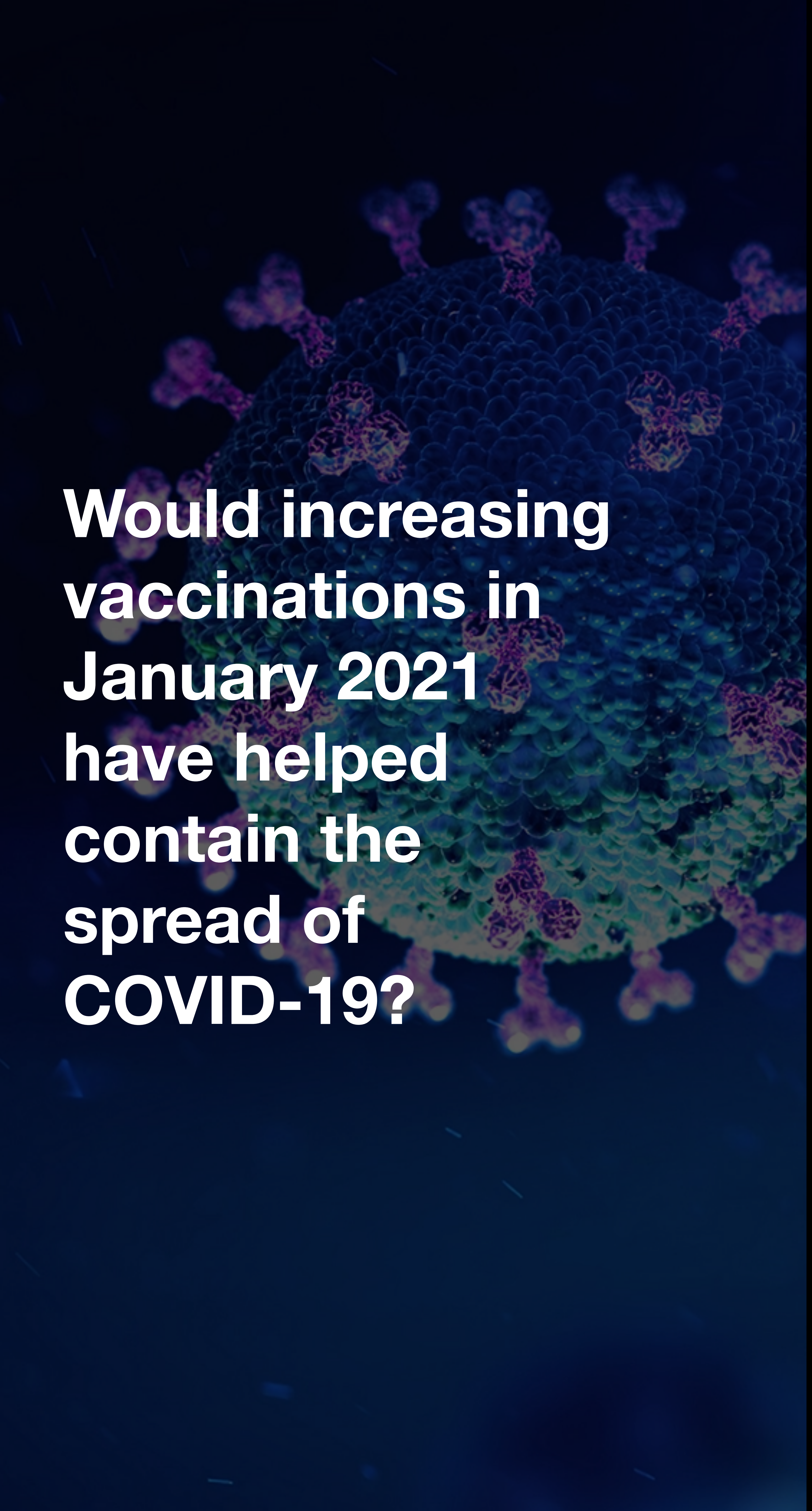
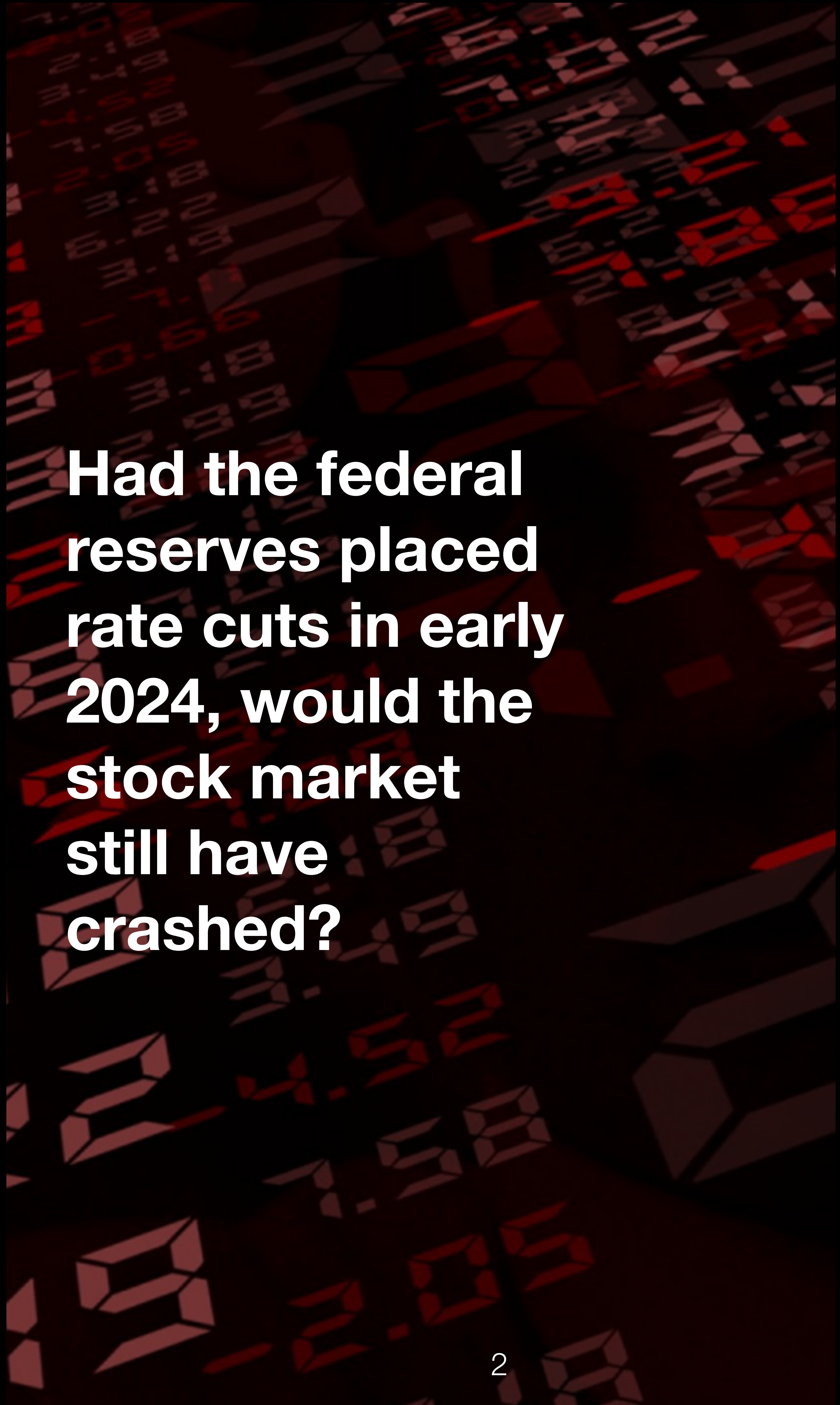
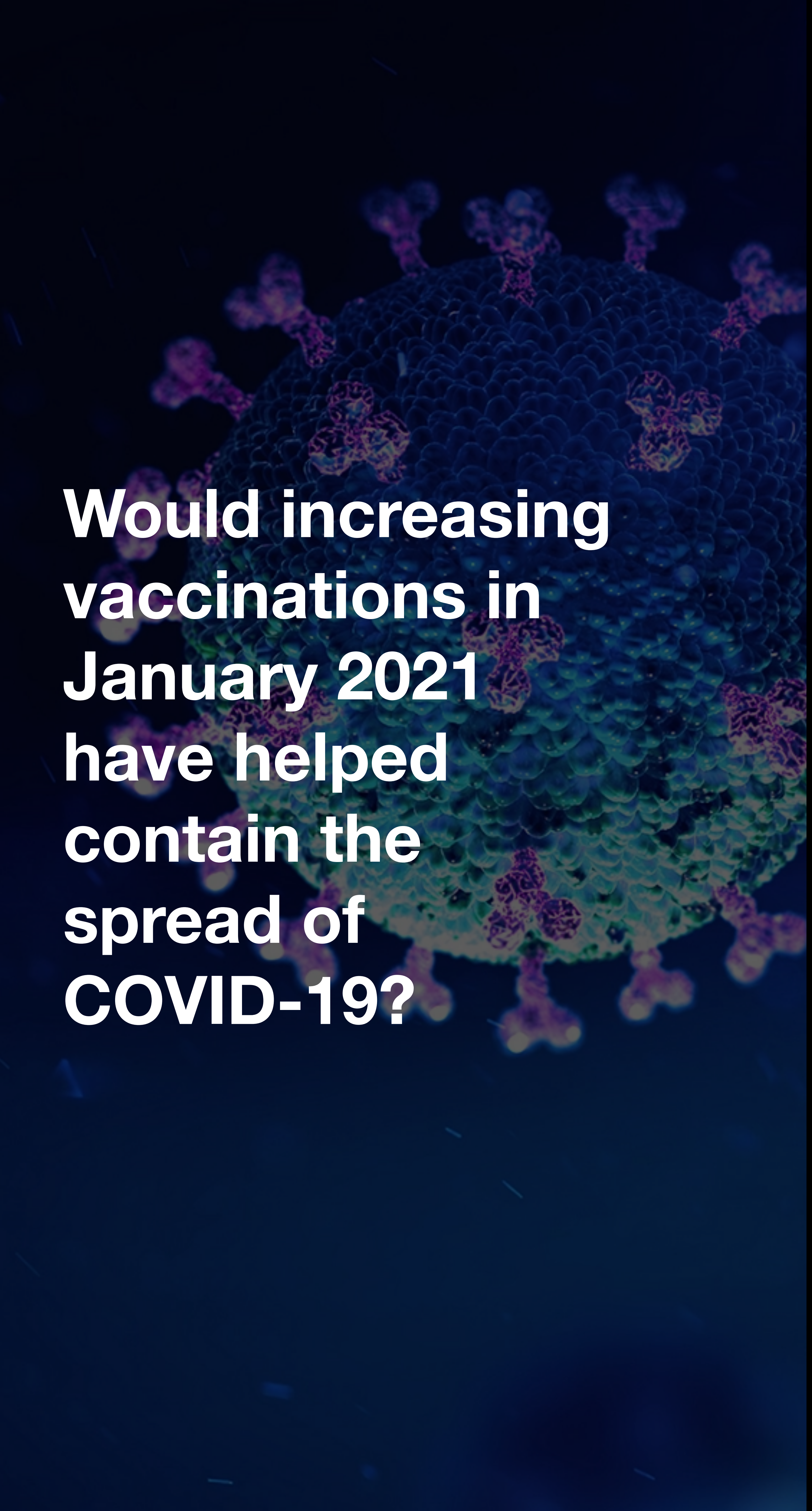
Peculiarities of Counterfactual Point Process Generation

Gerrit Großmann, **Sumantrak Mukherjee**, Sebastian Vollmer

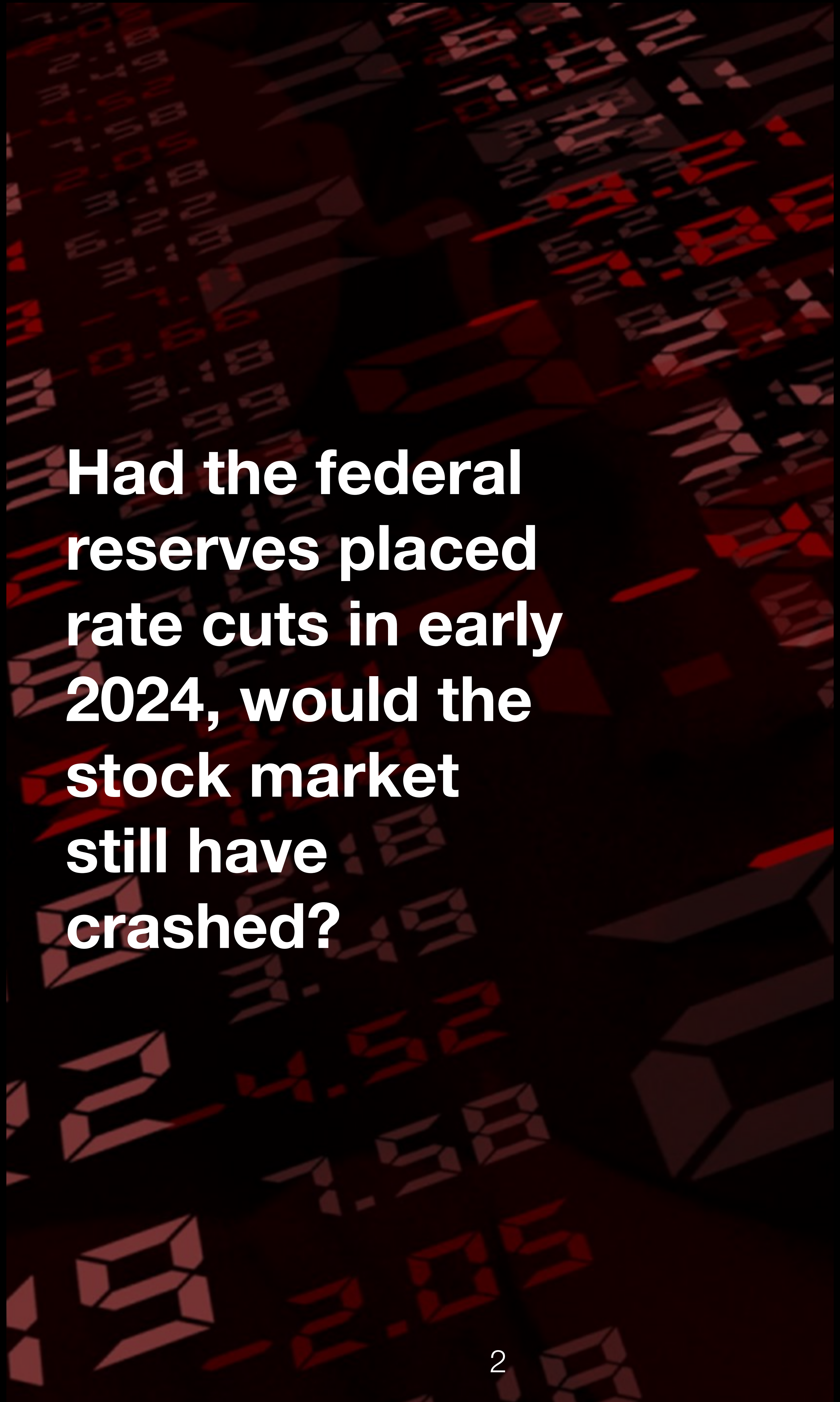
Would increasing vaccinations in January 2021 have helped contain the spread of COVID-19?


Would increasing vaccinations in January 2021 have helped contain the spread of COVID-19?

