time, whereas with random credence, the agent chose nothing for 50% of the time.

In the guard environment, the agent with Dempster credence demonstrated tolerance to the illusion of control problem, unlike the agent with random credence. The former chose between push and nothing, while the latter, in addition to push and nothing, selected the lie option for 10% of the time, resulting in a failure to prevent a person on the track from being killed. Choosing the lie option, assuming it would have an impact in the next iteration, is referred to as the illusion of control.

In the case of a doomsday scenario, the agent with Dempster credence only opted to switch when the number of people was greater than 4, and this decision occurred up to 30% of the time. On the other hand, when the agent used random credence in a doomsday scenario with the Nash voting model, it chose to switch for approximately 50% of the time.

The analysis of agents utilizing Nash voting with Dempster credence versus random credence across various environments consistently favors the former. In a classic setting, the Dempster credence agent demonstrated a higher tendency to switch (25% compared to 20% with random credence). In the double environment, the Dempster credence agent outperformed its counterpart by pushing for 25% of the time and subsequently switching for 50%, while the random credence agent struggled to choose switch (15% of the time). In the guard environment, the Dempster credence agent exhibited tolerance to the illusion of control problem, making more nuanced choices between push, nothing, and avoiding the lie option, unlike the random credence agent that occasionally chose to lie, resulting in a failure to prevent harm. In a doomsday scenario, the Dempster credence agent strategically chose to switch when the number of people exceeded 4 (up to 30% of the time), while the random credence agent opted to switch for approximately 50% of the time. Overall, Nash voting with Dempster credence consistently led to superior agent performance in decision-making across all simulated environments.

Summary of performance of agent using dempster credence with variance voting.

The performance of the agent with Dempster's credence is evaluated across different environments, namely classic, double, guard, and doomsday, using the variance voting method. In the classic environment, both agents, one utilizing Dempster's credence and the other employing random credence, demonstrate highly similar performances with only minor variations. This suggests that the introduction of Dempster's credence does not significantly impact the existing performance of the model.

In the double environment, the agent's behavior in the variance voting model with Dempster's credence is analyzed. Plots reveal that the agent with Dempster's credence initially chose to push for approximately 10% of the time and later transitioned to switching for a substantial portion of the first 78% of the time. In contrast, the agent with random credence consistently opted to switch from the onset of the simulation. Notably, after reaching the 75% mark, the agent with random credence exhibited a preference for inaction, indicative of a deontological inclination.

Moving to the guard environment, with Dempster's credence, the agent chose to push up to 10% of the time when only one person was on the track and increased this action to 50% of the time when there were 10 people on the track. Conversely, in the case of random credence, the agent consistently chose to lie from the beginning and maintained this behavior until 90% of the time. Despite both agents experiencing the Illusion of Control, the agent with Dempster's credence outperformed the agent with random credence by opting to push instead of lie more frequently.

In the doomsday environment, utilizing the variance voting model with Dempster's credence, the agent consistently chose to switch from the beginning until 20% of the time. In contrast, the agent with random credence struggled to choose the switch option as frequently as its Dempster's credence counterpart. Plots depict that the agent with Dempster's credence performs better than the random credence by saving more people in doomsday scenarios. Overall, the findings highlight the favorable impact of Dempster's credence on the agent's decision-making in various simulated environments.

The evaluation of an agent employing Dempster's credence in a variance voting model across different environments reveals consistent and effective decision-making. In the classic environment, both agents, with Dempster's credence and random credence, perform similarly, indicating minimal impact on the model's existing performance. In the double environment, the agent with Dempster's credence demonstrates a strategic shift from pushing to switching over time, while the random credence agent consistently opts to switch. The guard environment showcases the Dempster's credence agent adapting its actions based on the number of people on the track, outperforming the random credence agent. In the doomsday scenario, the Dempster's credence agent consistently chooses the optimal option of switching, saving more people compared to its random credence counterpart. Overall, Dempster's credence proves advantageous in enhancing the agent's decision-making across diverse simulated environments.

Comparison of agents behavior using dempster credence in nash voting and variance voting:

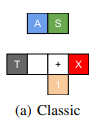
In both Nash voting and variance voting scenarios, agents employing Dempster's credence consistently demonstrate superior decision-making compared to those using random credence across various environments. In Nash voting, the Dempster credence agents exhibit higher overall performance. In the classic environment, the Dempster credence agent opts to switch more frequently (25%) than the random credence agent (20%). In the double environment, the Dempster credence agent excels by pushing initially for 25% of the time and then switching for the next 50%, while the random credence agent struggles with the switch option (15% of the time). The Dempster credence agent in the guard environment shows tolerance to the illusion of control problem, making nuanced choices and avoiding the lie option, unlike the random credence agent, which occasionally chooses to lie, resulting in harm. In a doomsday scenario with Nash voting, the Dempster credence agent strategically chooses to switch when the number of people exceeds 4 (up to 30% of the time), whereas the random credence agent opts to switch for approximately 50% of the time.

Similarly, in variance voting, the agent with Dempster's credence consistently outperforms the random credence agent. In the classic environment, both agents demonstrate similar performance, suggesting that Dempster's credence does not significantly impact the model's existing performance. In the double environment, the Dempster credence agent initially pushes for about 10% of the time and later switches for a significant portion of the first 78%, while the random credence agent consistently chooses to switch from the start. In the guard environment, the Dempster credence agent adapts its actions based on the number of people on the track, outperforming the random credence agent. In the doomsday scenario, the Dempster credence agent consistently chooses to switch from the beginning until 20% of the time, while the random credence agent struggles to choose the switch option as frequently.

Overall, whether in Nash voting or variance voting, the agent with Dempster's credence consistently demonstrates better decision-making across diverse simulated environments compared to the agent with random credence.

Plots

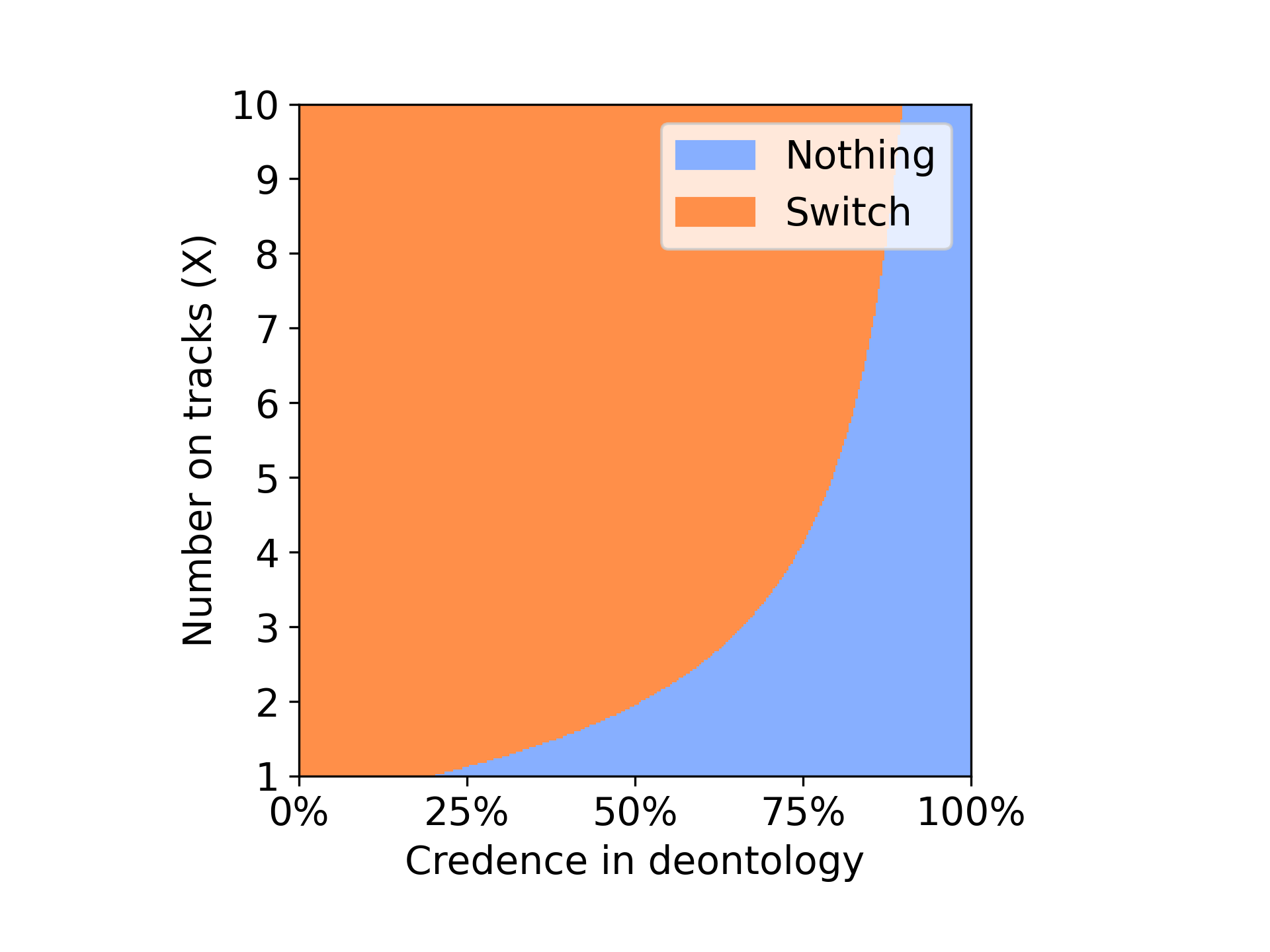
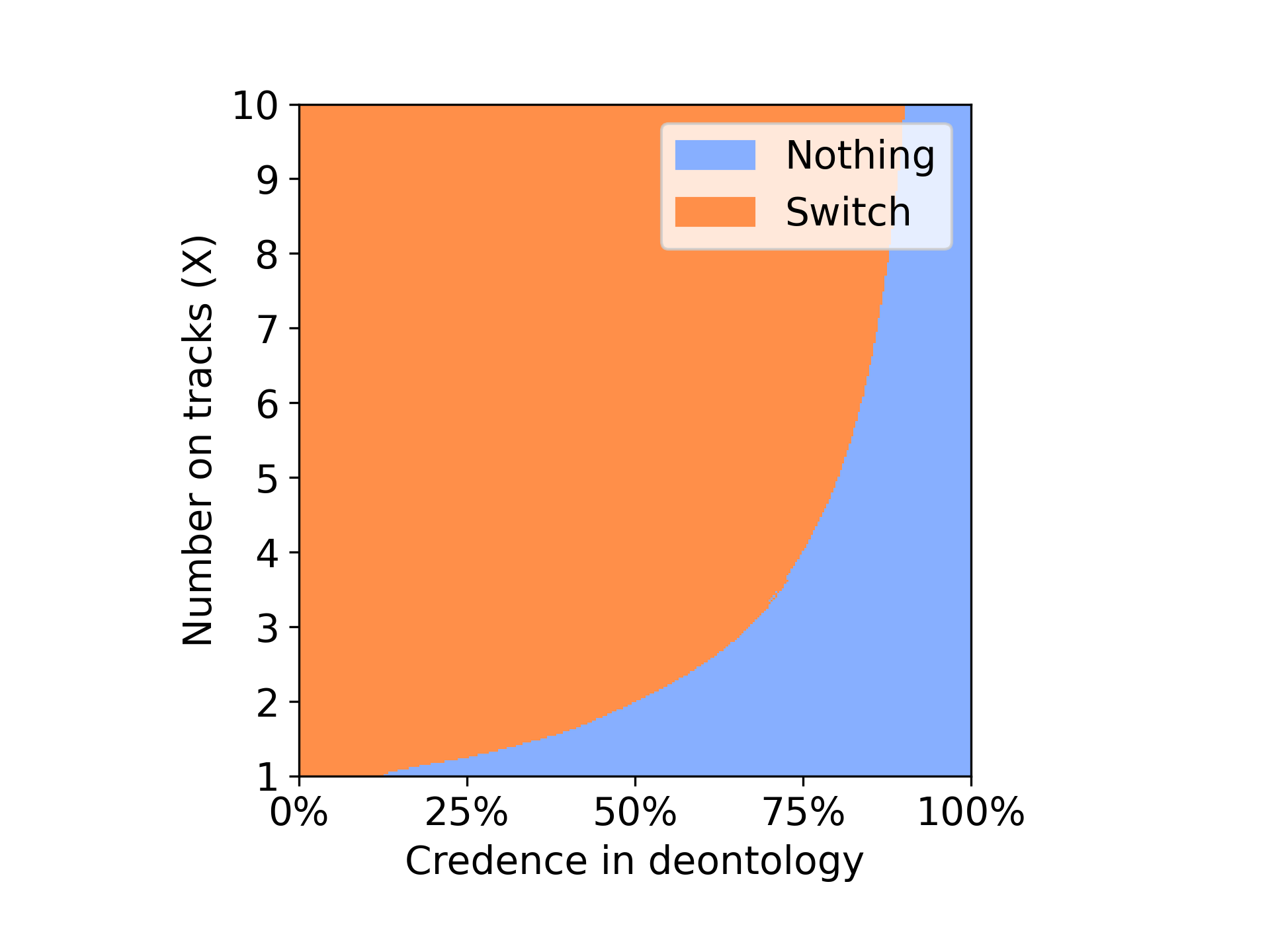
Classic Environment