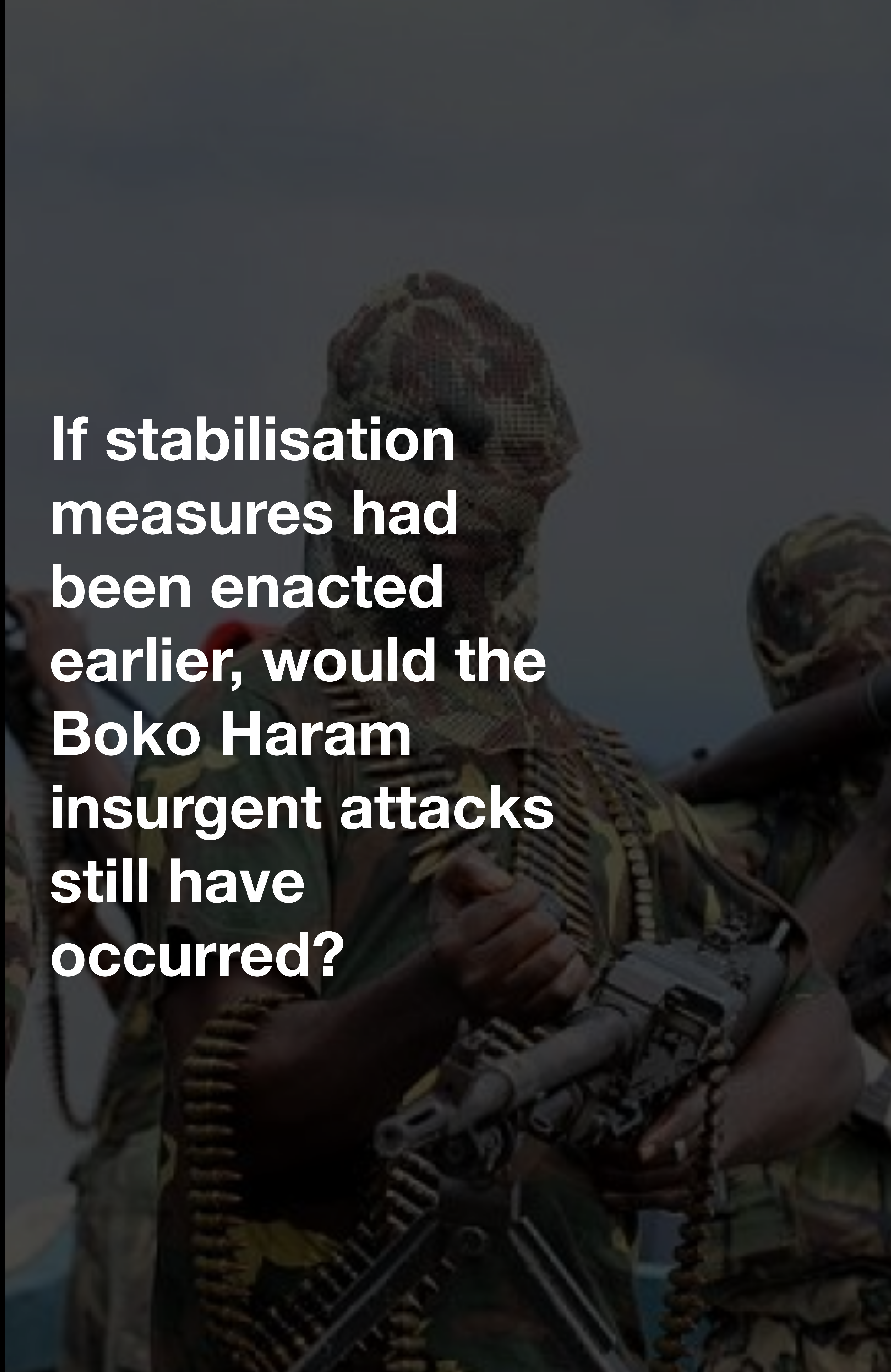

Had the federal reserves placed rate cuts in early 2024, would the stock market still have crashed?



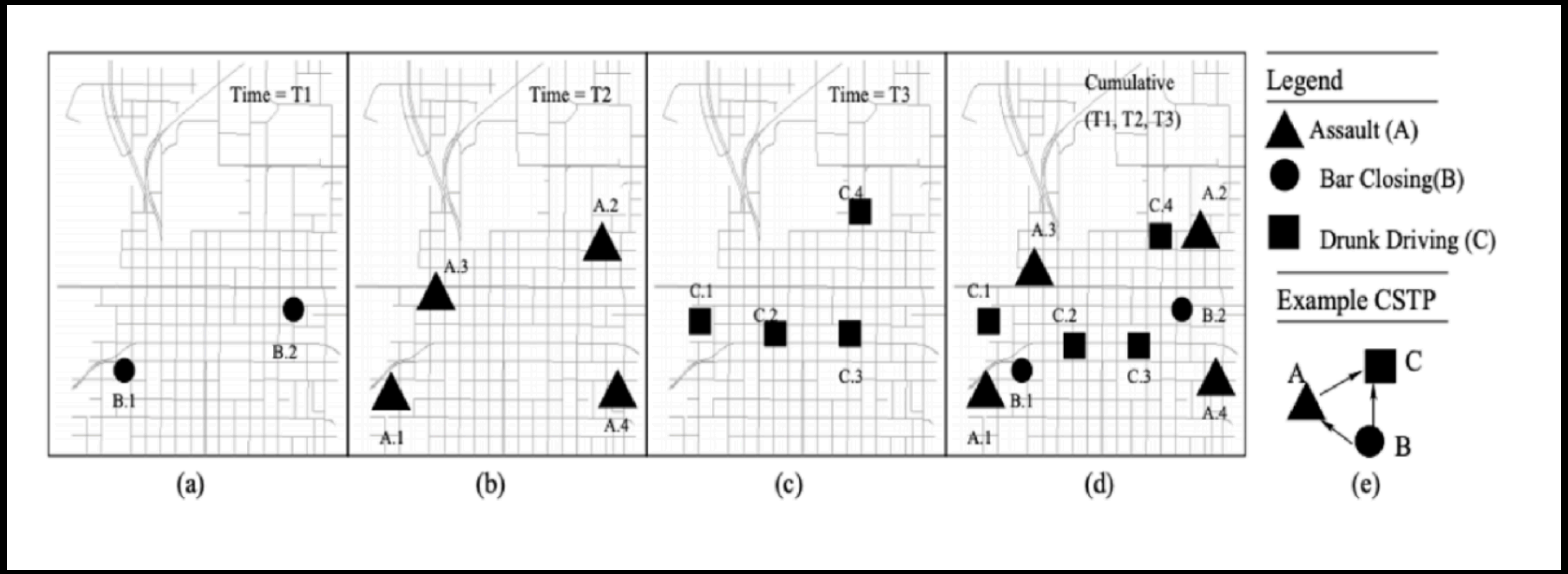
Would increasing vaccinations in January 2021 have helped contain the spread of COVID-19?



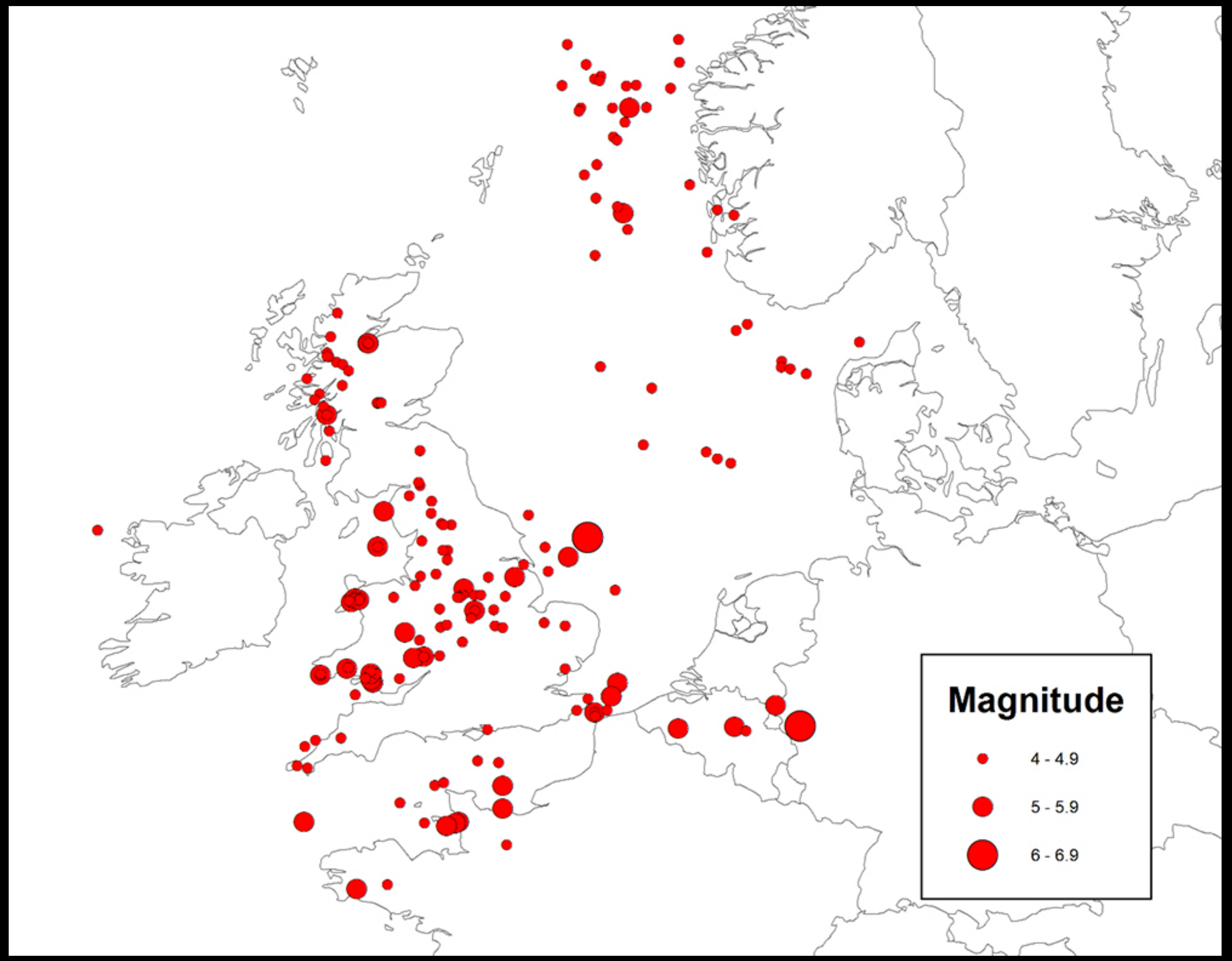
Had the federal reserves placed rate cuts in early 2024, would the stock market still have crashed?



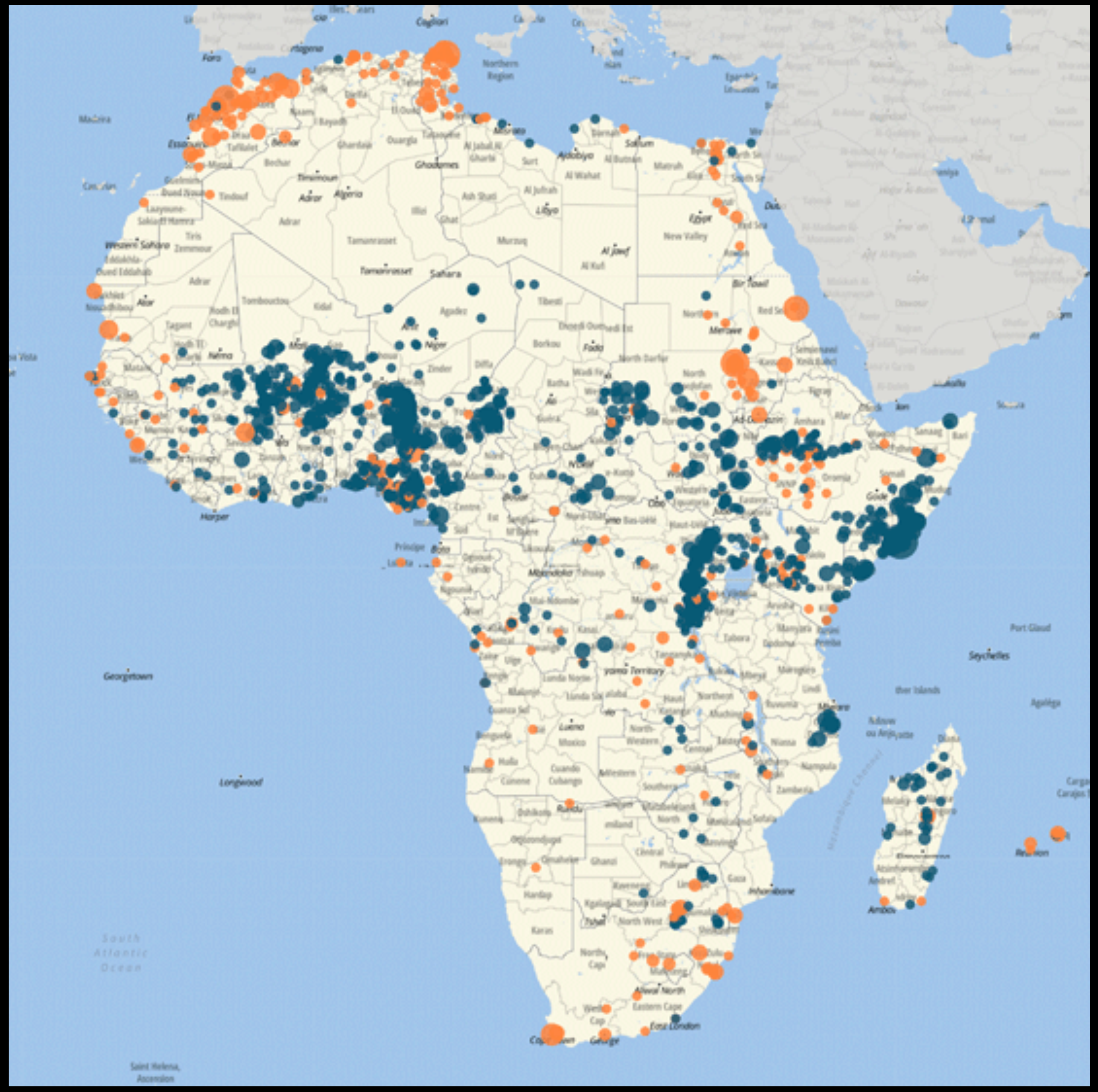
If stabilisation measures had been enacted earlier, would the Boko Haram insurgent attacks still have occurred?



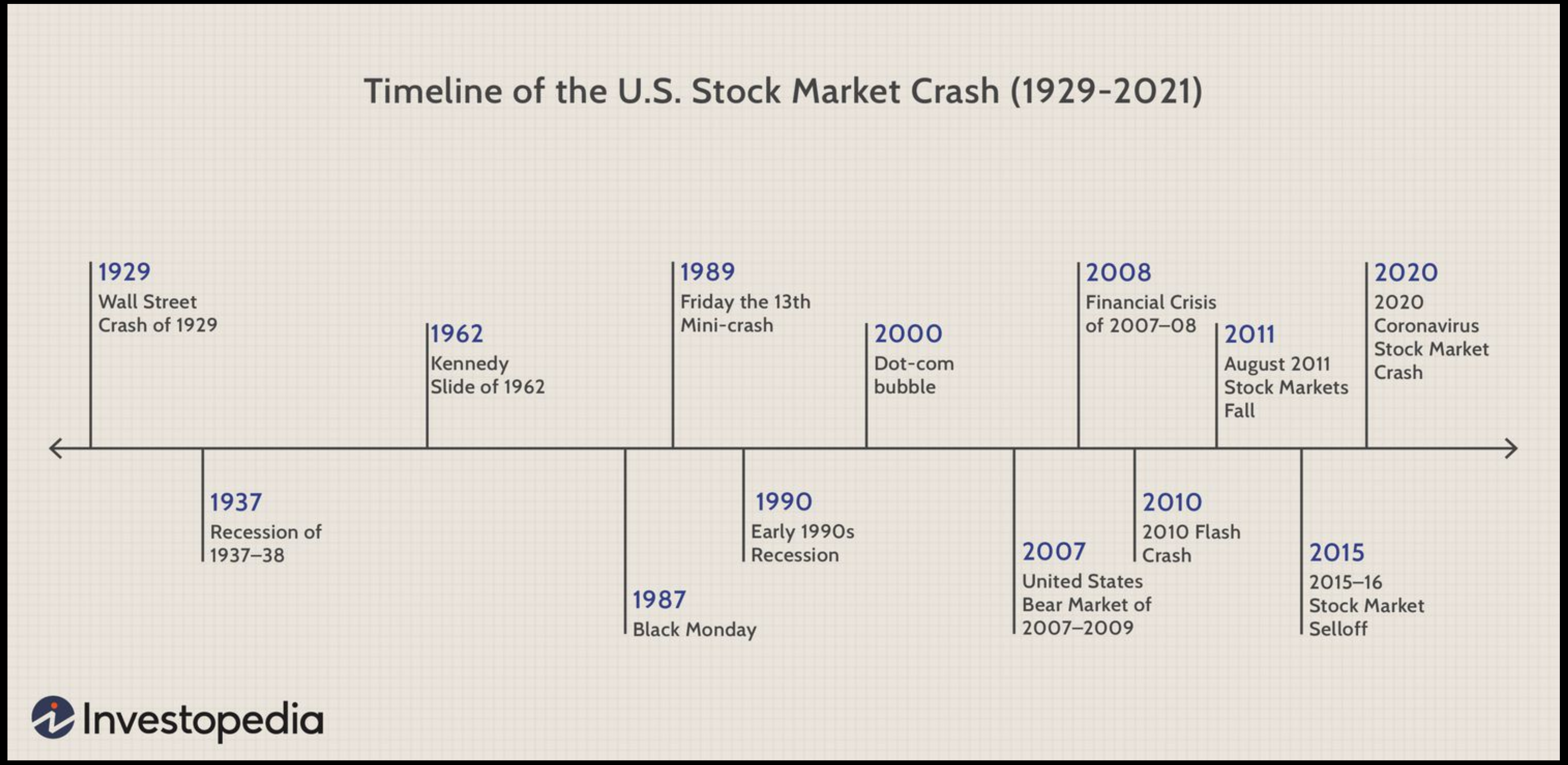
Emergency Calls



Earthquakes

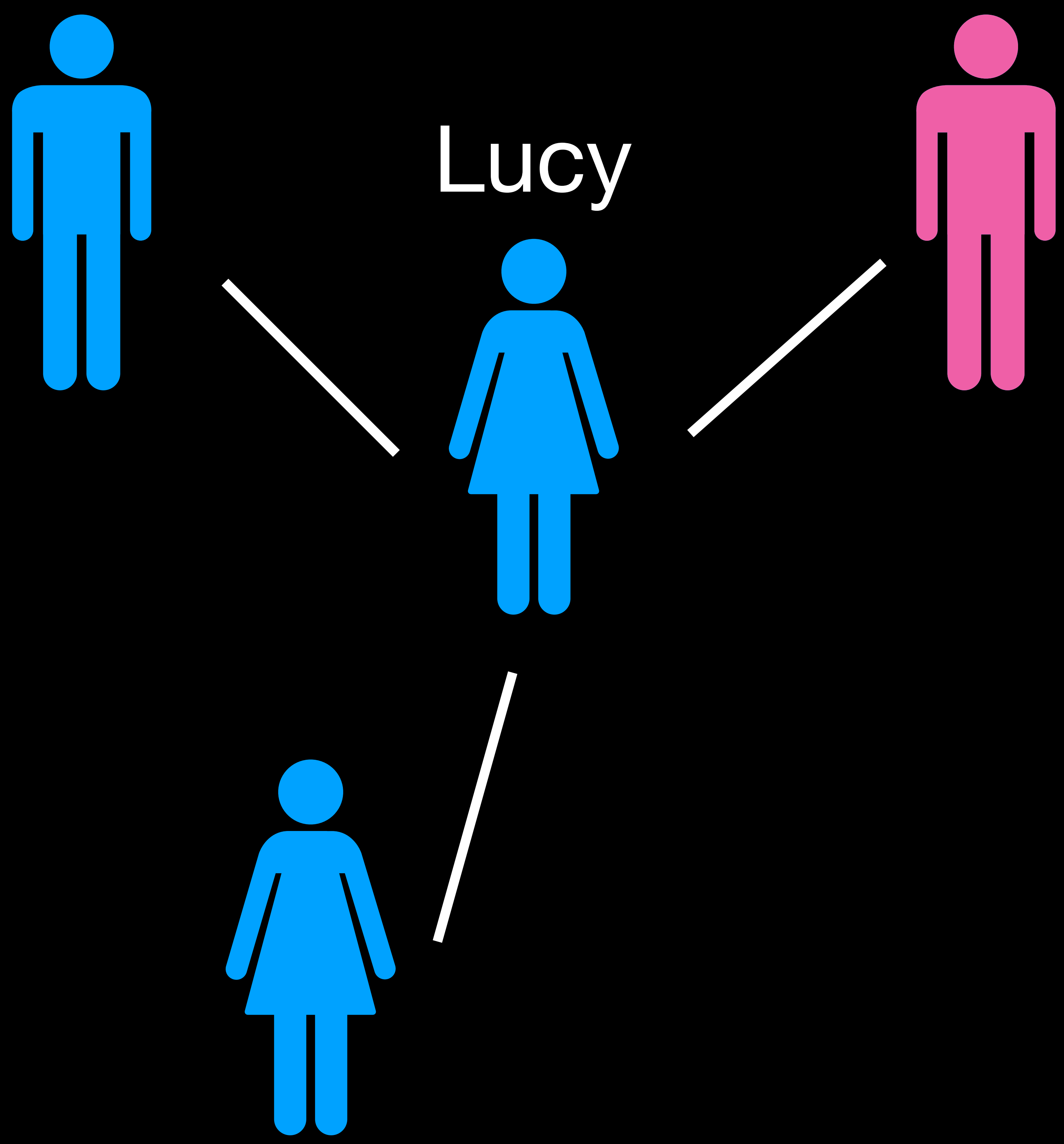


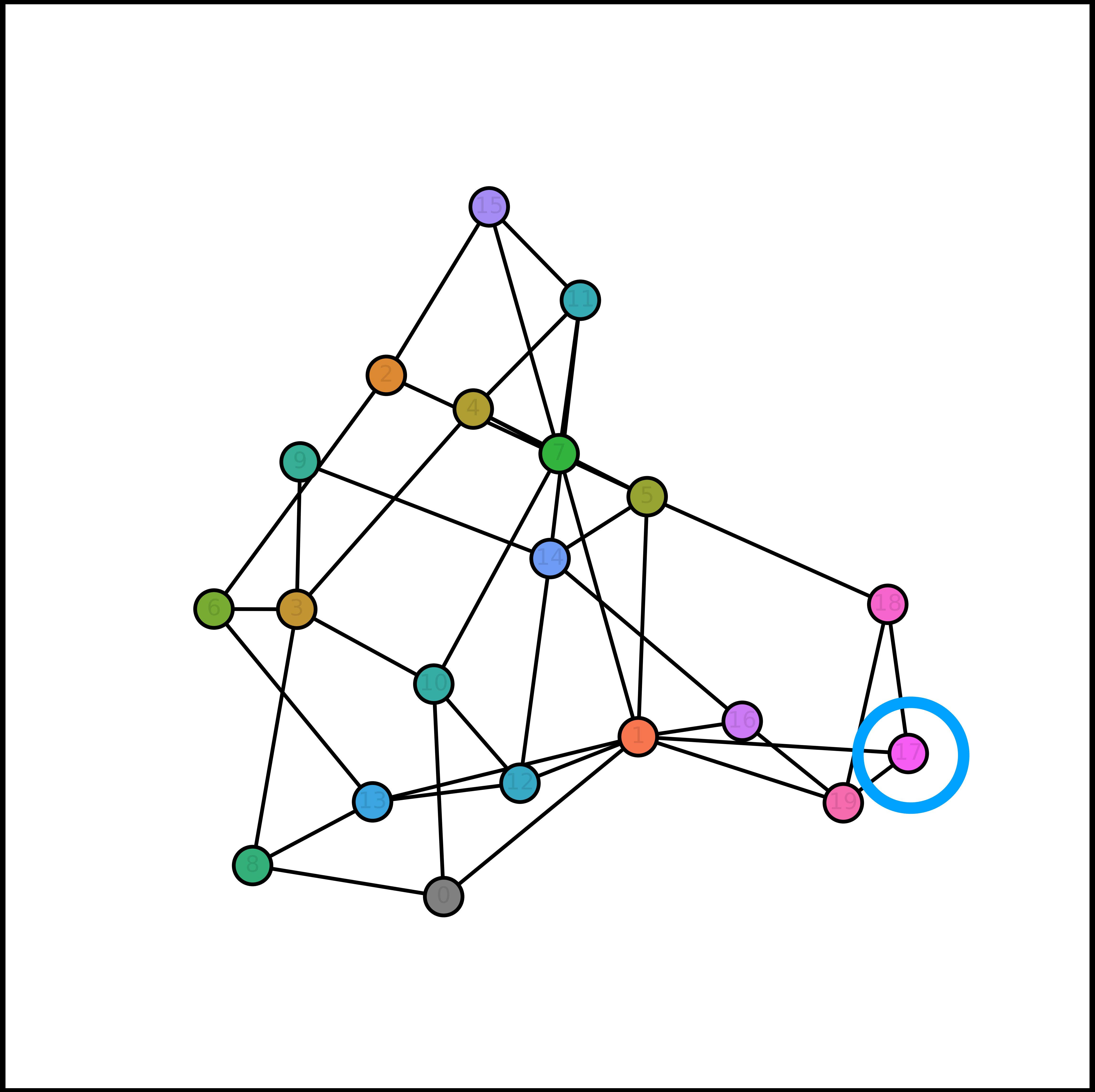
Armed Conflicts



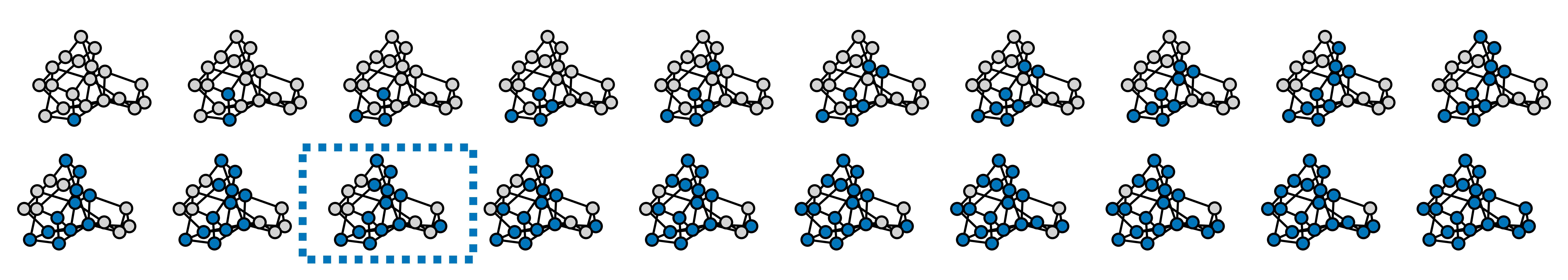
Stock market crashes

When Will Lucy Get Infected?

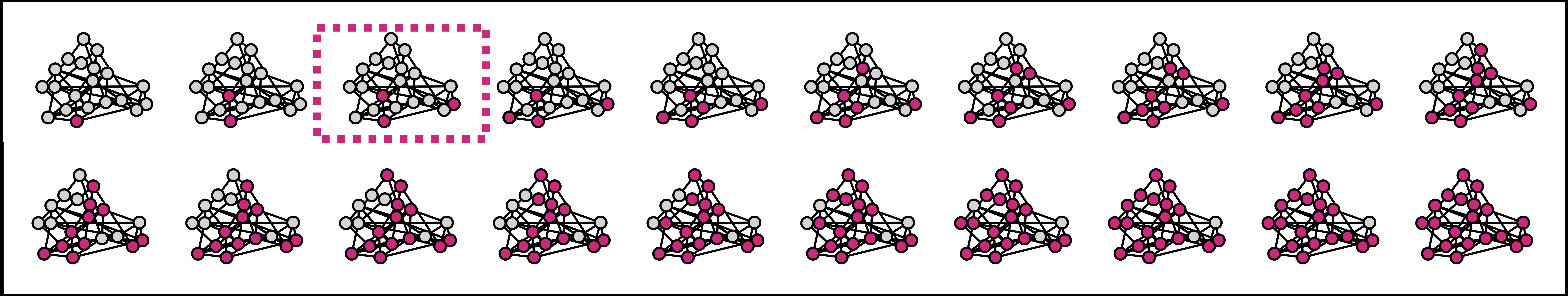




Would Lucy Have Gotten Infected Earlier ?