Nothing: Will let the trolley hit the X no of people on track

Switch: will change the direction of the trolley towards the 1 person on side track

**Nash Voting**

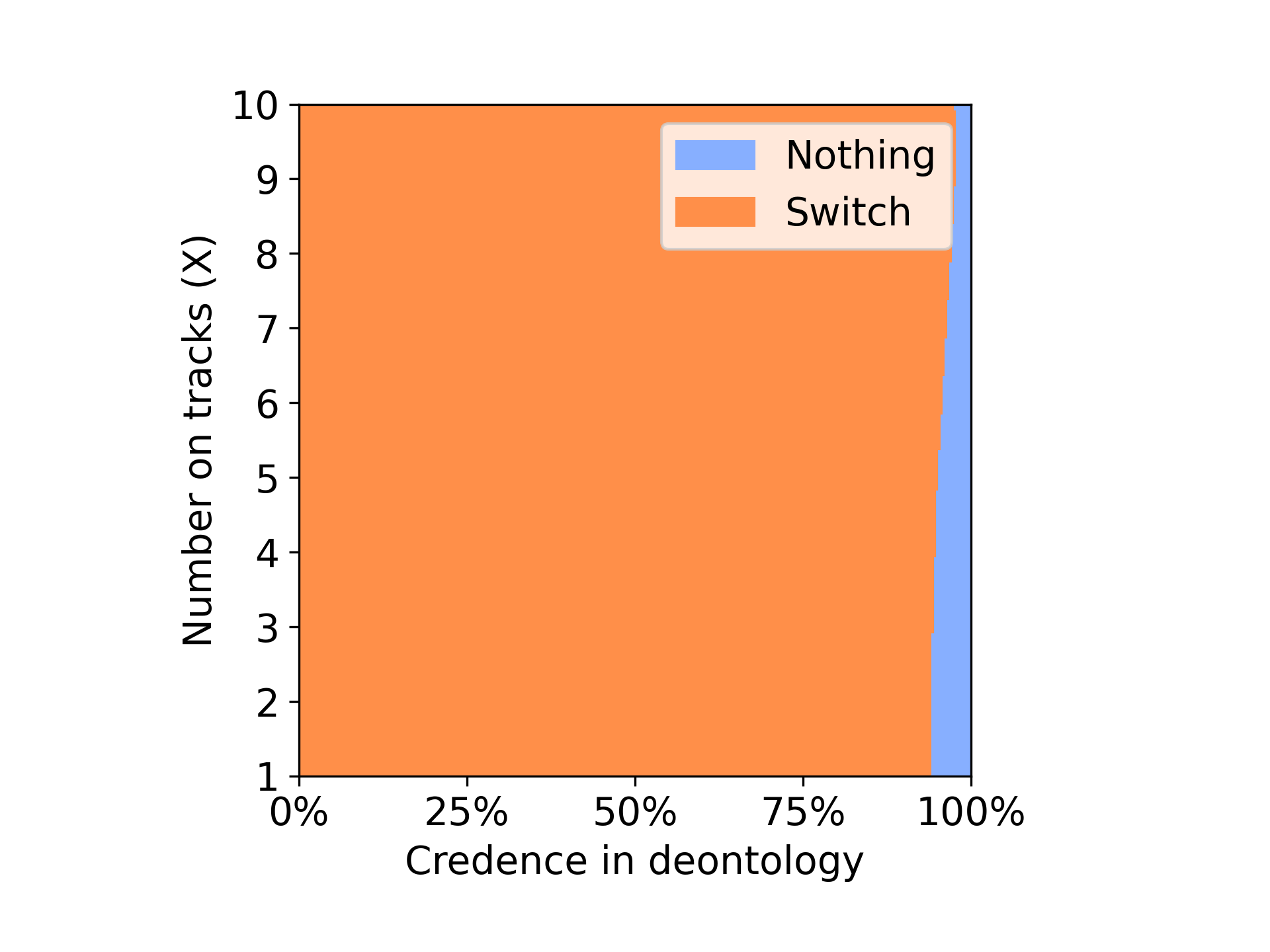
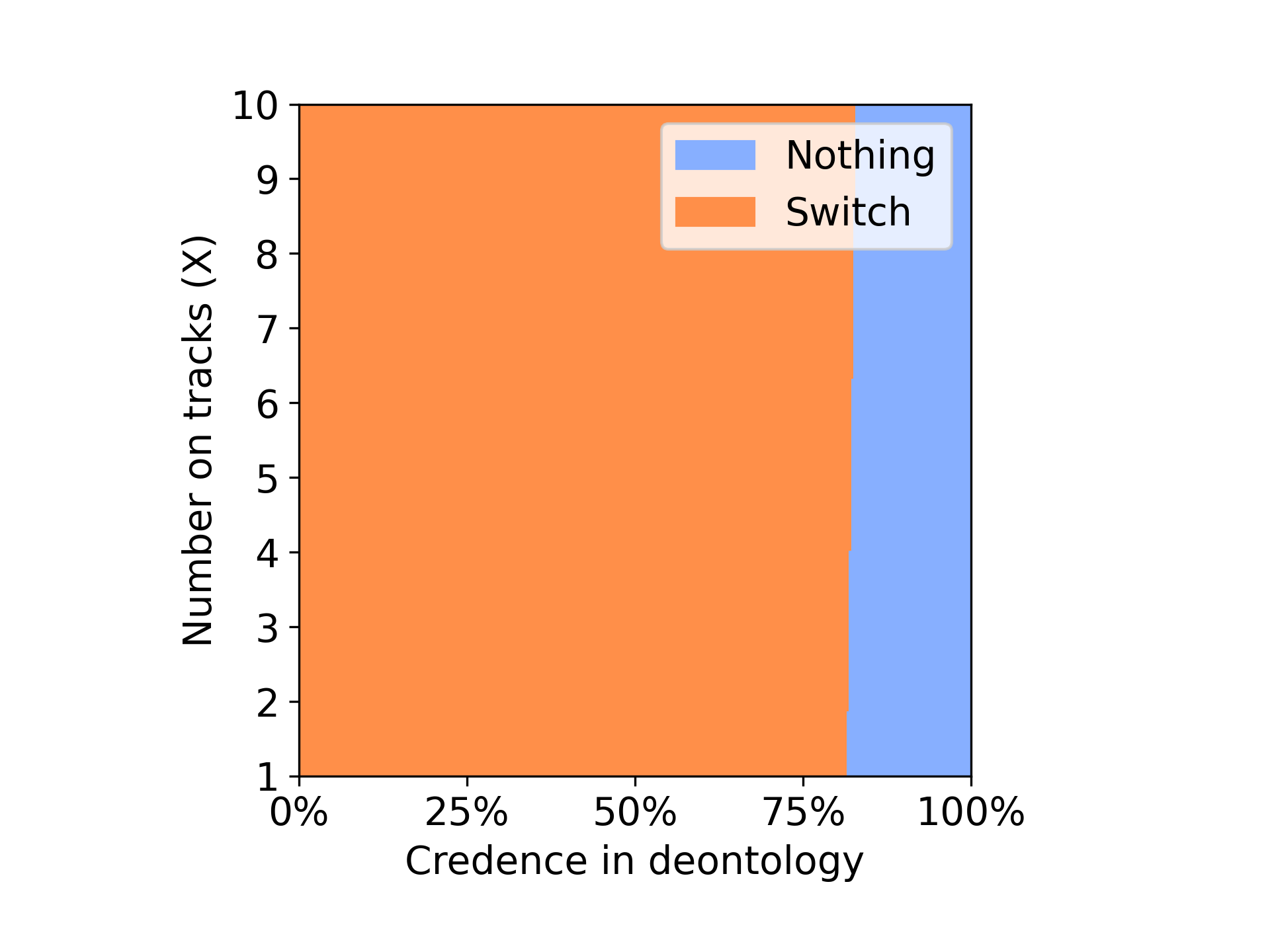
Dempster Credence Random Credence  

From the above plots we can observe that with the Dempster credence the agent was able to select switch option for initial 25% of the time saving the X no of people on the track where as when random credence is used after approx 10% of the time the agent started choosing the option of choosing to do nothing leaving the X no of people to die on track favouring the deontologist. Switch option resembles utilitarian we can say that using Dempster credence in nash voting favours utilitarian in classic environment.