Observed Trajectory



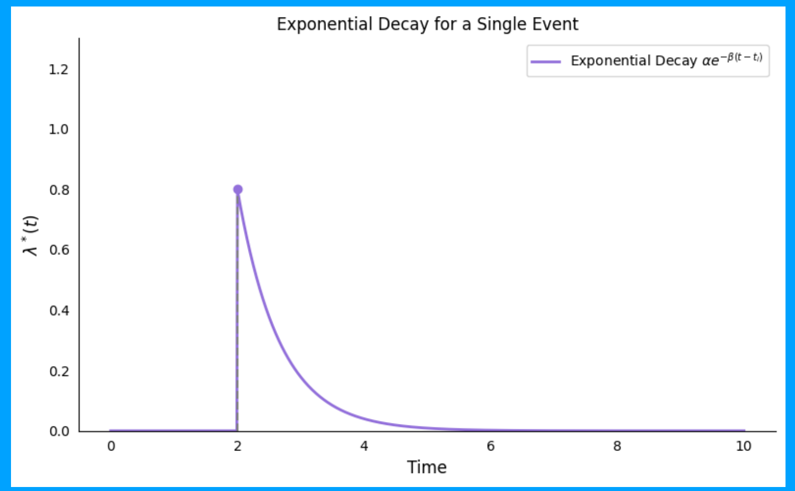
Counterfactual Trajectory

Modelling Events Using Point Processes

Modelling Events Using Point Processes

History

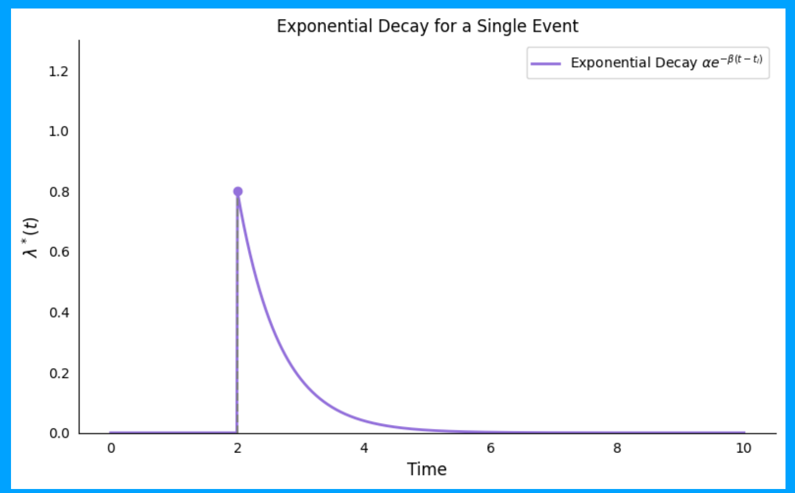
Dynamics and
Background rate

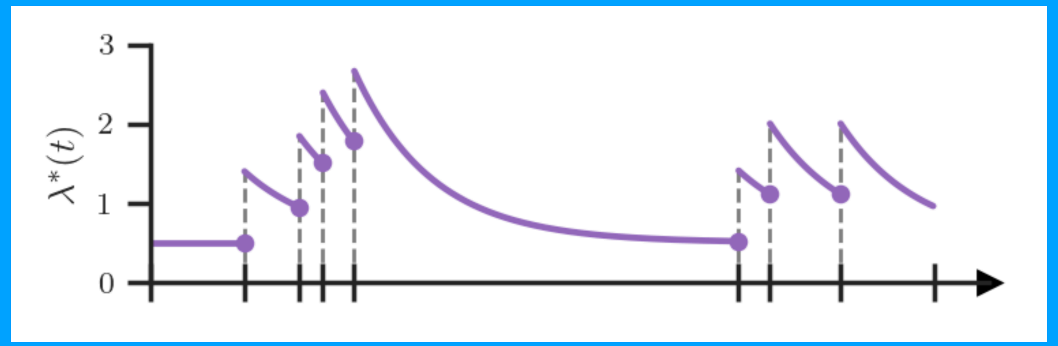


Modelling Events Using Point Processes

History

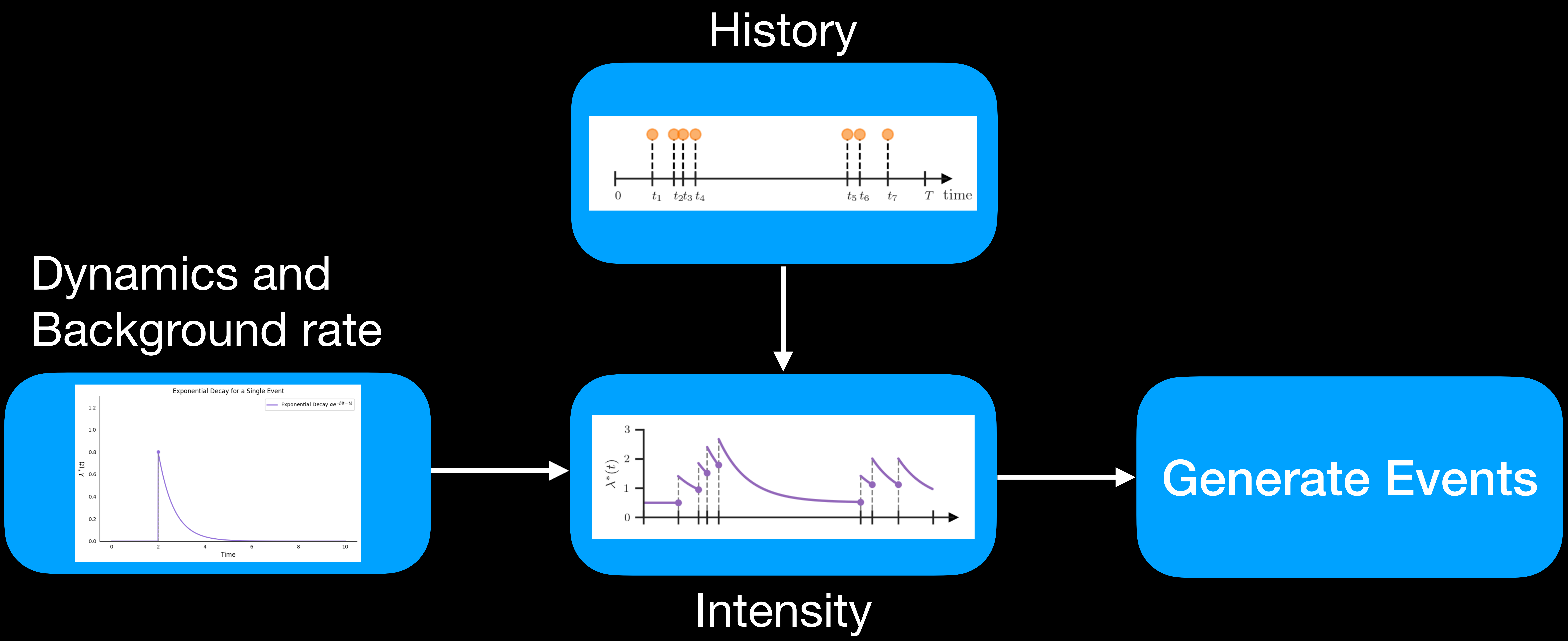
Dynamics and
Background rate



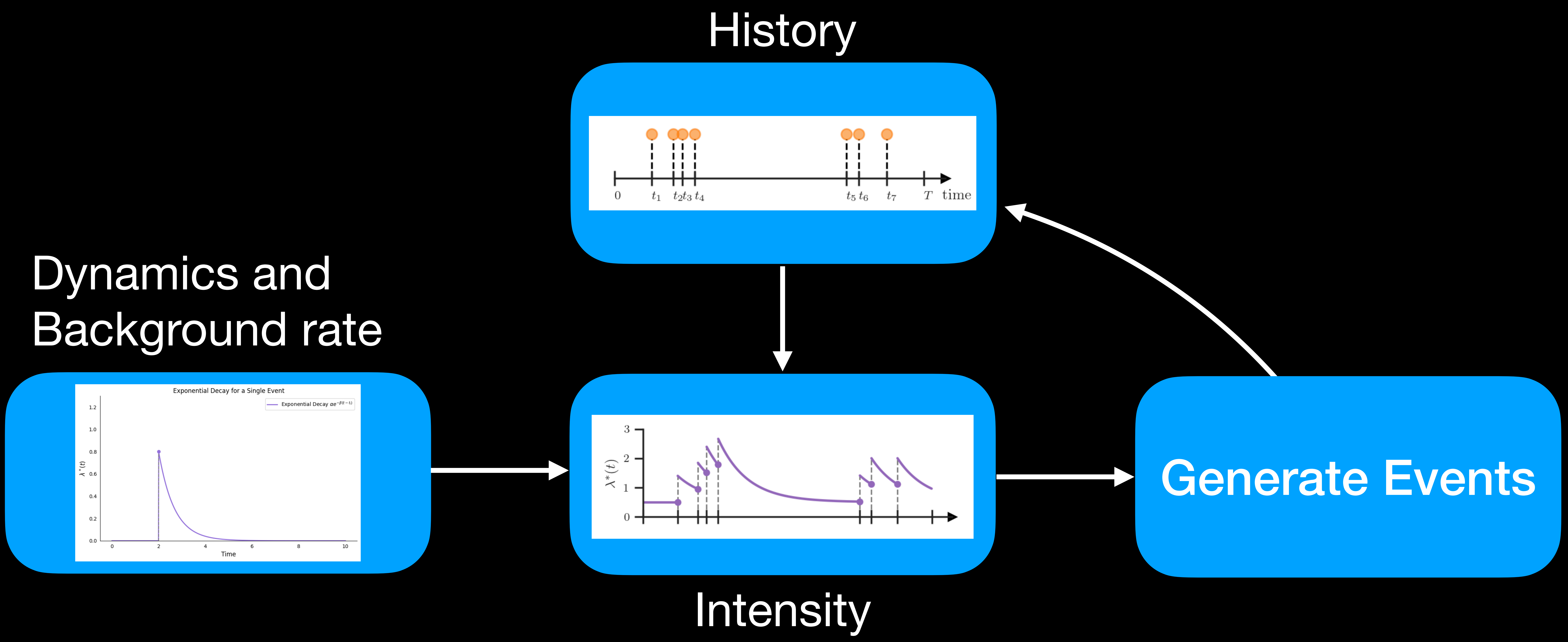


Intensity

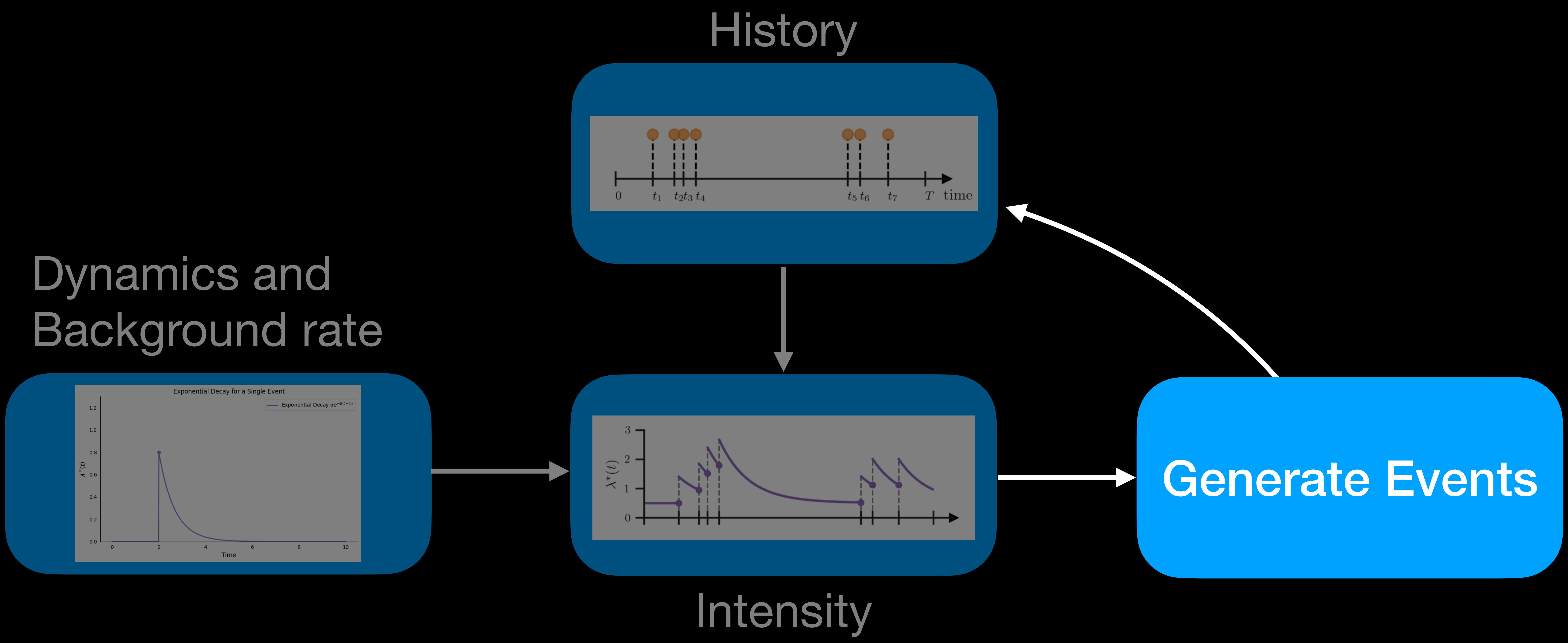
Modelling Events Using Point Processes



Modelling Events Using Point Processes



Modelling Events Using Point Processes



In our paper, we focus on the methods used to generate events

Does the method(+assumptions) used to generate events from the intensity affect the counterfactual sequence?

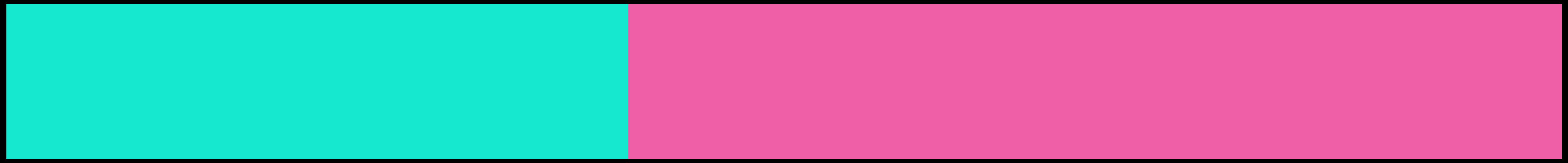
Yes, lets explore how.

Counterfactuals

Counterfactuals

$P(H) = 0.4$

$P(T) = 0.6$

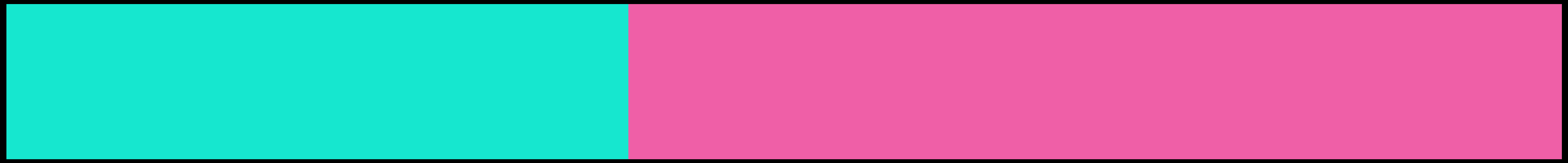


Counterfactuals

$$u \sim U([0,1])$$

P(H) = 0.4

P(T) = 0.6

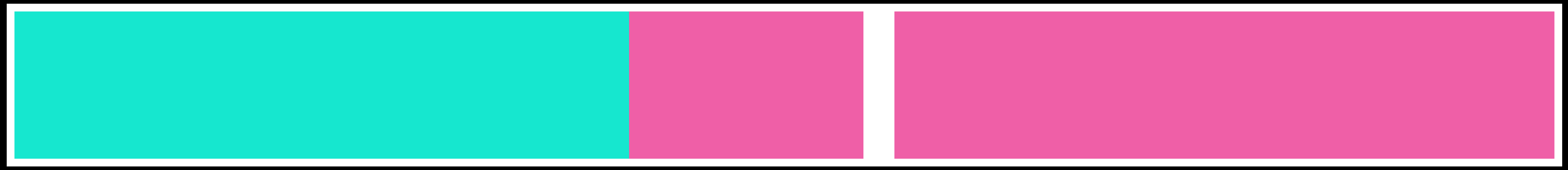


Counterfactuals

$$u \sim U([0,1])$$

P(H) = 0.4

P(T) = 0.6



Counterfactuals

$$u \sim U([0,1])$$

P(H) = 0.4

P(T) = 0.6

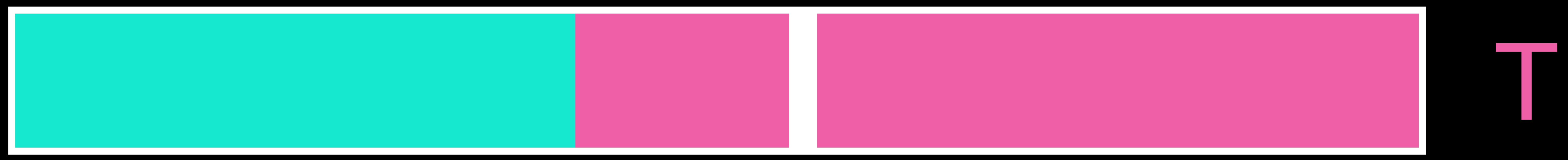
T

Counterfactuals

$$u \sim U([0,1])$$

P(H) = 0.4

P(T) = 0.6



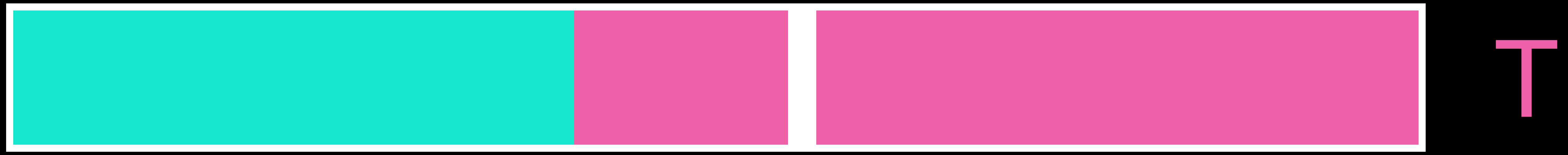
What if?

Counterfactuals

$$u \sim U([0,1])$$

P(H) = 0.4

P(T) = 0.6



What if?

P(H) = 0.6

P(T) = 0.4

Counterfactuals

$u \sim U([0,1])$

P(H) = 0.4

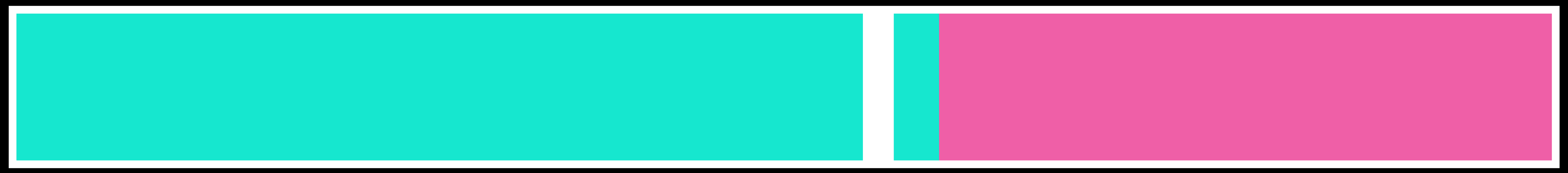
P(T) = 0.6

T

What if?

P(H) = 0.6

P(T) = 0.4



H

Counterfactuals

$$u \sim U([0,1])$$

P(H) = 0.4

P(T) = 0.6

T

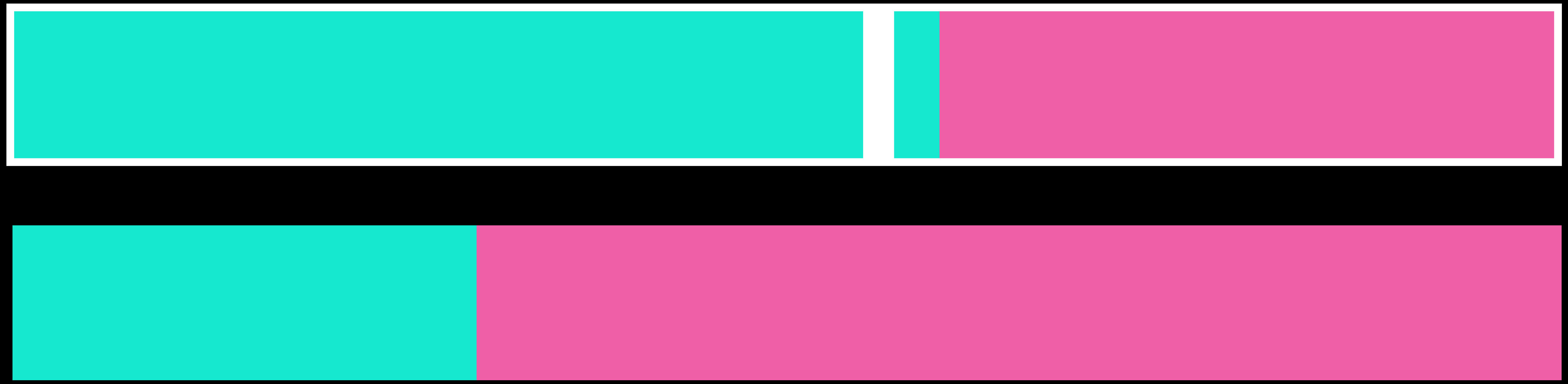
What if?

P(H) = 0.6

P(T) = 0.4

P(H) = 0.3

P(T) = 0.7



H

Counterfactuals

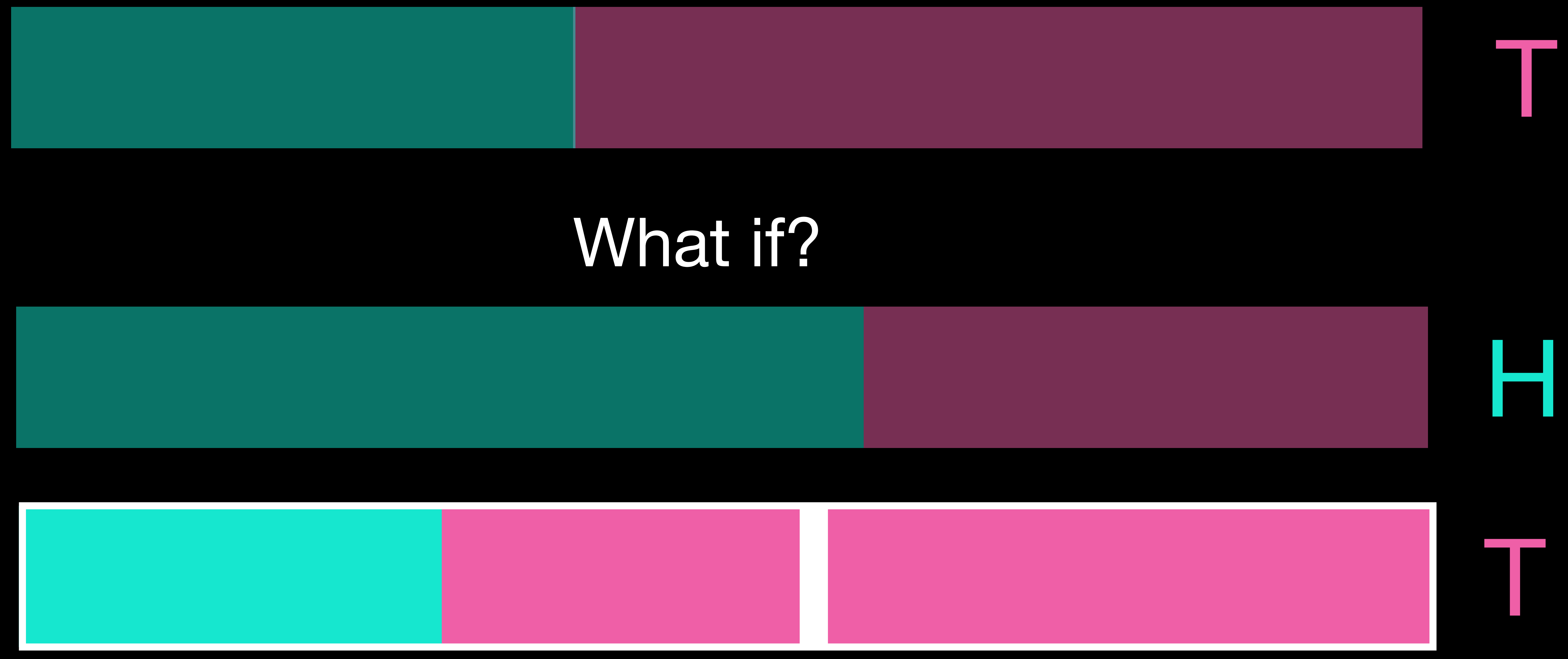
$$u \sim U([0,1])$$

P(H) = 0.4
P(T) = 0.6

What if?