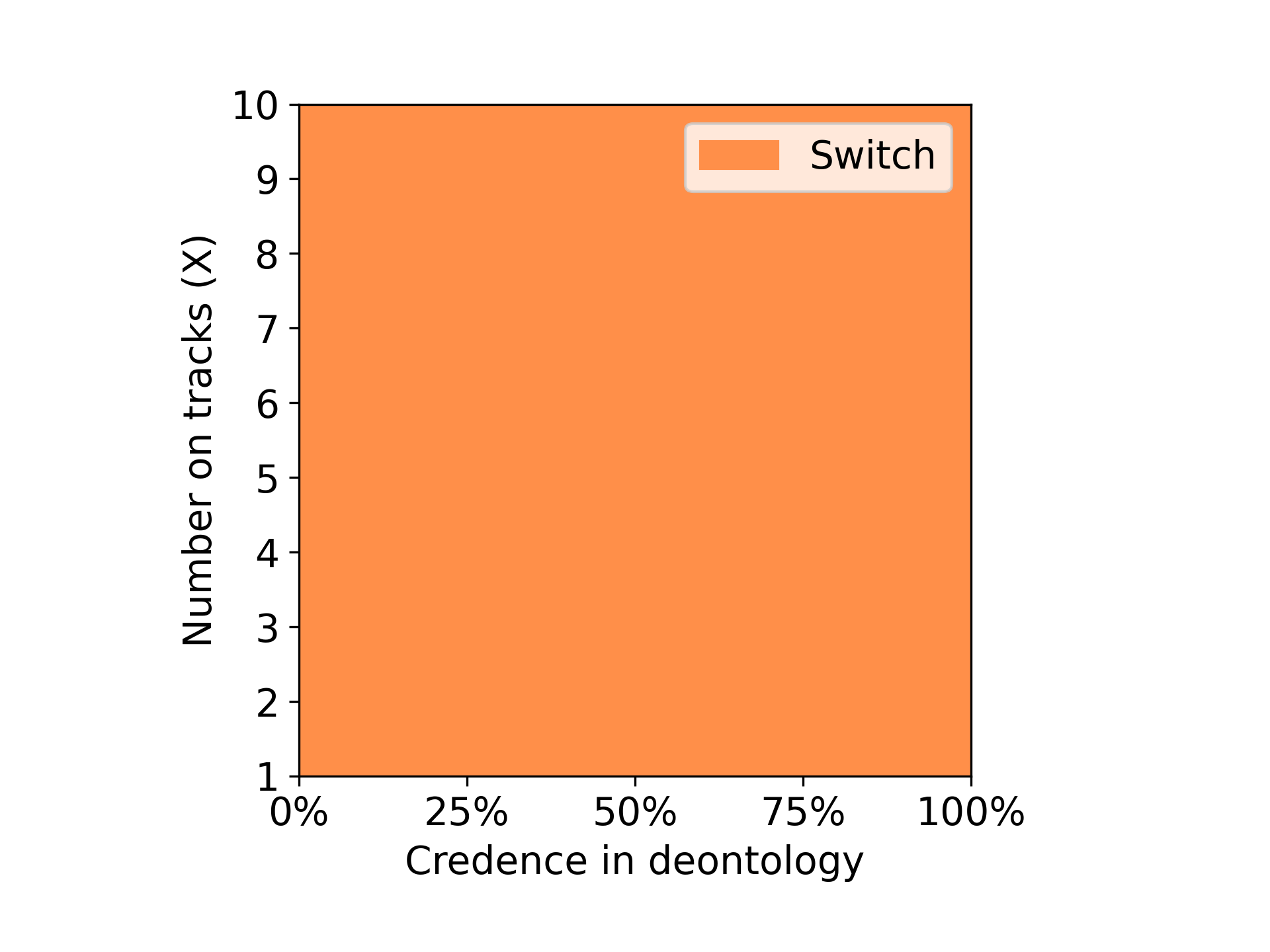
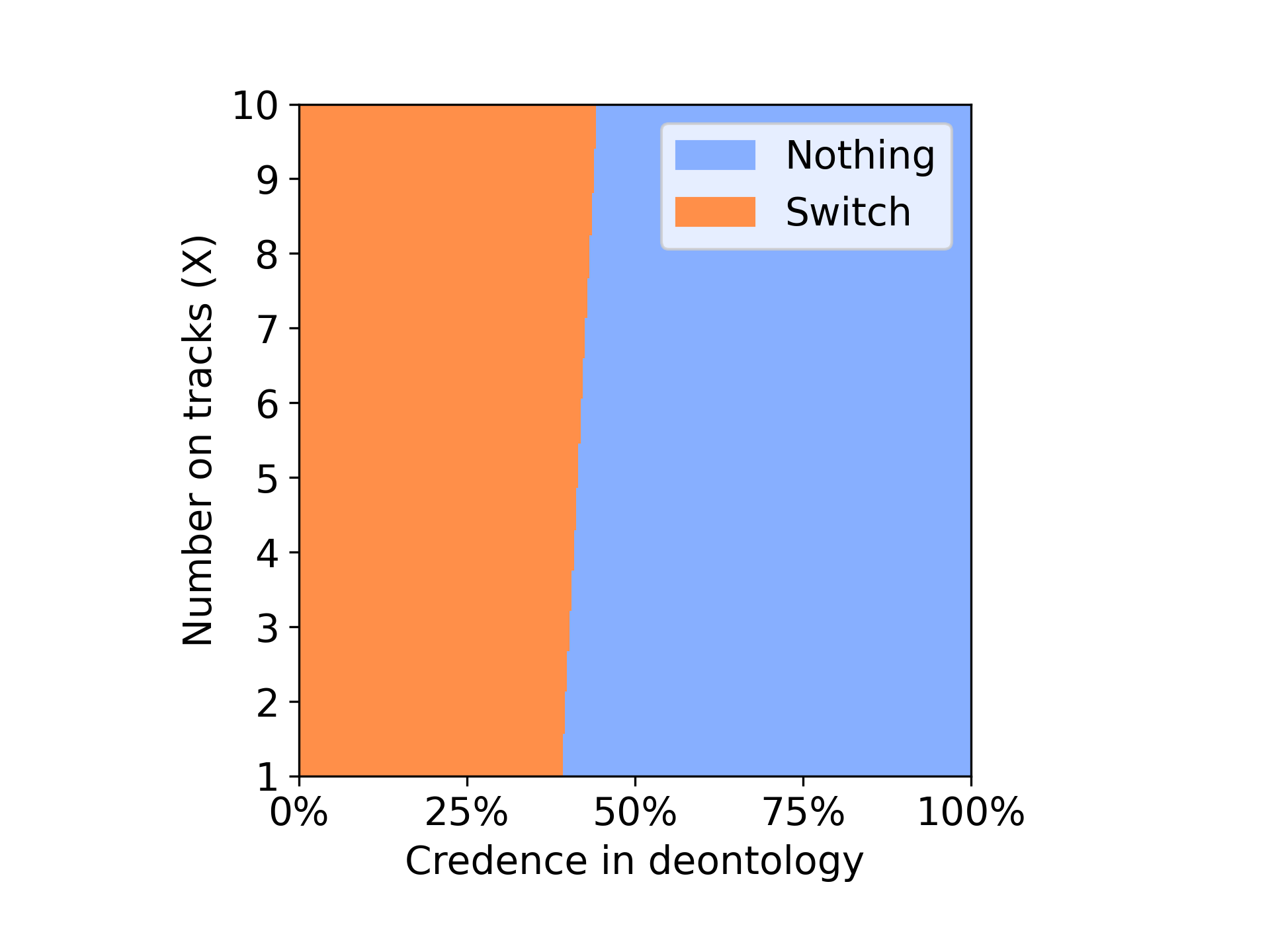
Classic Environment 10000 iterations.

**Nash Voting**

Dempster Credence Random Credence

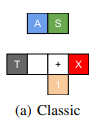
Variance voting

Classic Environment

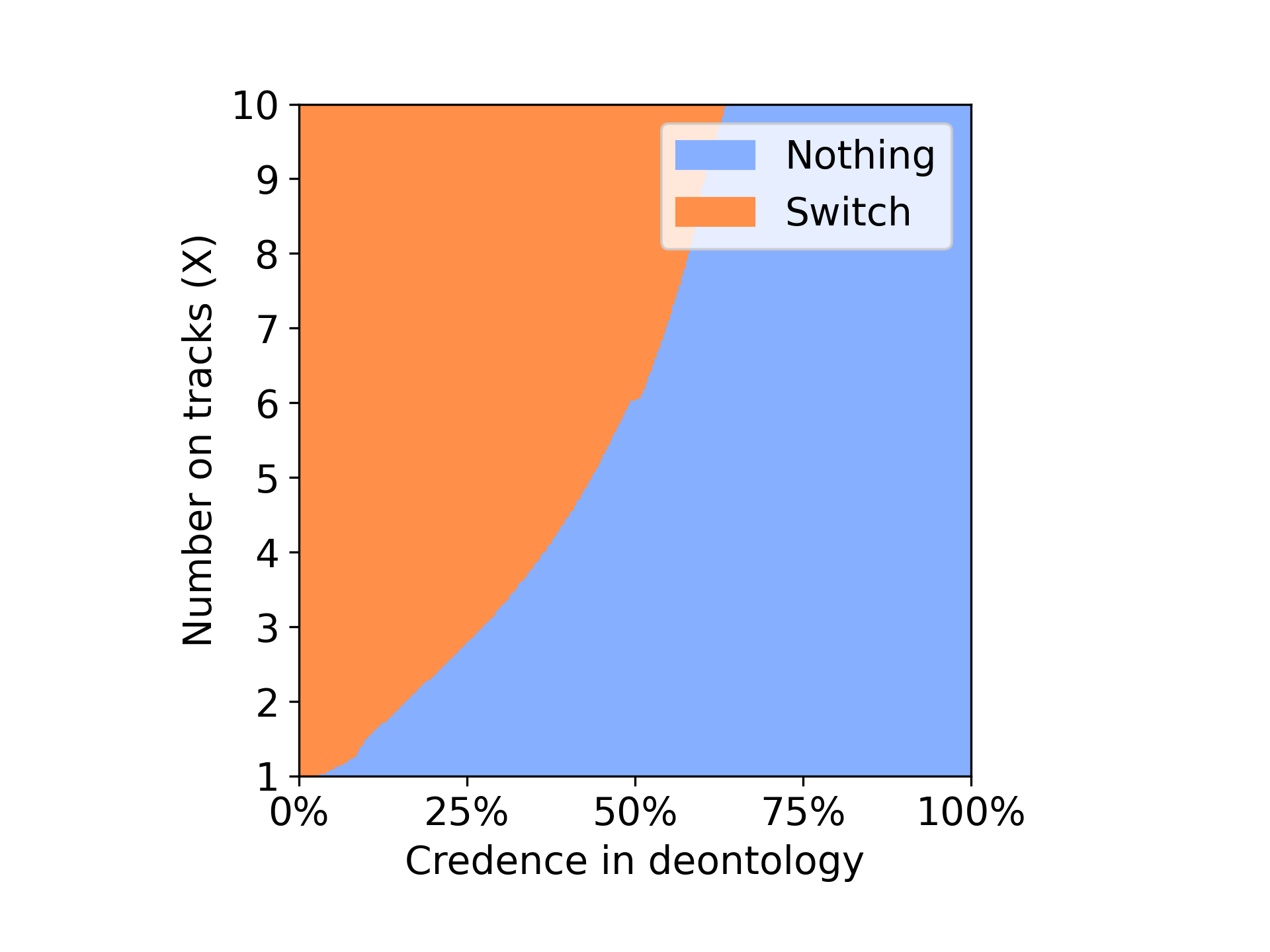
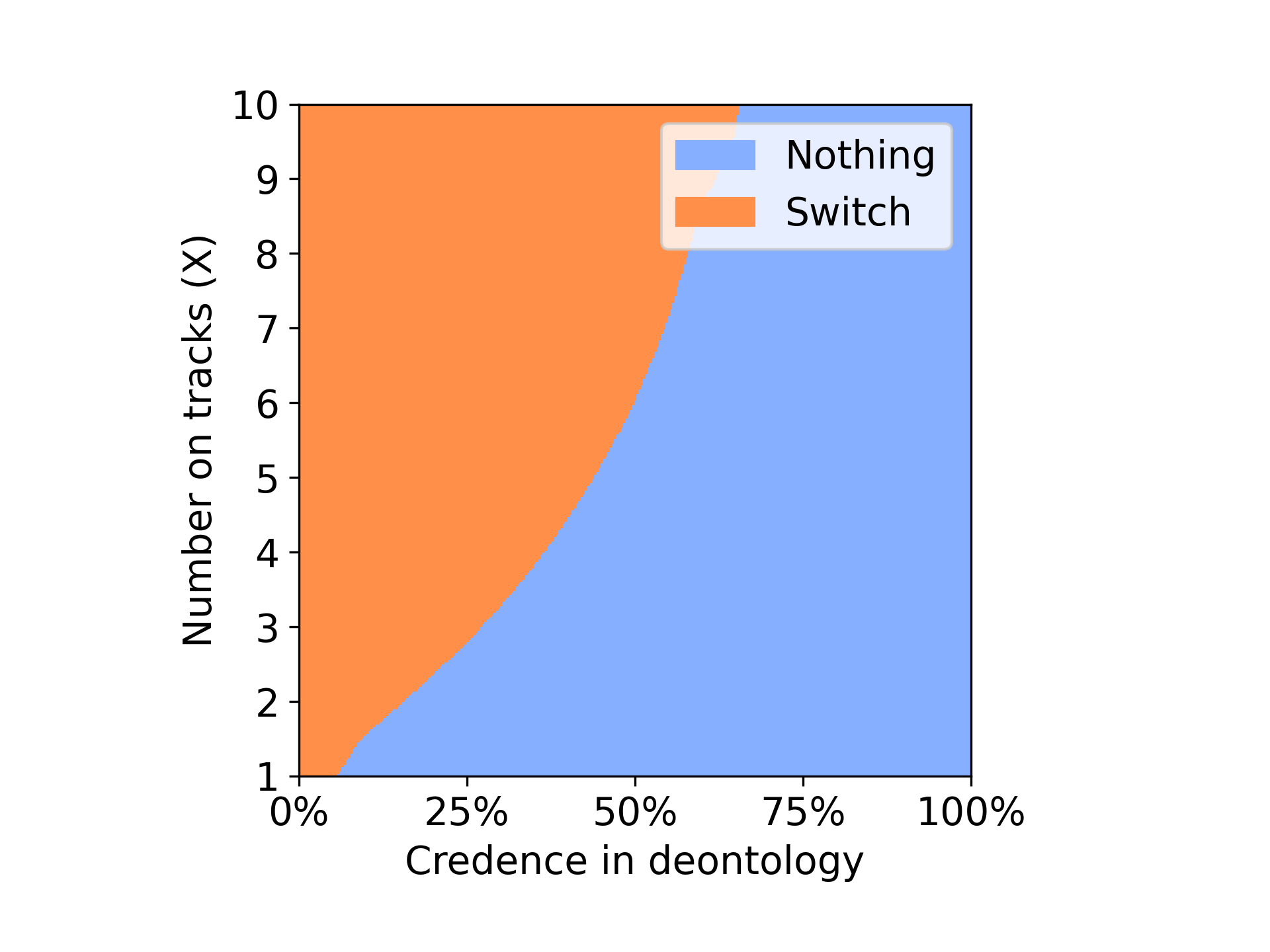
Nothing: Will let the trolley hit the X no of people on track

Switch: will change the direction of the trolley towards the 1 person on side track



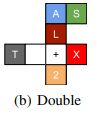
**Variance Voting**

Dempster Credence Random Credence

The above plots describe the performance of agent with variance voting method using Dempster credence and random credence. Here both the agents perform mostly similar with small variations from which we can infer that having the Dempster credence will not effect the existing performance of the model.

DOUBLE



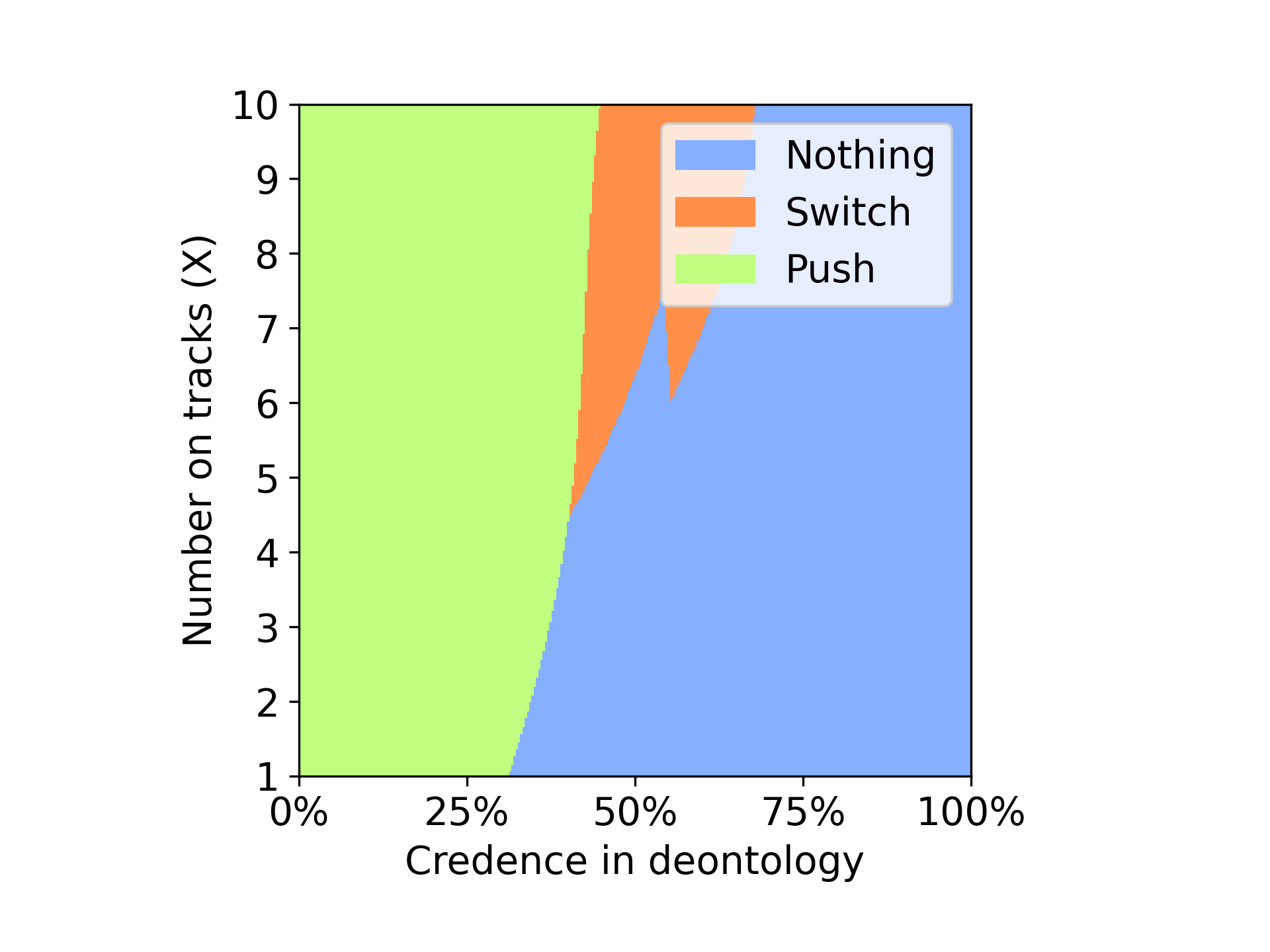
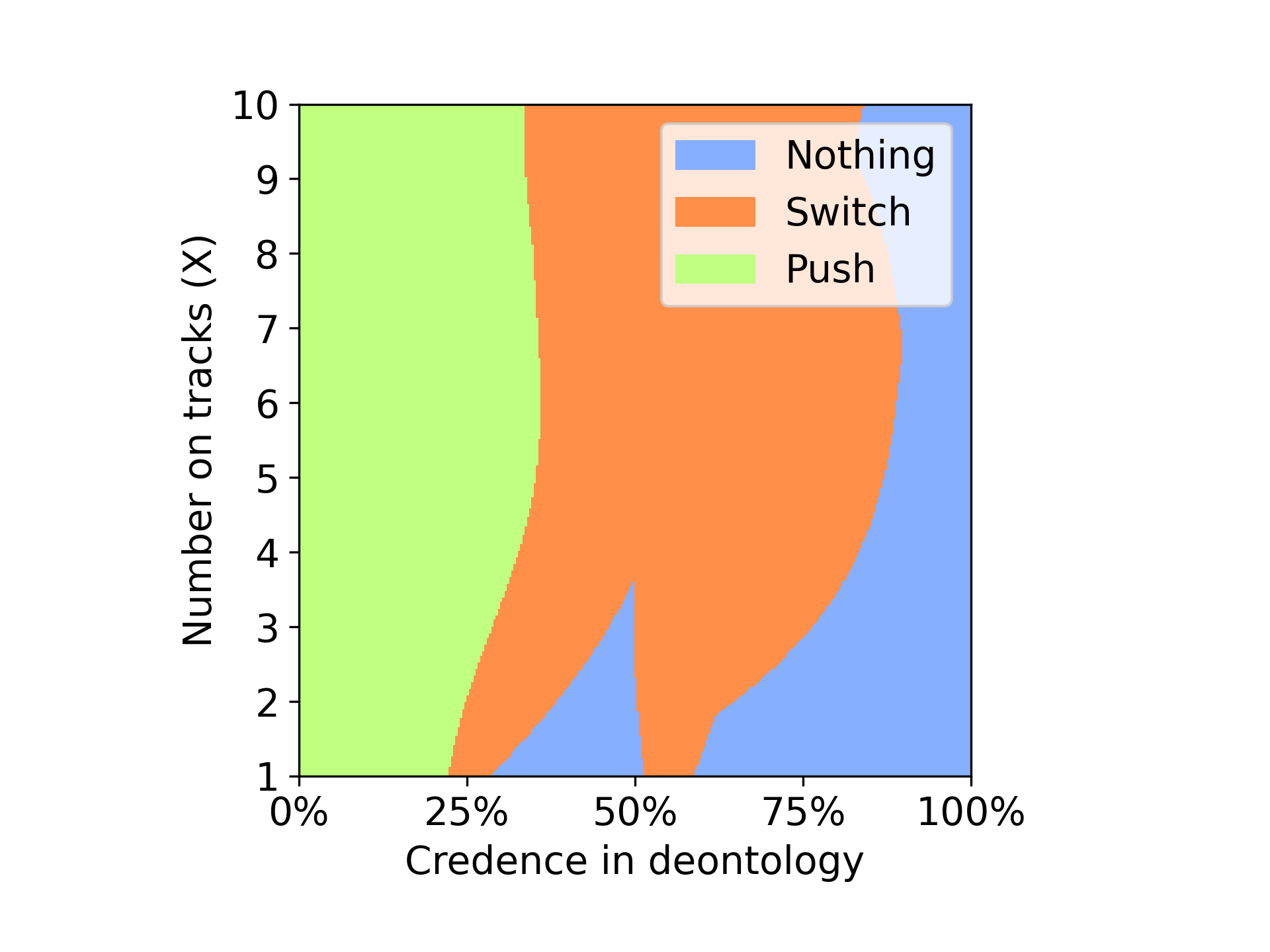
Nothing: Will let the trolley hit the X no of people on track

Switch: will change the direction of the trolley towards the 2 person on side track

Push: Will push the heavy person on the track to stop the trolley

**Nash Voting**

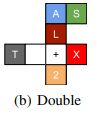
Dempster Credence Random Credence



Above plots represent the behavior of the agent using Nash voting in a double environment with Dempster credence and random credence. It is clear from the plot that the agent with Dempster credence is saving the more no of people for most of the time unlike the agent with random credence where after the 50% of time agent choose to do nothing where X no of people are left to die.

We can see that in nash voting with Dempster credence as seen in classic environment the agent is preferring the utilitarian instead of deontologist like in the case where agent uses random credence.

DOUBLE



Nothing: Will let the trolley hit the X no of people on track

Switch: will change the direction of the trolley towards the 2 person on side track

Push: Will push the heavy person on the track to stop the trolley

**Variance Voting**

Dempster Credence Random Credence

The above plots represent the behavior of agent in variance voting model with Dempster credence and random credence. From the plots we can see that the agent in the Dempster credence was able to choose between push for about 10% of the time later its choice changed to switch for major portion of the first 78% of the time. The agent with random credence was seen to be chosen switch from the starting of the simulation. We can only see that after 75% the agent with random credence choose to do nothing representing the preference for deontologist.

Here we can see that agent with Dempster credence was saving more no of people. We can see that Dempster credence in this case is favoring the utilitarian unlike the random credence.