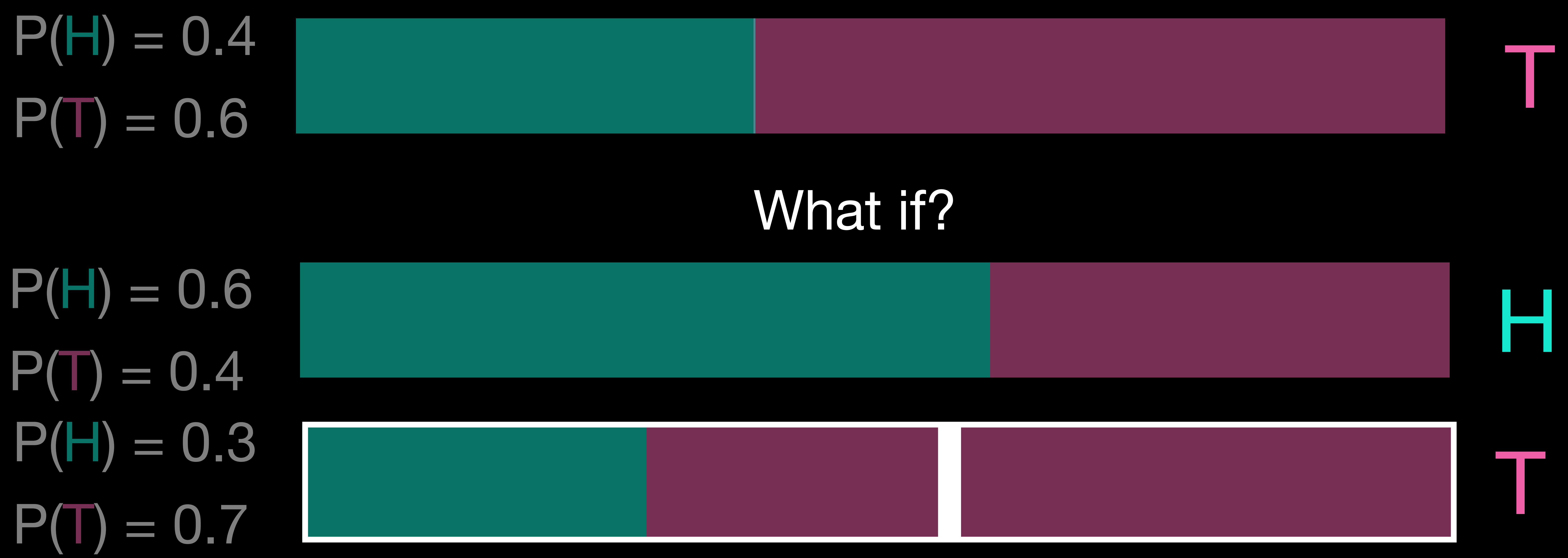
P(H) = 0.6
P(T) = 0.4

P(H) = 0.3
P(T) = 0.7



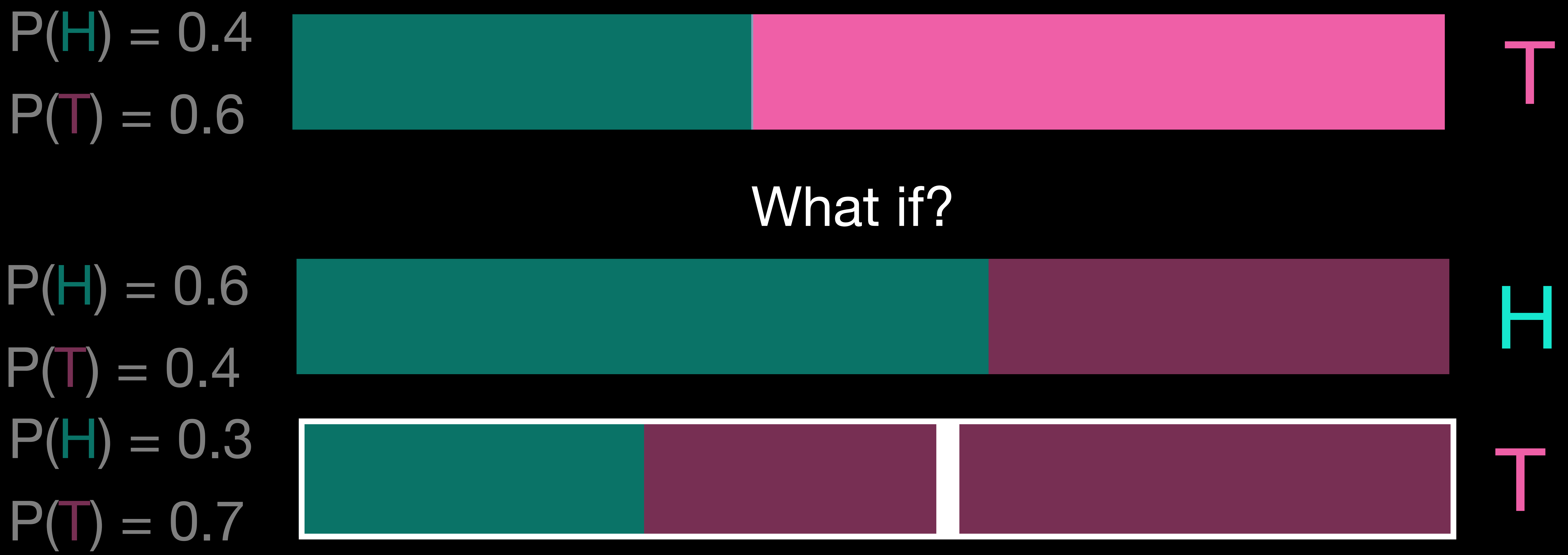
Counterfactuals

$$u \sim U([0,1])$$



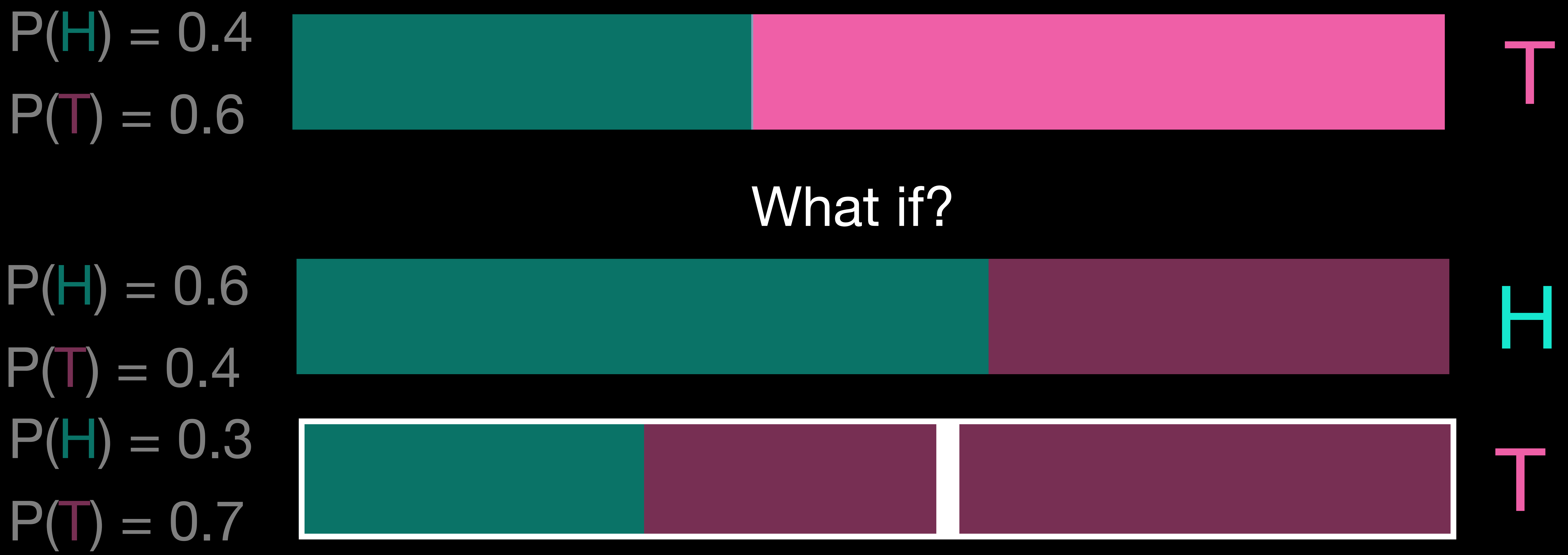
Counterfactuals

$$u \sim U([0,1])$$



Counterfactuals

$u \sim U([0,1])$



If outcome is T $u^{cf} \sim U([P(H),1])$

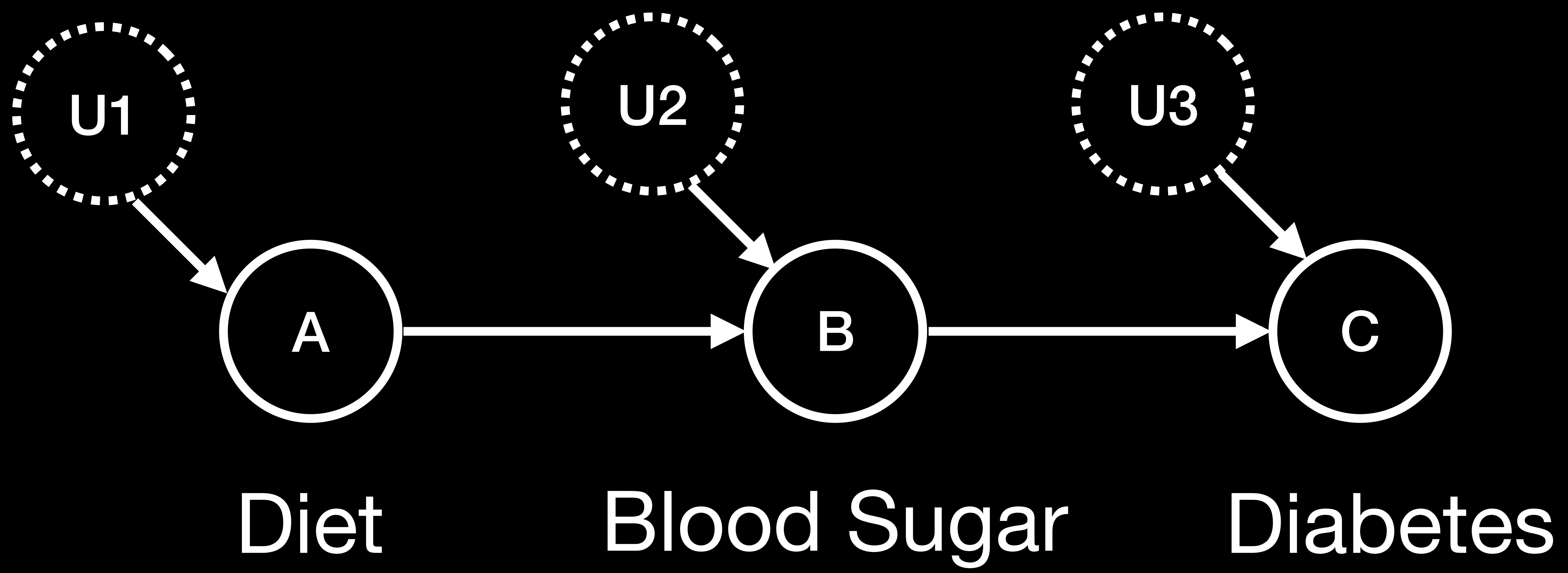
If outcome is H $u^{cf} \sim U([0,P(H)])$

Structural Causal Model (SCM)

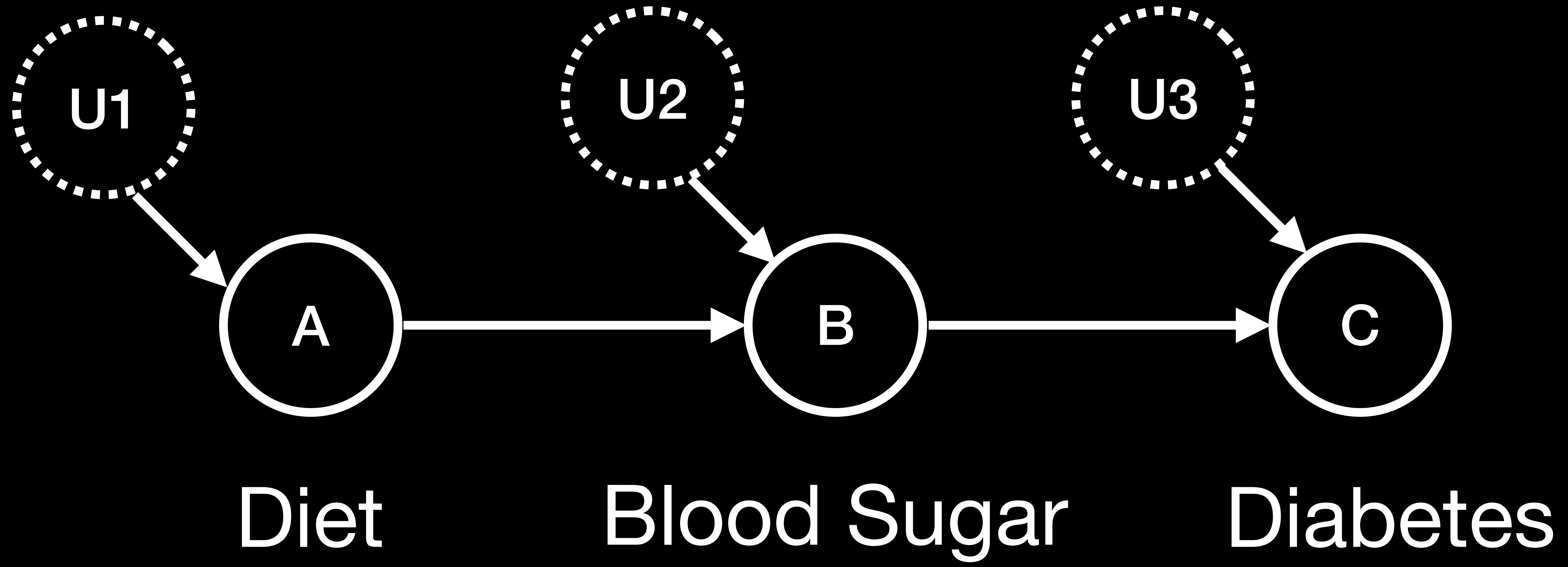
Structural Causal Model (SCM)

Structural Causal Model (SCM)

Structural Causal Model (SCM)



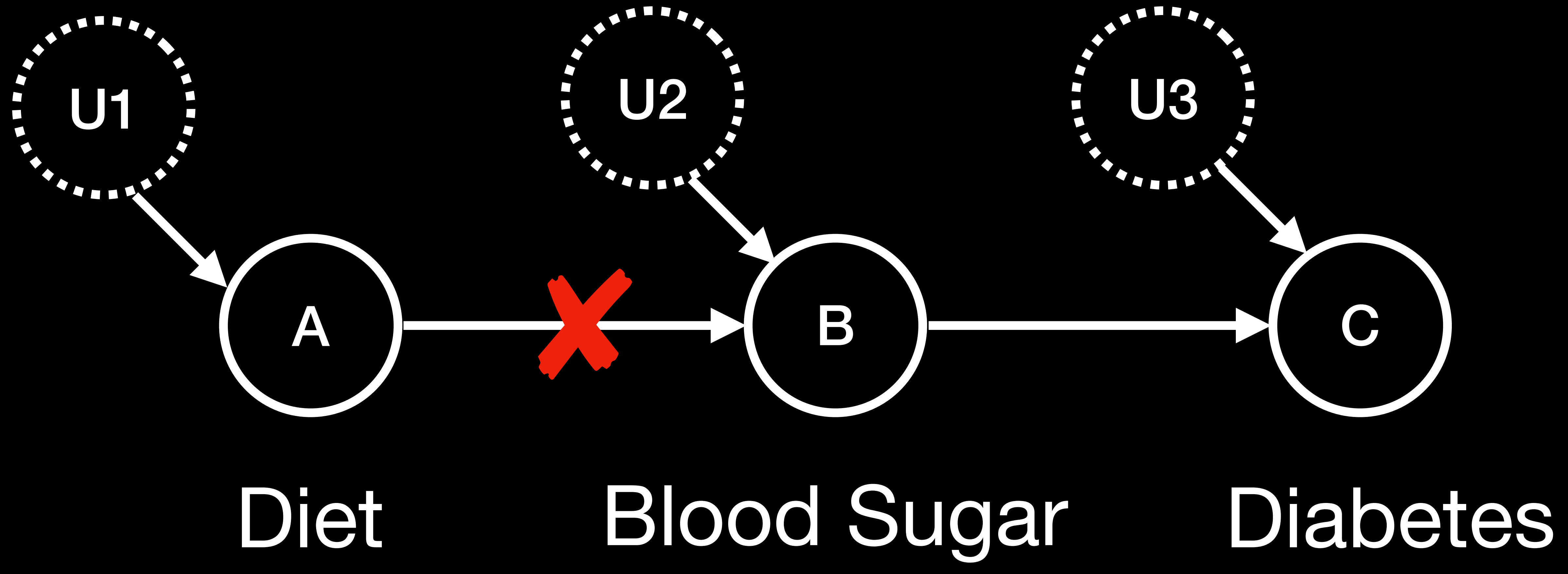
Structural Causal Model (SCM)



How does diet influence diabetes?

Observations : Seeing

Structural Causal Model (SCM)



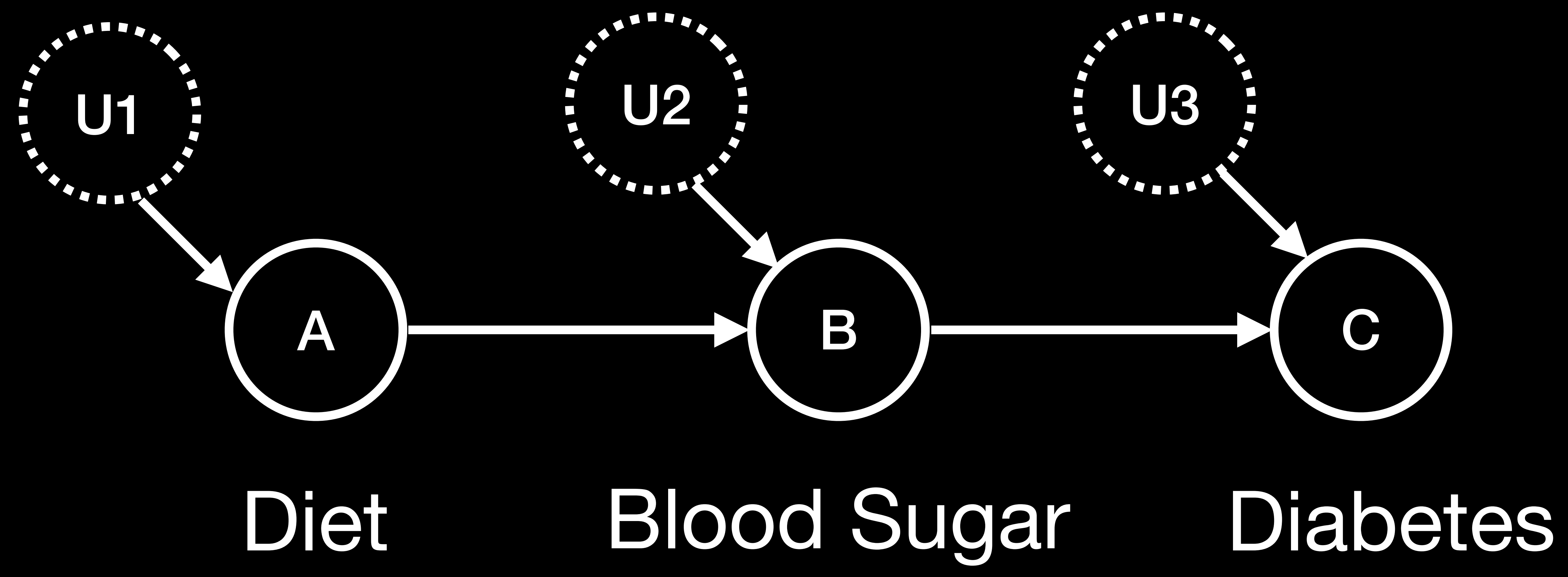
How does diet influence diabetes?

Observations : Seeing

Does treating blood sugar prevent diabetes

Interventions : Doing

Structural Causal Model (SCM)



How does diet influence diabetes?

Observations : Seeing

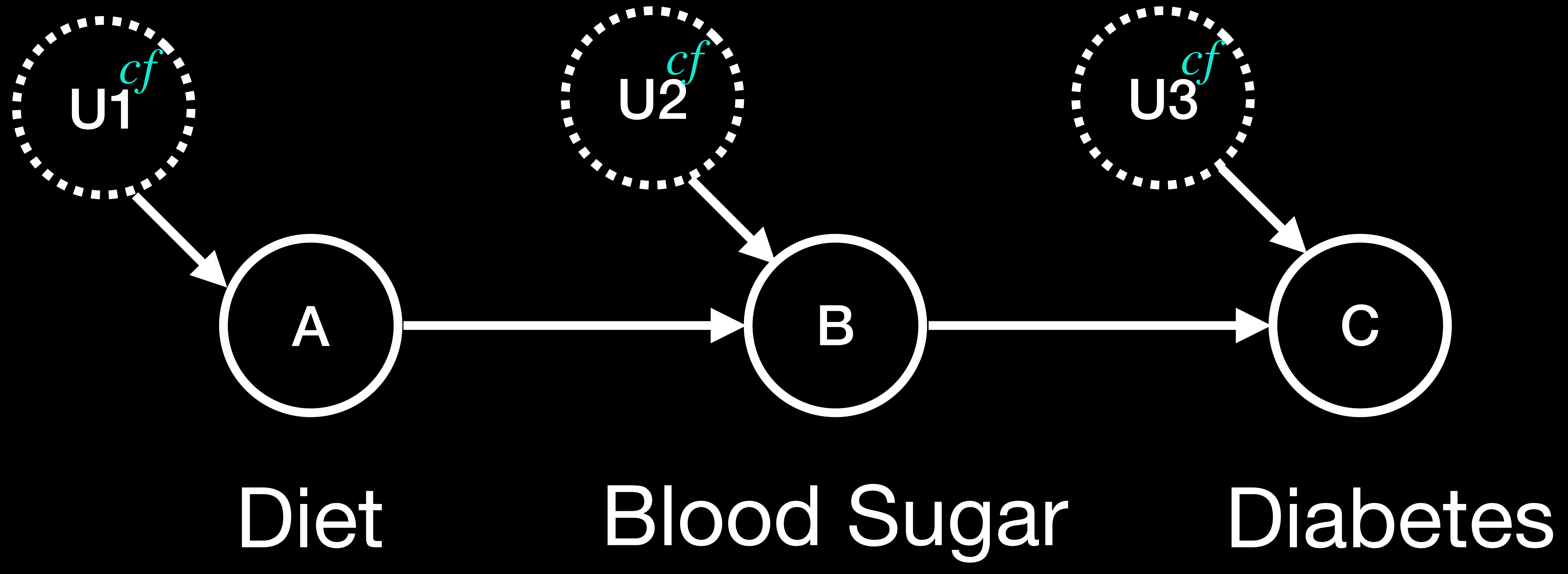
Does treating blood sugar prevent diabetes

Interventions : Doing

Would Lucy still have gotten diabetes if her blood sugar was controlled

Counterfactuals : Imagining

Structural Causal Model (SCM)



How does diet influence diabetes?