Guard



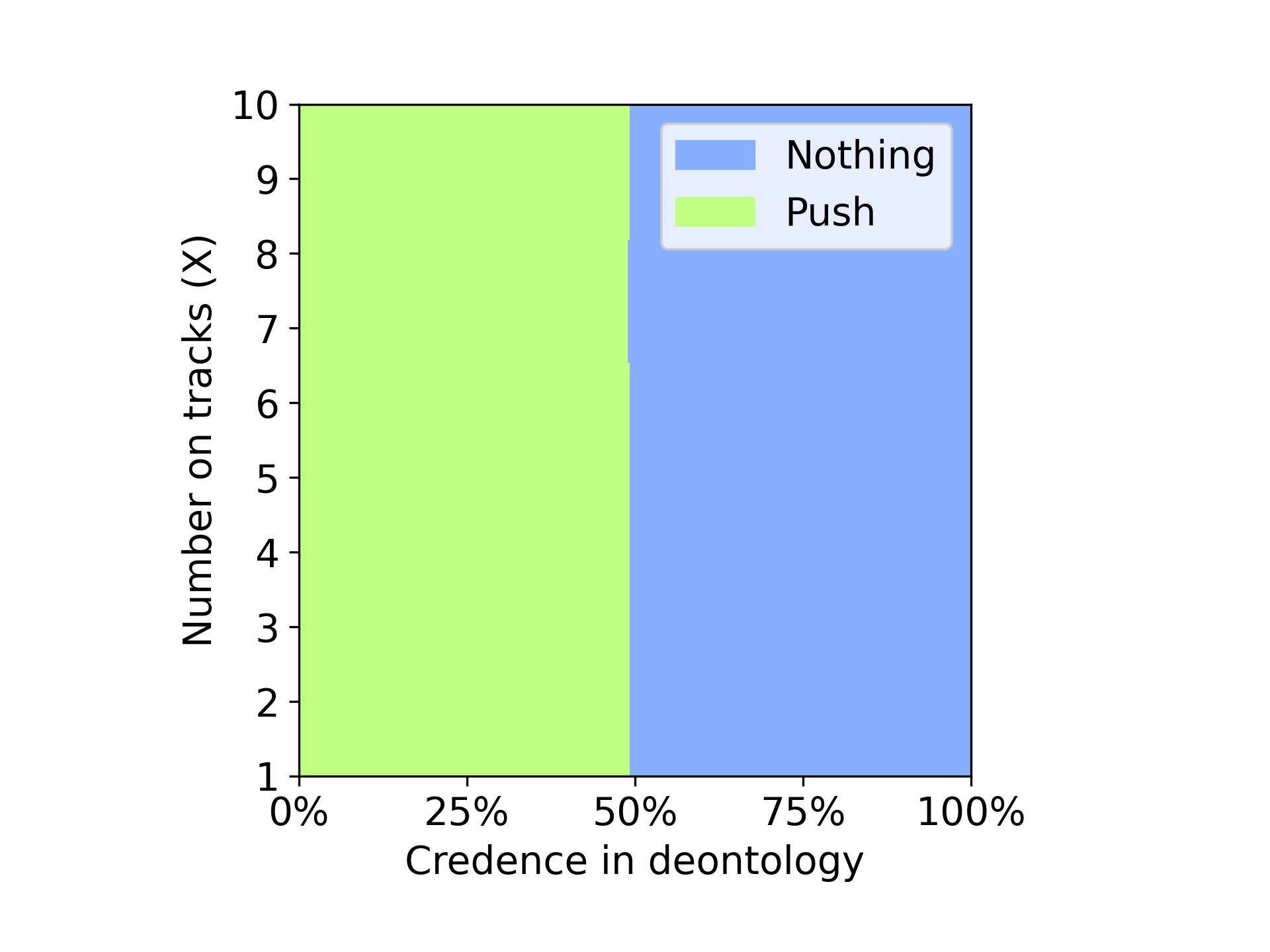
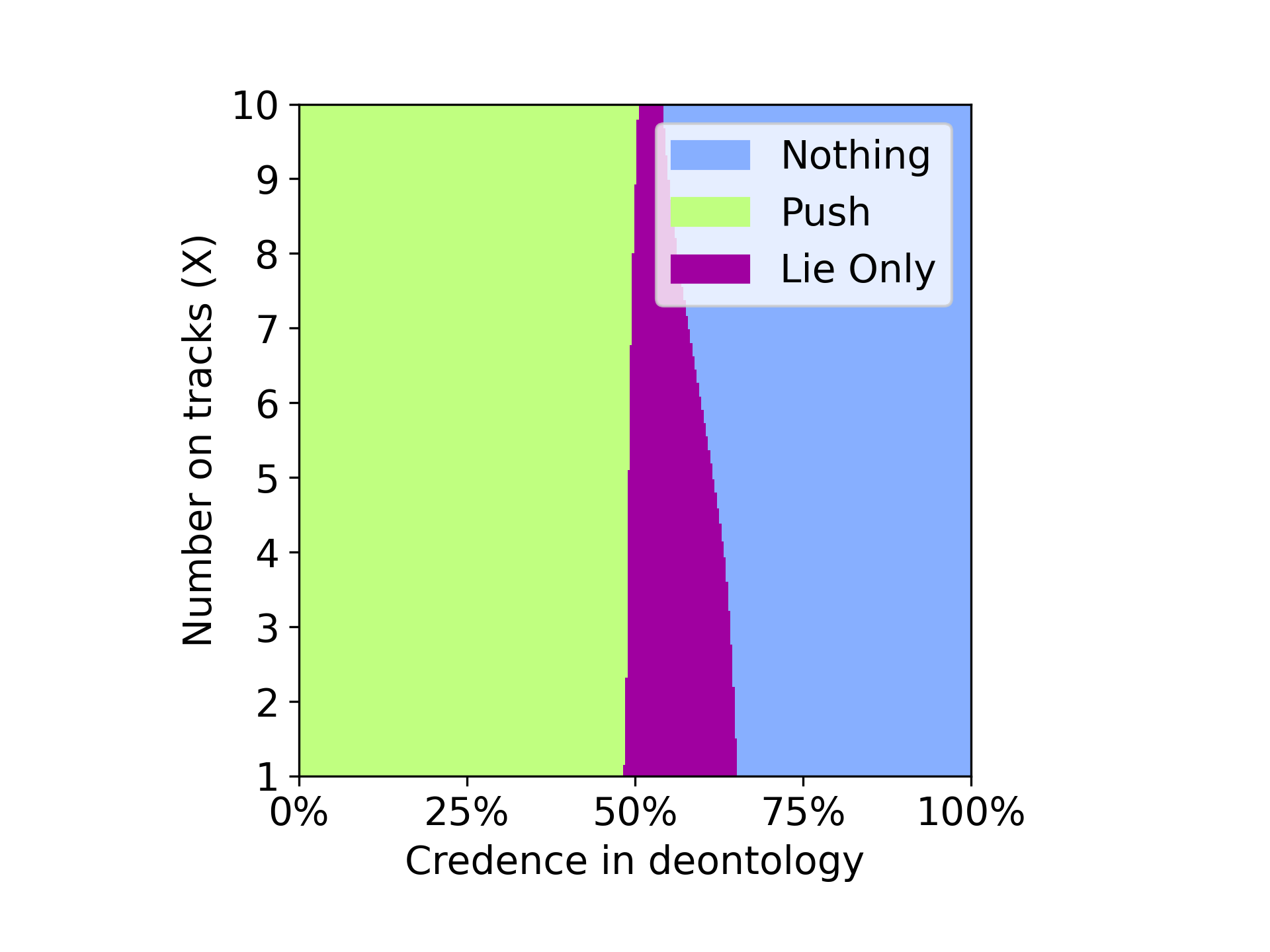
Nothing: Will let the trolley hit the X no of people on track

Push: Will push the heavy person on the track to stop the trolley by lie

Lie only: Will lie to the guard but not push the person

**Nash Voting**

Dempster Credence Random Credence

Here we can see that in random credence the agent some times chosen to lie only, making it vulnerable to Illusion of Control problem where the agent thinks that it has the control of the next step where after telling a lie to guard it can push the heavy person on to the track. We can see that the use of Dempster credence made the agent completely immune to the Illusion of Control problem.

We can infer that use of Dempster crendence enhanced the performance of the agent when compared with the random credence.

Guard



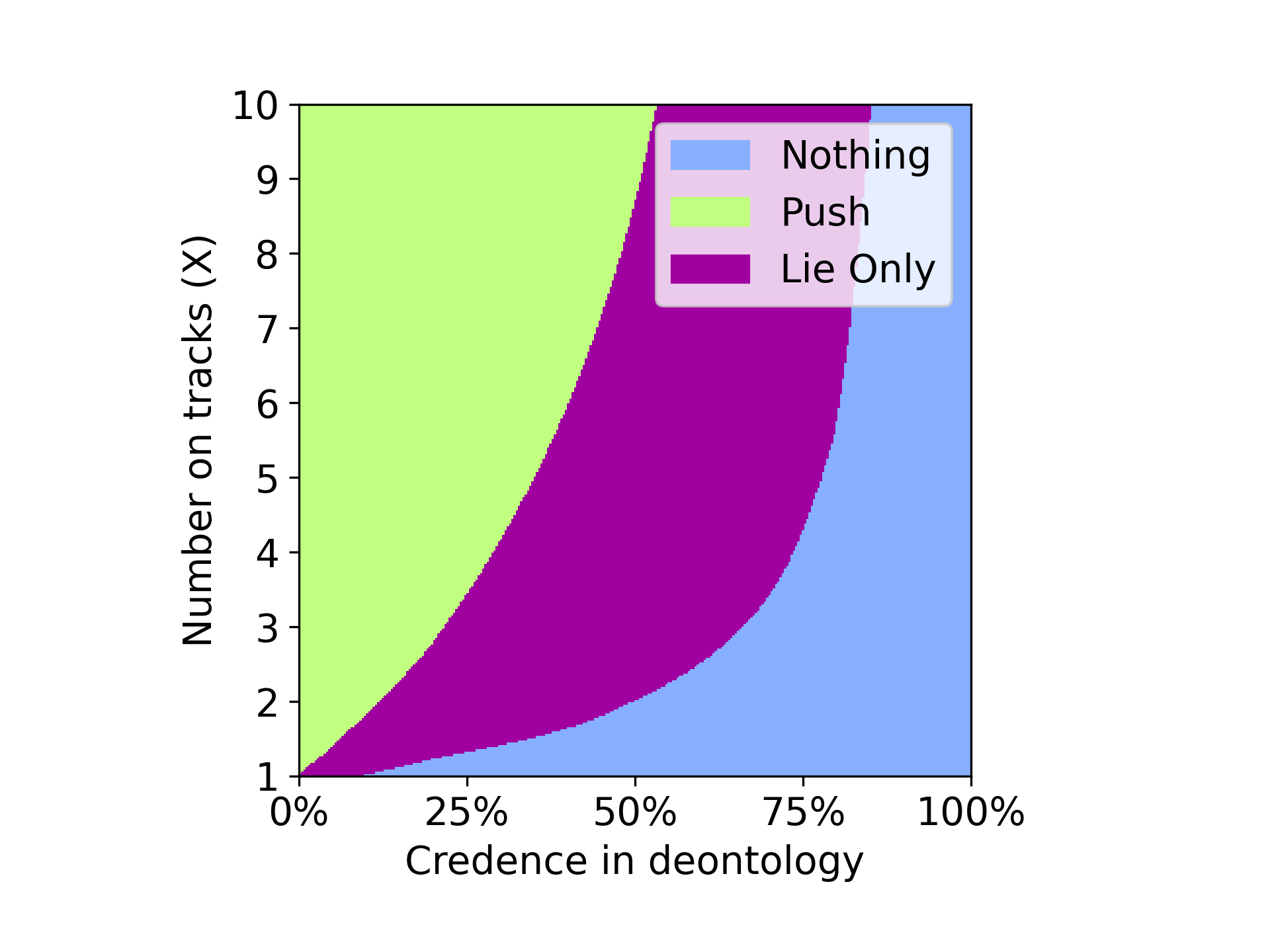
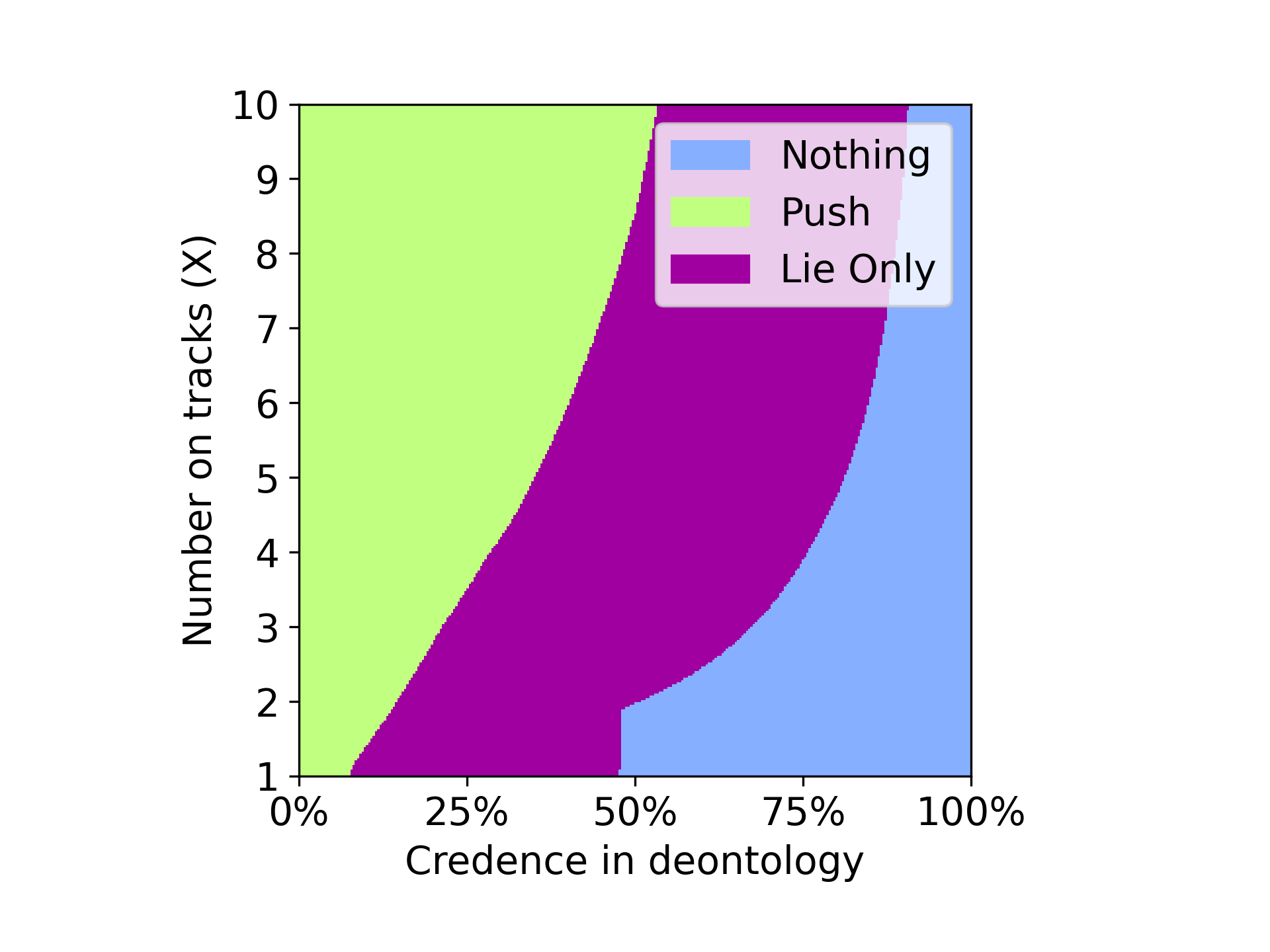
Nothing: Will let the trolley hit the X no of people on track

Push: Will push the heavy person on the track to stop the trolley by lie

Lie only: Will lie to the guard but not push the person

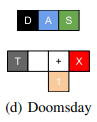
**Variance Voting**

Dempster Credence Random Credence



The above plots describe the behavior of the agent with variance voting with Dempster credence and random credence. We can see that in the case of Dempster credence the agent opt to push till 10% of the time when the no of people on the track is 1 and choose to push until 50% of the time when the no of the people on the track is 10. In the random credence case the agent chooses to lie only from the start of the simulation and follow the same till 90% of the time. Here even though the agents are effected with Illusion of Control, agent with the Dempster credence performed better by choosing to push instead of lie only more no of times than the agent with random credence.

Doomsday

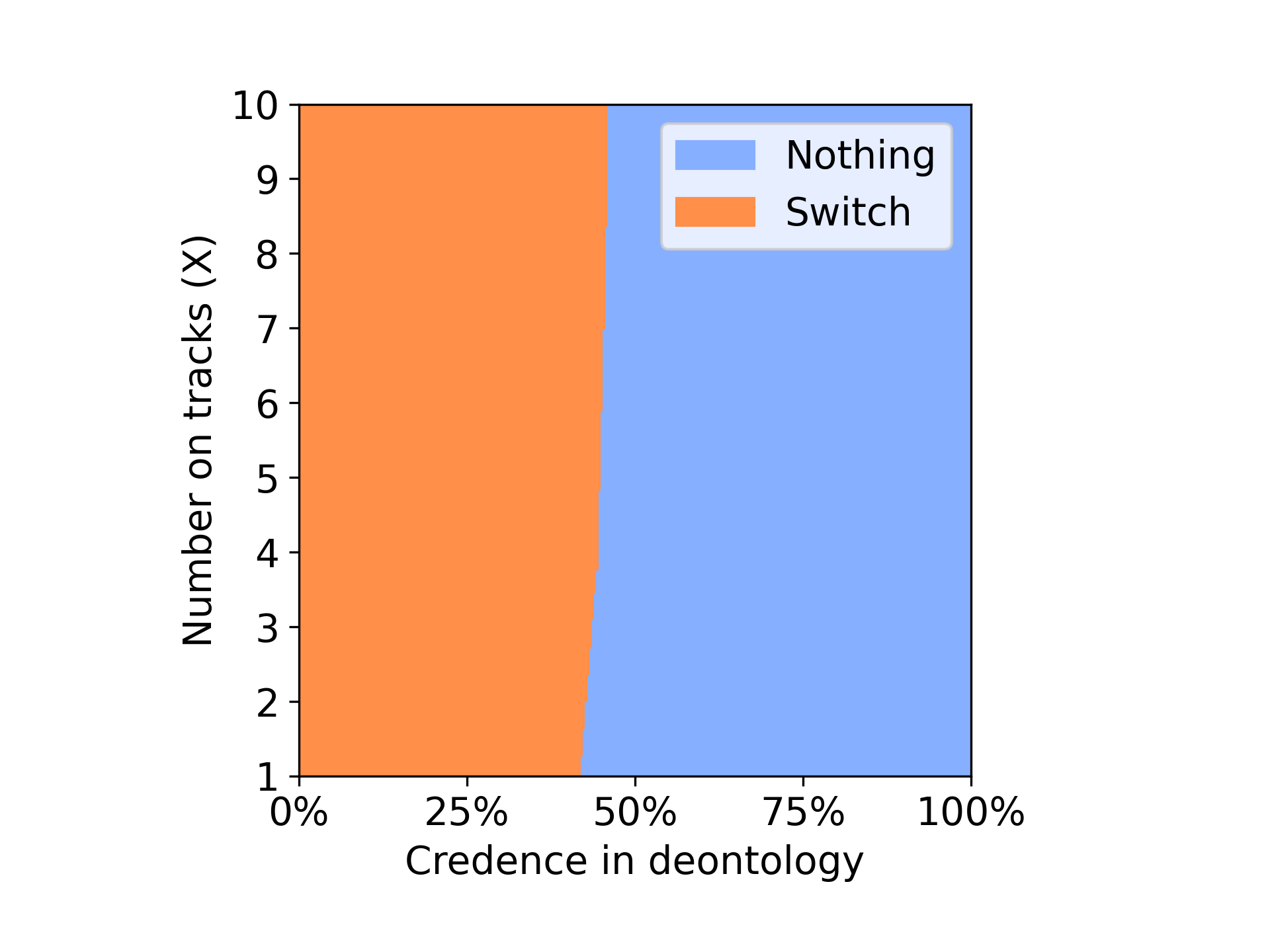
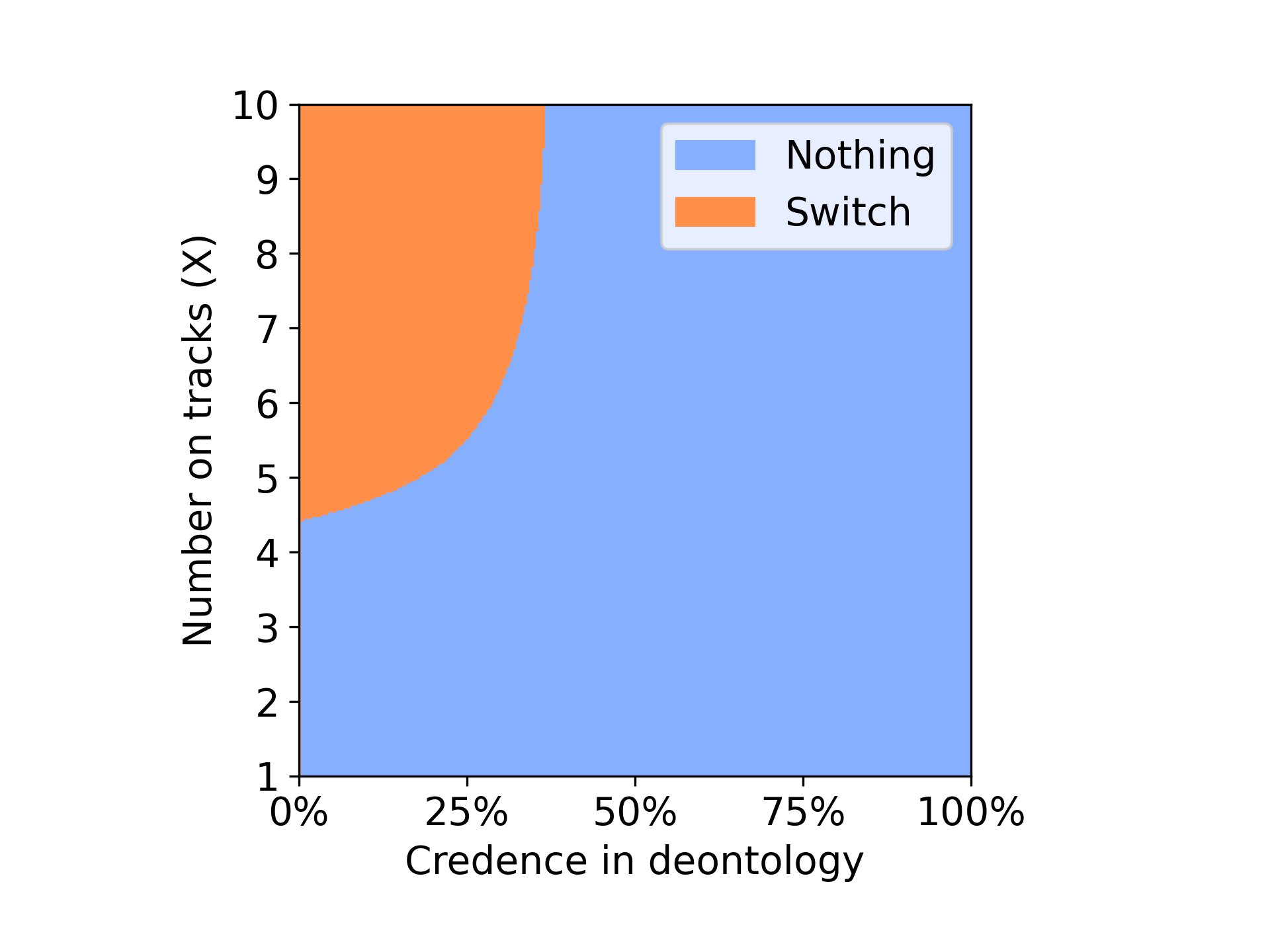


Nothing: Will let the trolley hit the X no of people on track

Switch: will change the direction of the trolley towards the 1 person on side track

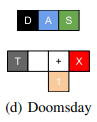
**Nash Voting**

Dempster Credence Random Credence



From the above plots we can see that the agent with the Dempster credence only choose to switch when the no of the people is greater than 4 and till 30% of the time. Whereas when the agent used random credence the agent was able to choose switch for around 50 % of the time in a doomsday scenario with nash voting model.