Observations : Seeing

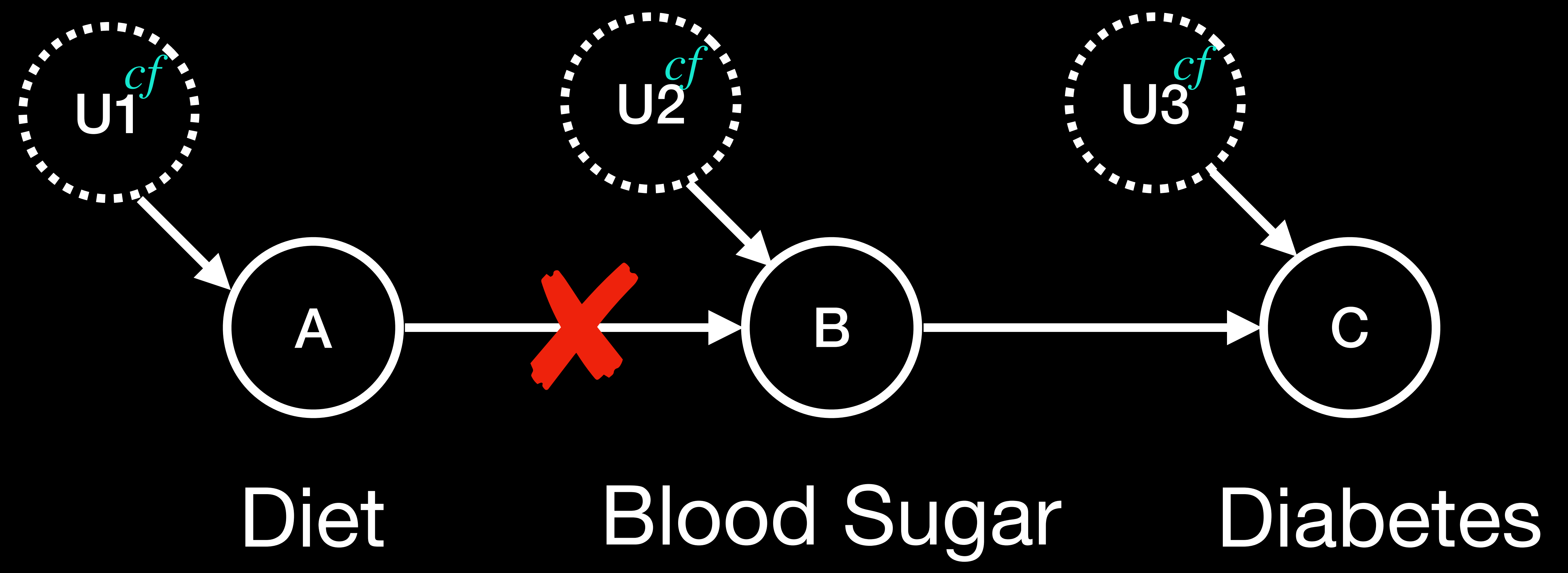
Does treating blood sugar prevent diabetes

Interventions : Doing

Would Lucy still have gotten diabetes if her blood sugar was controlled

Counterfactuals : Imagining

Structural Causal Model (SCM)



How does diet influence diabetes?

Observations : Seeing

Does treating blood sugar prevent diabetes

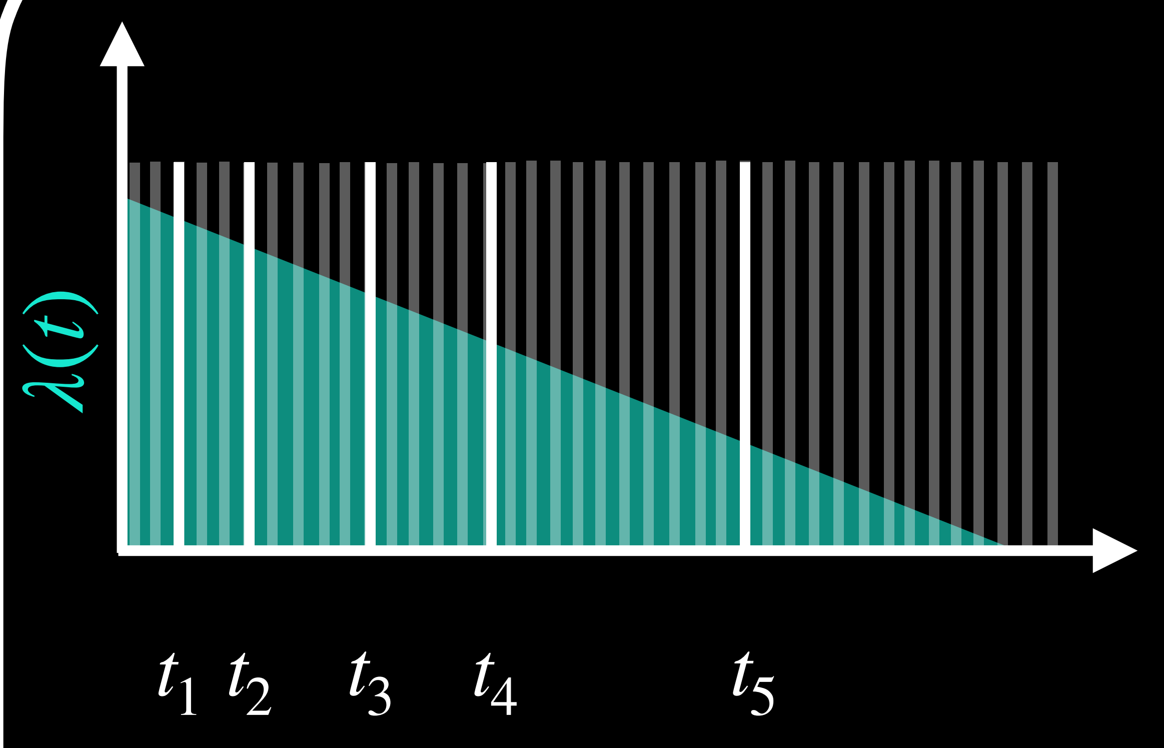
Interventions : Doing

Would Lucy still have gotten diabetes if her blood sugar was controlled

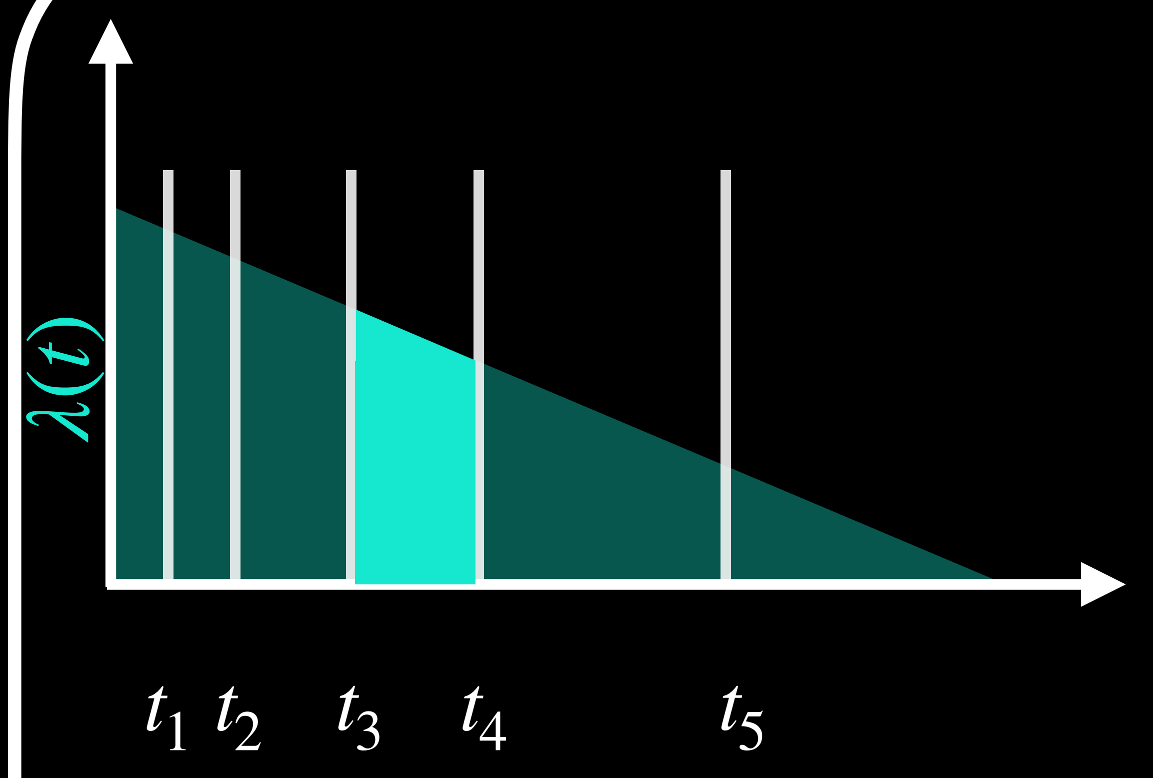
Counterfactuals : Imagining

Simulation algorithms and their SCMs

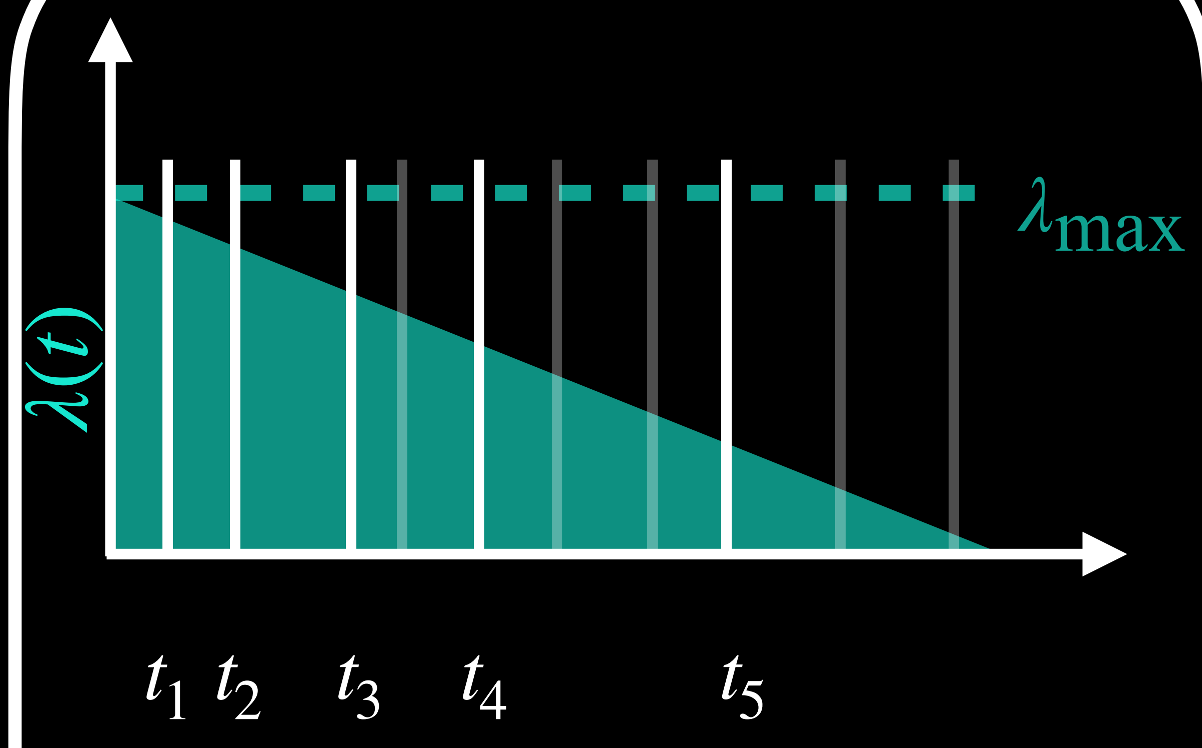
Statistically Equivalent Event Generation