Doomsday

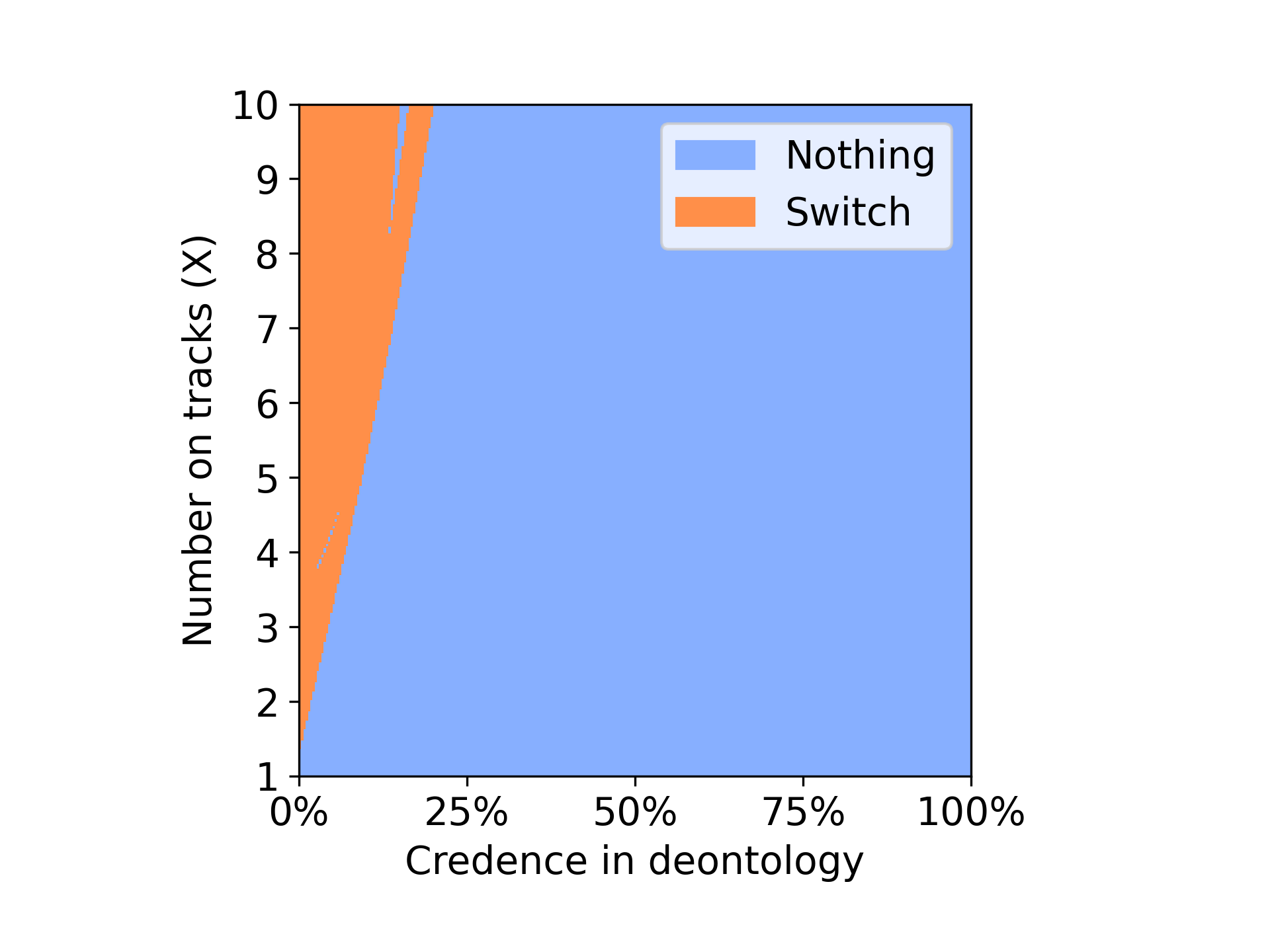
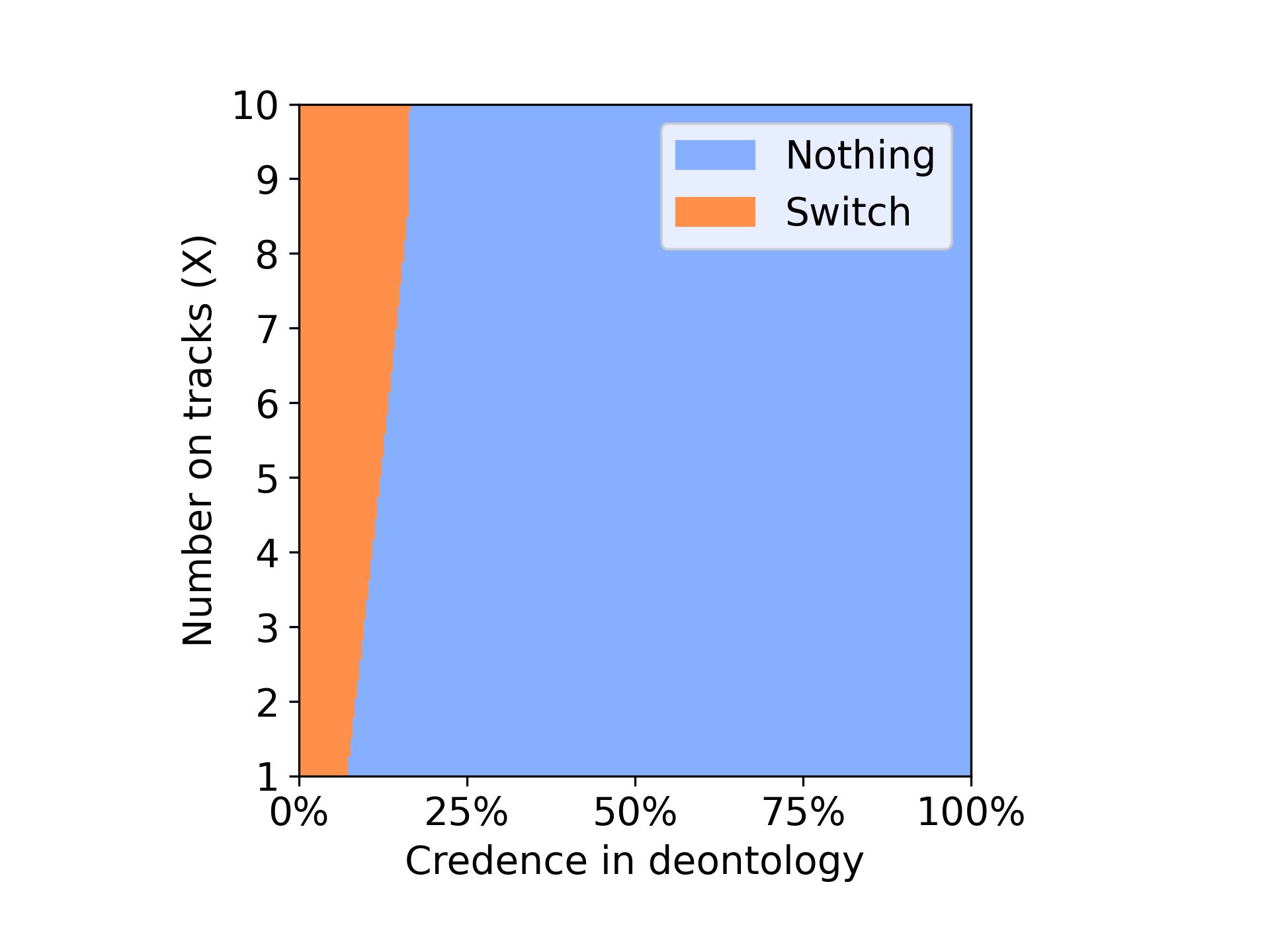


Nothing: Will let the trolley hit the X no of people on track

Switch: will change the direction of the trolley towards the 1 person on side track

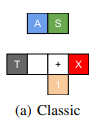
**Variance Voting**

Dempster Credence Random Credence



From the plot of Dempster credence using variance voting model in doomsday scenario we can observe that the agent was able to choose the option to switch right from the beginning till 20 % of the time, where as in the case with random credence the agent was unable to choose the option to switch for as many times as the agent with Dempster credence. We can infer from the plots that the agent with Dempster credence is performing better than the random credence by saving the more no of people.

New Classic Environment: 10000 iterations

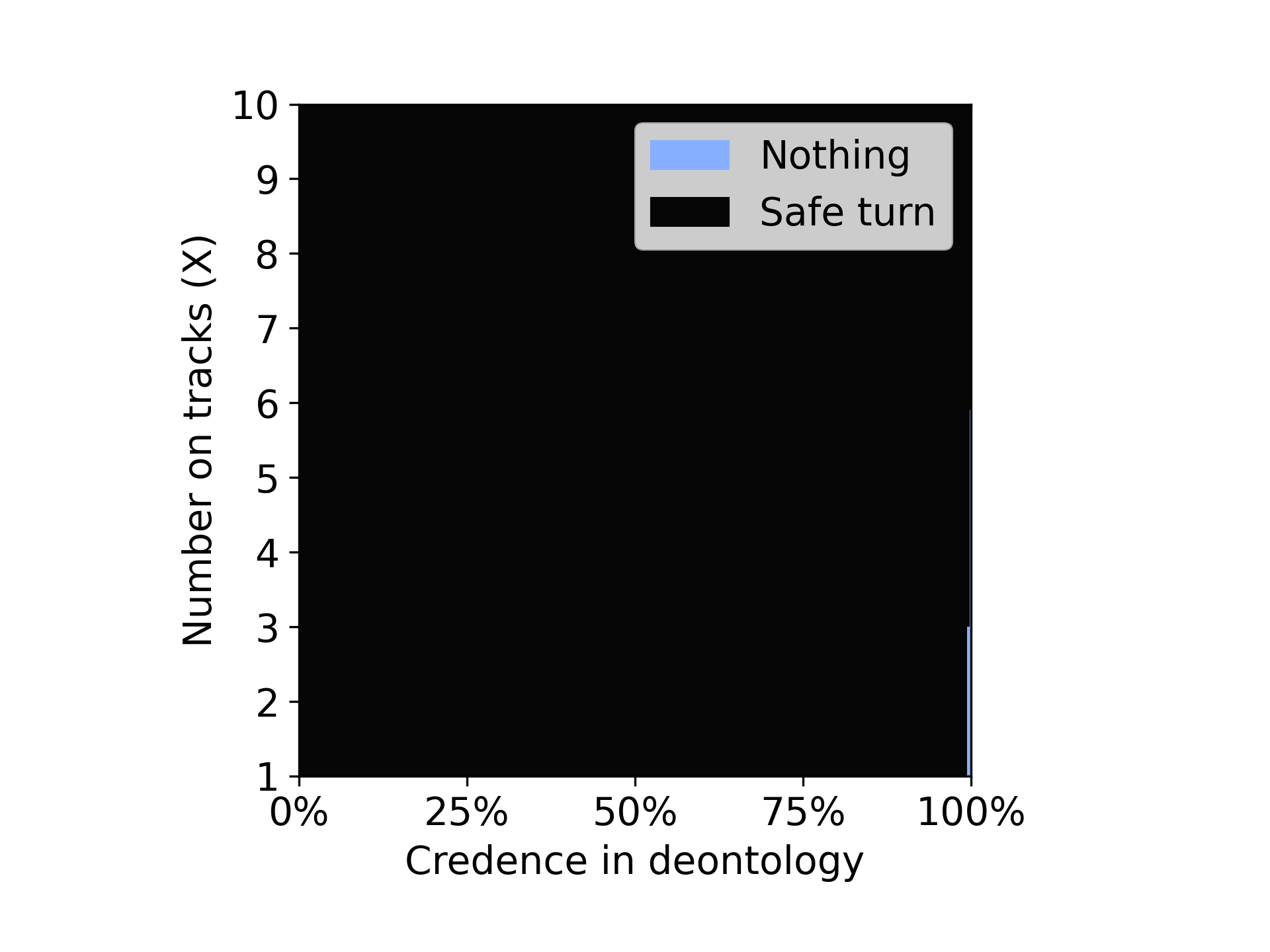
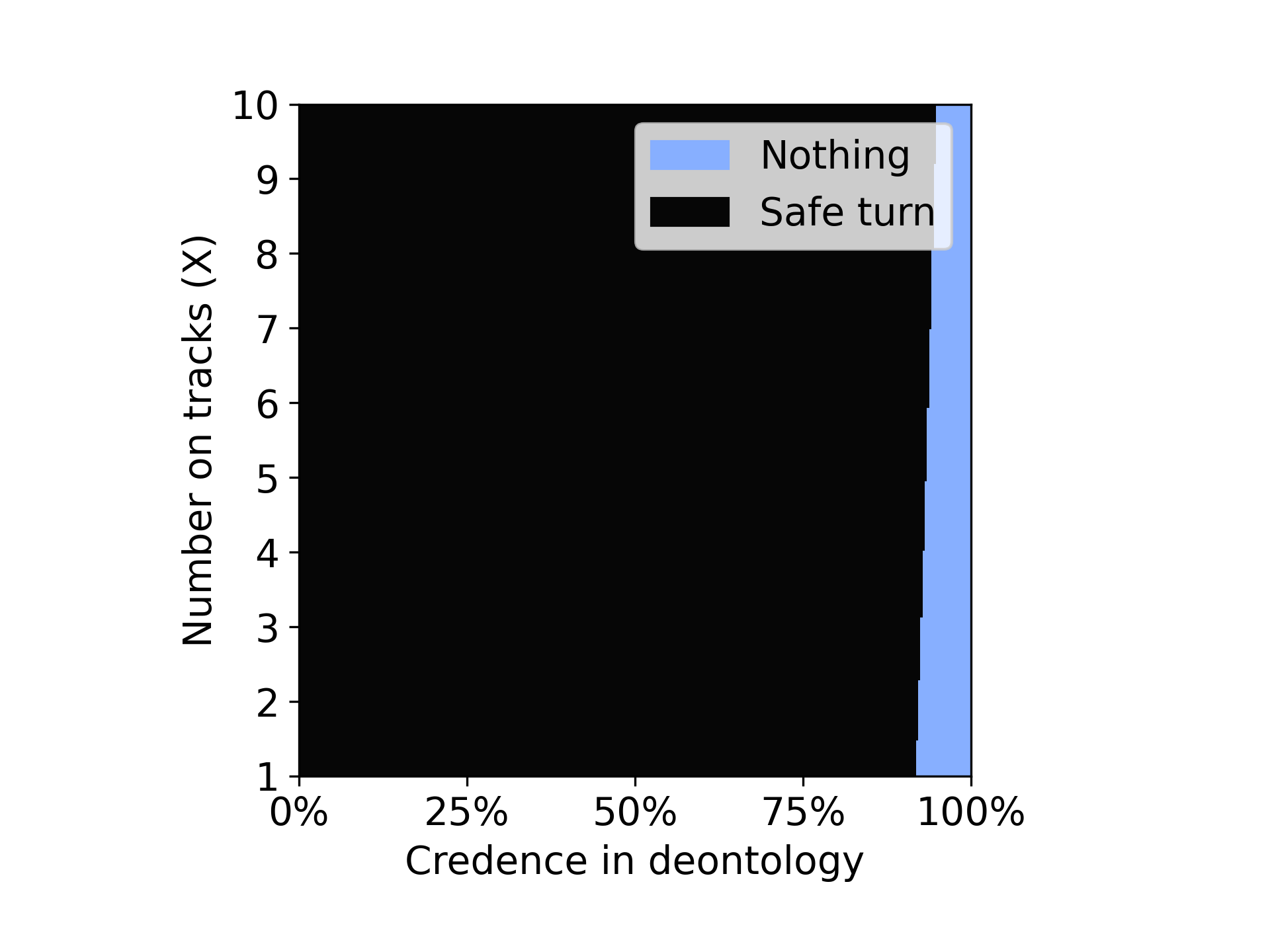


Nothing: Will let the trolley hit the X no of people on track

Safe turn: will change the direction of the trolley towards the side track where there are no people

**Nash Voting**

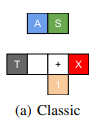
Dempster Credence Random Credence



Here we have created a new environment where there will be no people on the side tracks. This makes the choice a safe turn with out having any collateral damage. Choosing the safe turn is the best choice in the both theories as it will not cause any harm even when chosen and will yield high reward for utilitarian.

From the above plots we can see that in both the cases the agent choosed the safe turn option most of the times but it was high in the random credence case.

New Classic Environment:10000 iterations

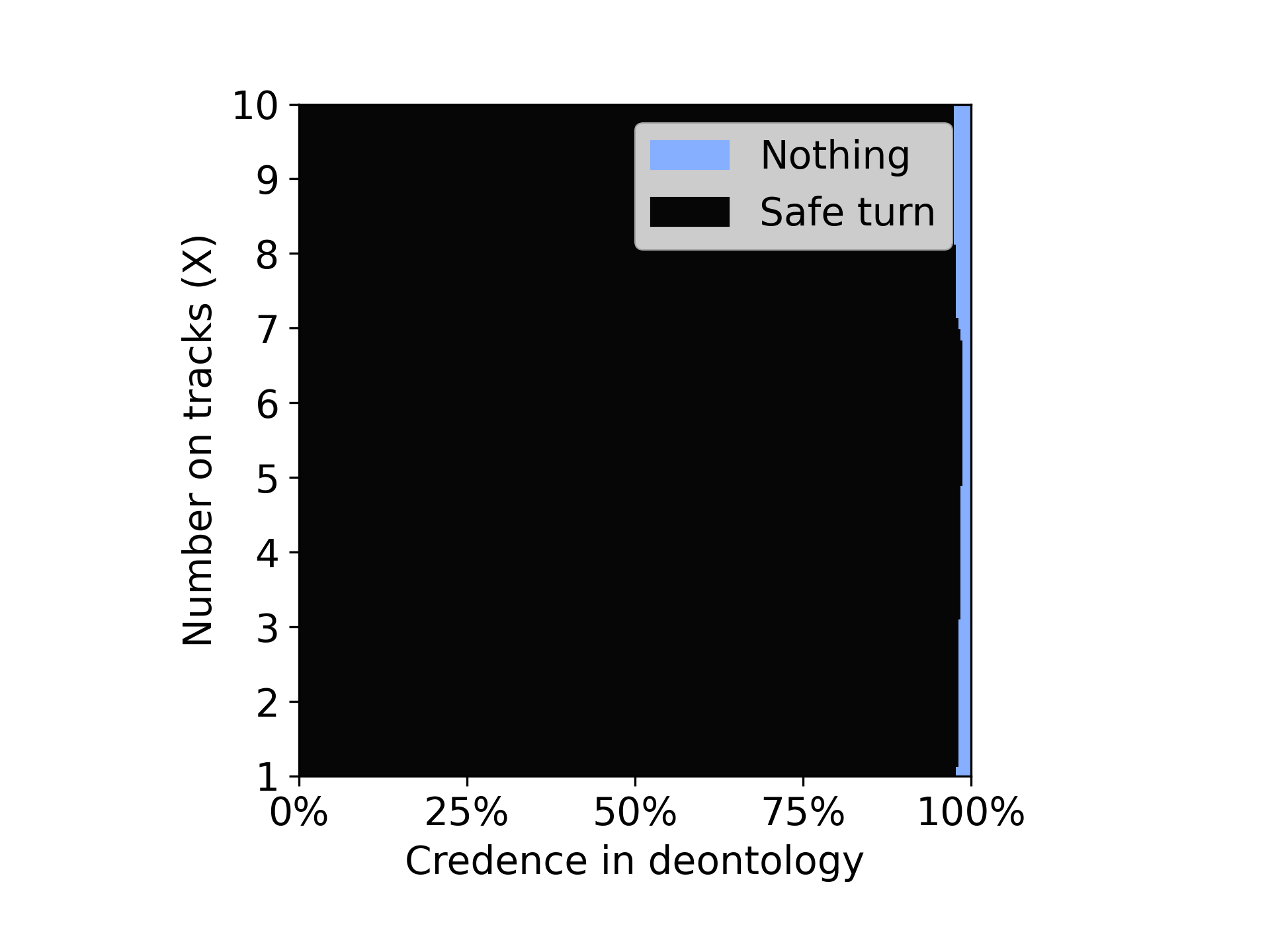
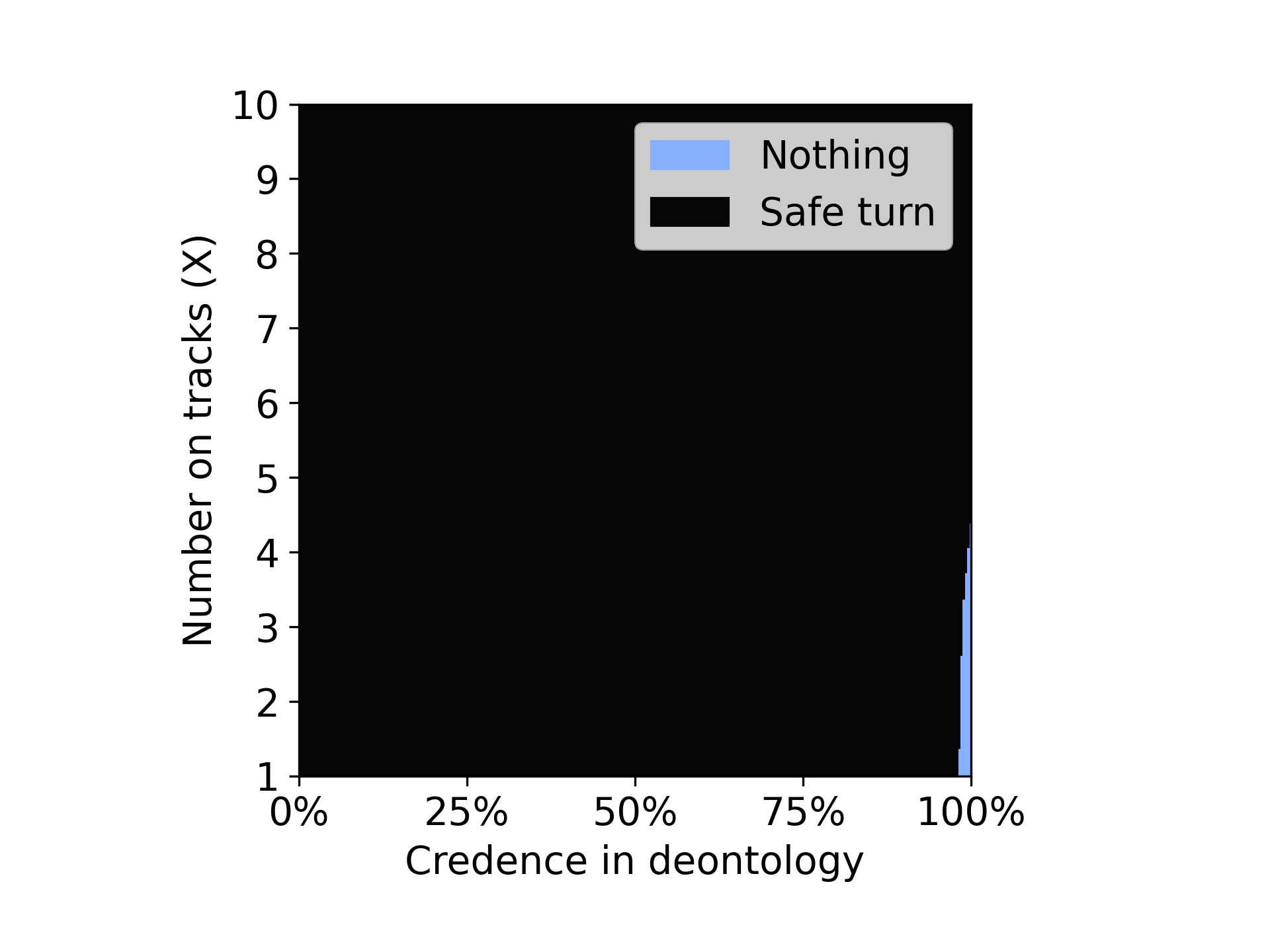


Nothing: Will let the trolley hit the X no of people on track

Safe turn: will change the direction of the trolley towards the side track where there are no people

**Variance Voting**

Dempster Credence Random Credence

From the above plots its clear that the agent with Dempster credence was able to choose nothing option even though choosing nothing is not any theory’s preference. Where as in the random credence case we can see that the agent has chosen safe turn the most of the times.