Naïve Method



Numerical Integration

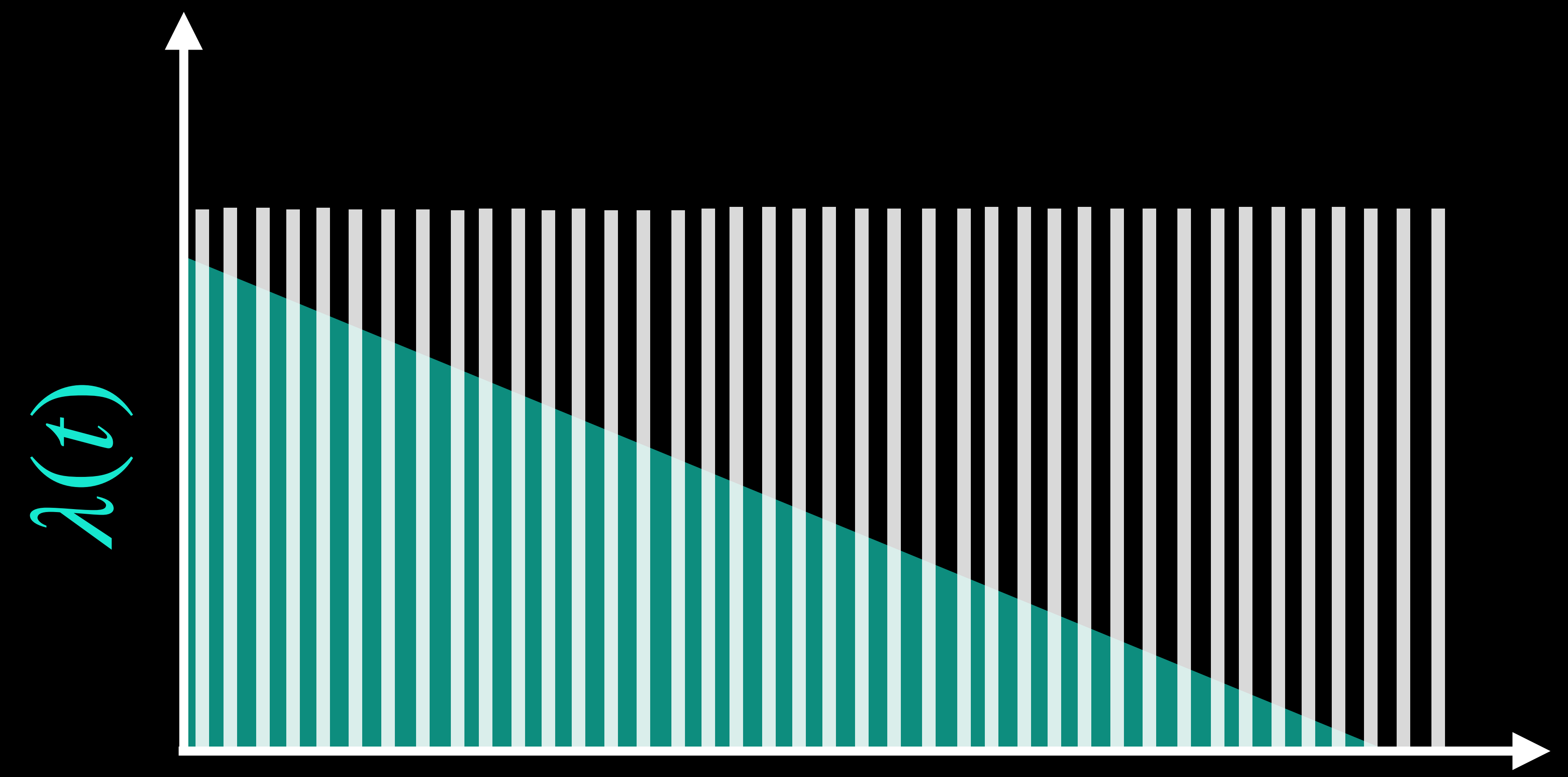


Thinning

Naïve Method

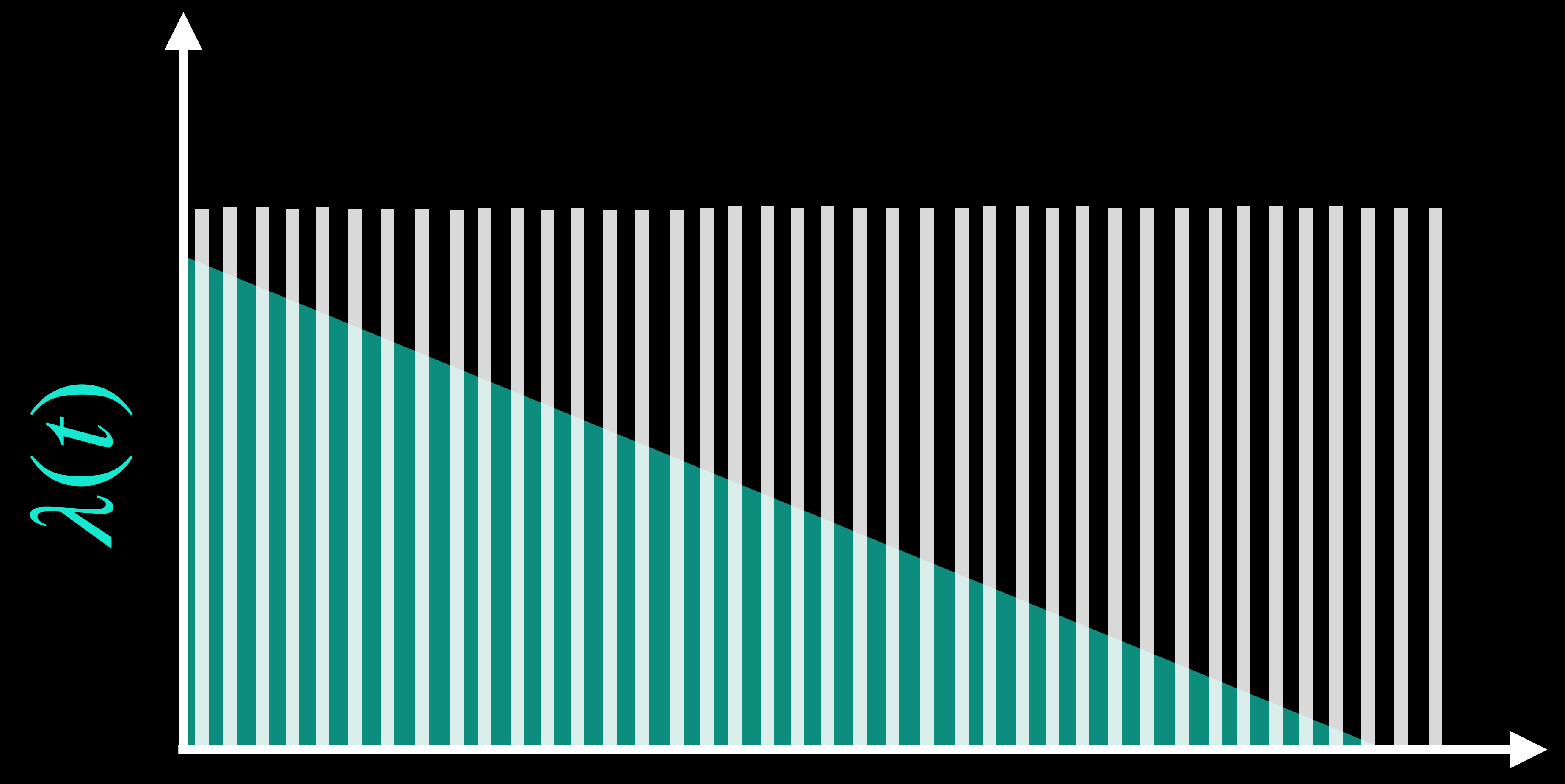
Naïve Method

Discretise the time axis into very small time intervals Δ and for each Δ calculate



Naïve Method

Discretise the time axis into very small time intervals Δ and for each Δ calculate



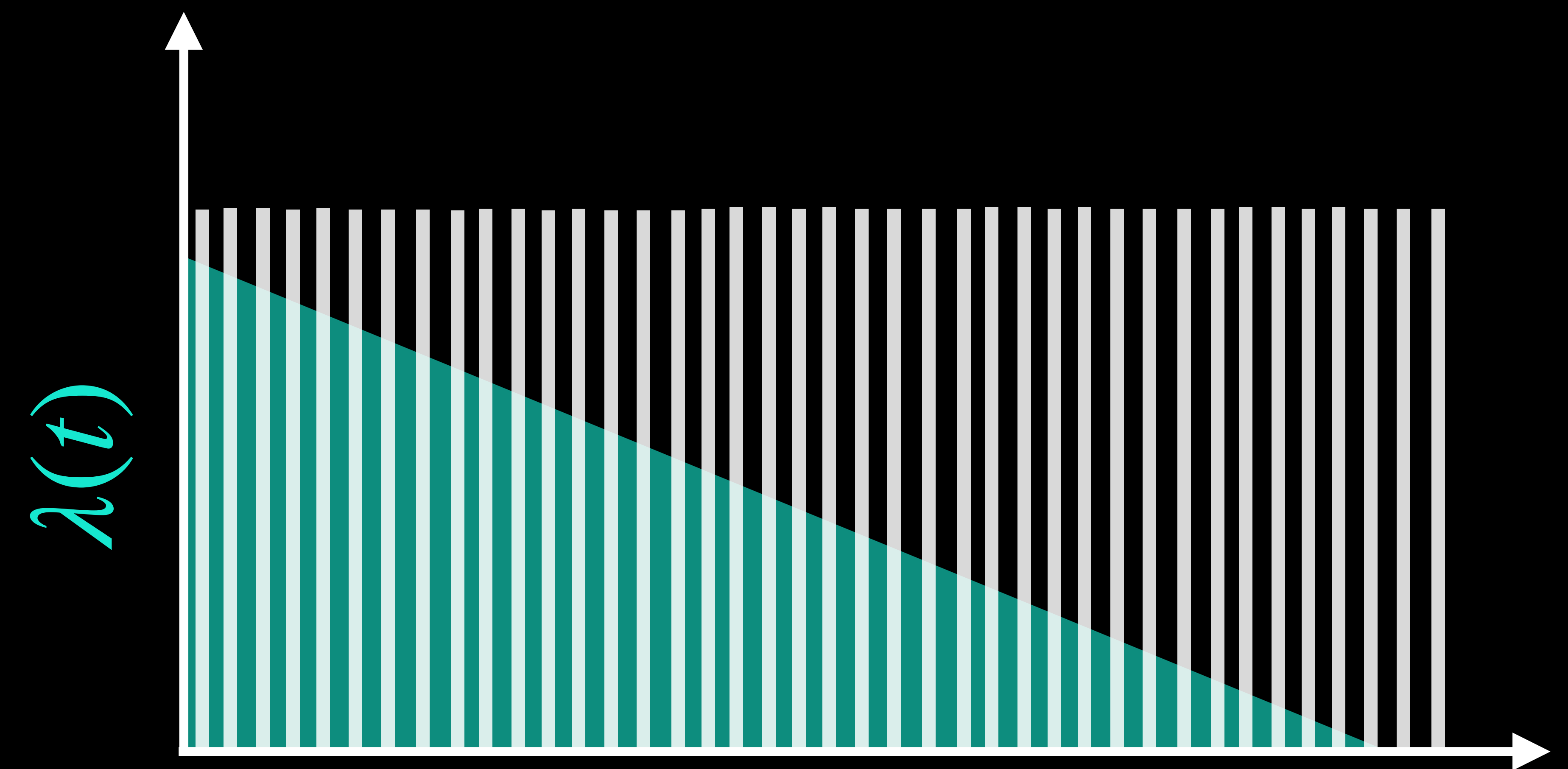
The Probability of the next event happening in Δ at time t $= \lambda(t)\Delta$

P(event) $= \lambda(t)\Delta$

P(no event) $= (1 - \lambda(t)\Delta)$

Naïve Method

Discretise the time axis into very small time intervals Δ and for each Δ calculate



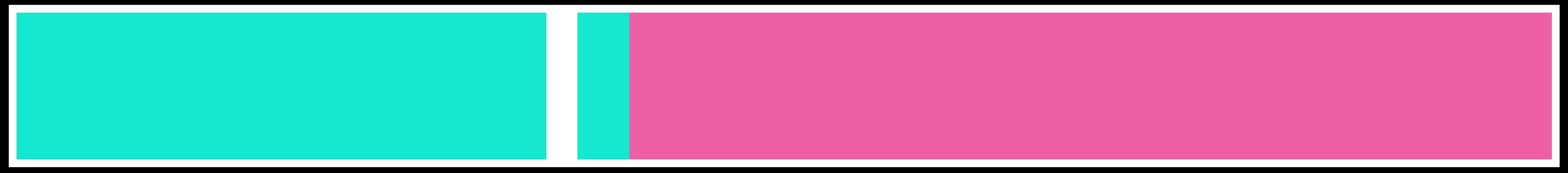
The Probability of the next event happening in Δ at time t $= \lambda(t)\Delta$

Sample from a uniform distribution

$$u \sim U([0,1)]$$

P(event) $= \lambda(t)\Delta$

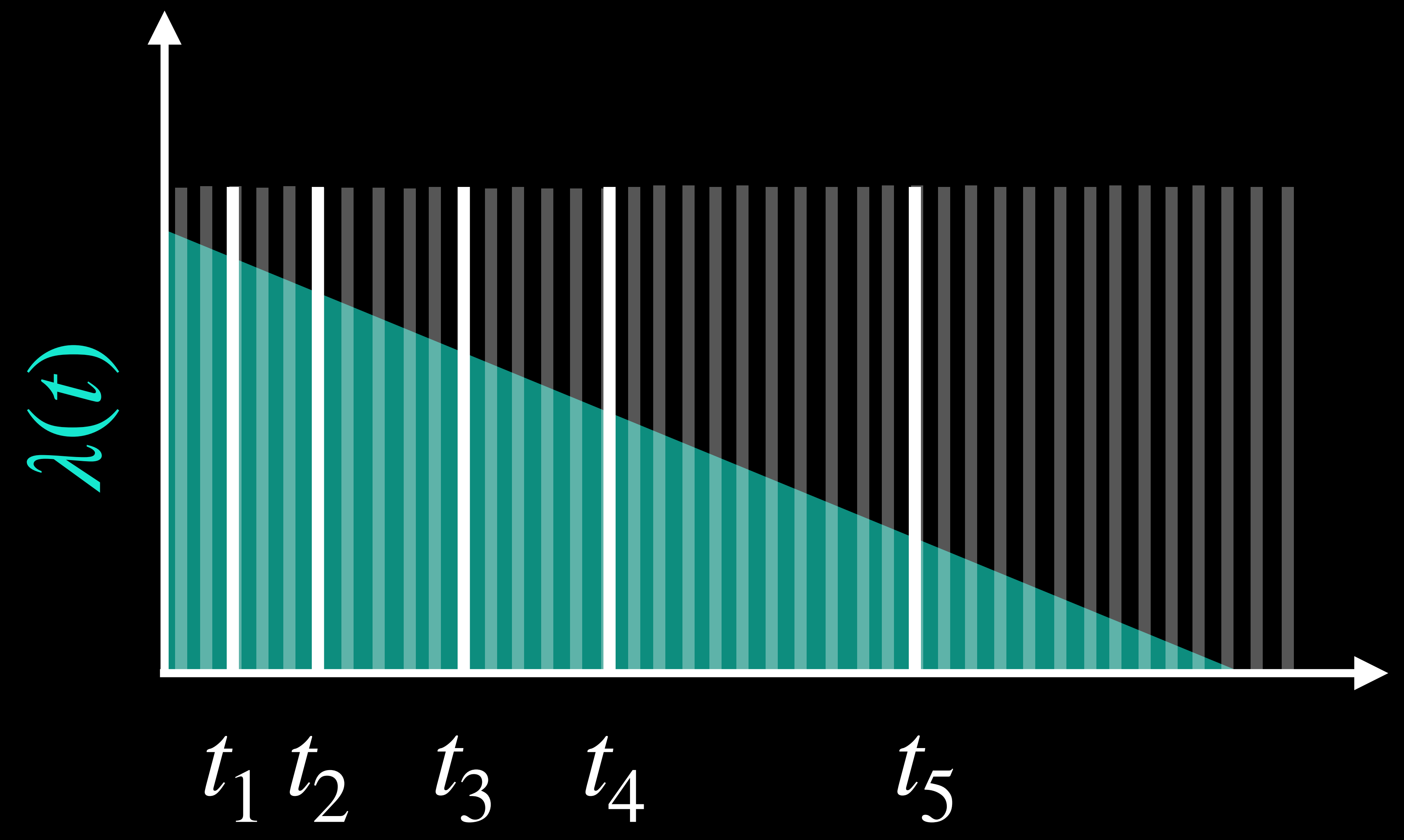
P(no event) $= (1 - \lambda(t)\Delta)$



Naïve Method

Discretise the time axis into very small time intervals Δ and for each Δ calculate

The Probability of the next event happening in Δ at time t $= \lambda(t)\Delta$

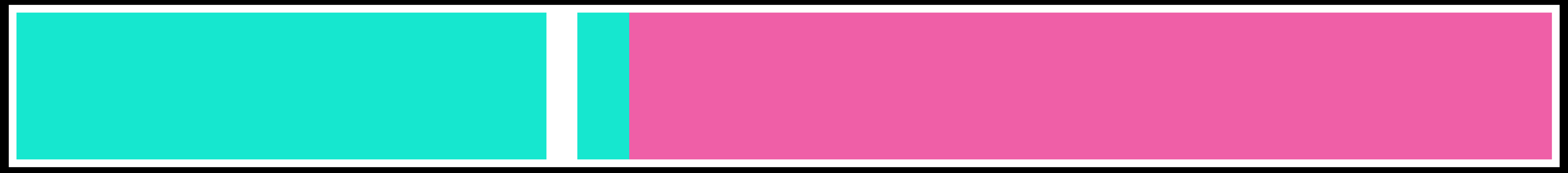


P(event) $= \lambda(t)\Delta$

P(no event) $= (1 - \lambda(t)\Delta)$

Sample from a uniform distribution

$$u \sim U([0,1)]$$



Event happens if : $u \leq \lambda(t)dt$

Naïve Method SCM

History

Past events up to t

Intensity

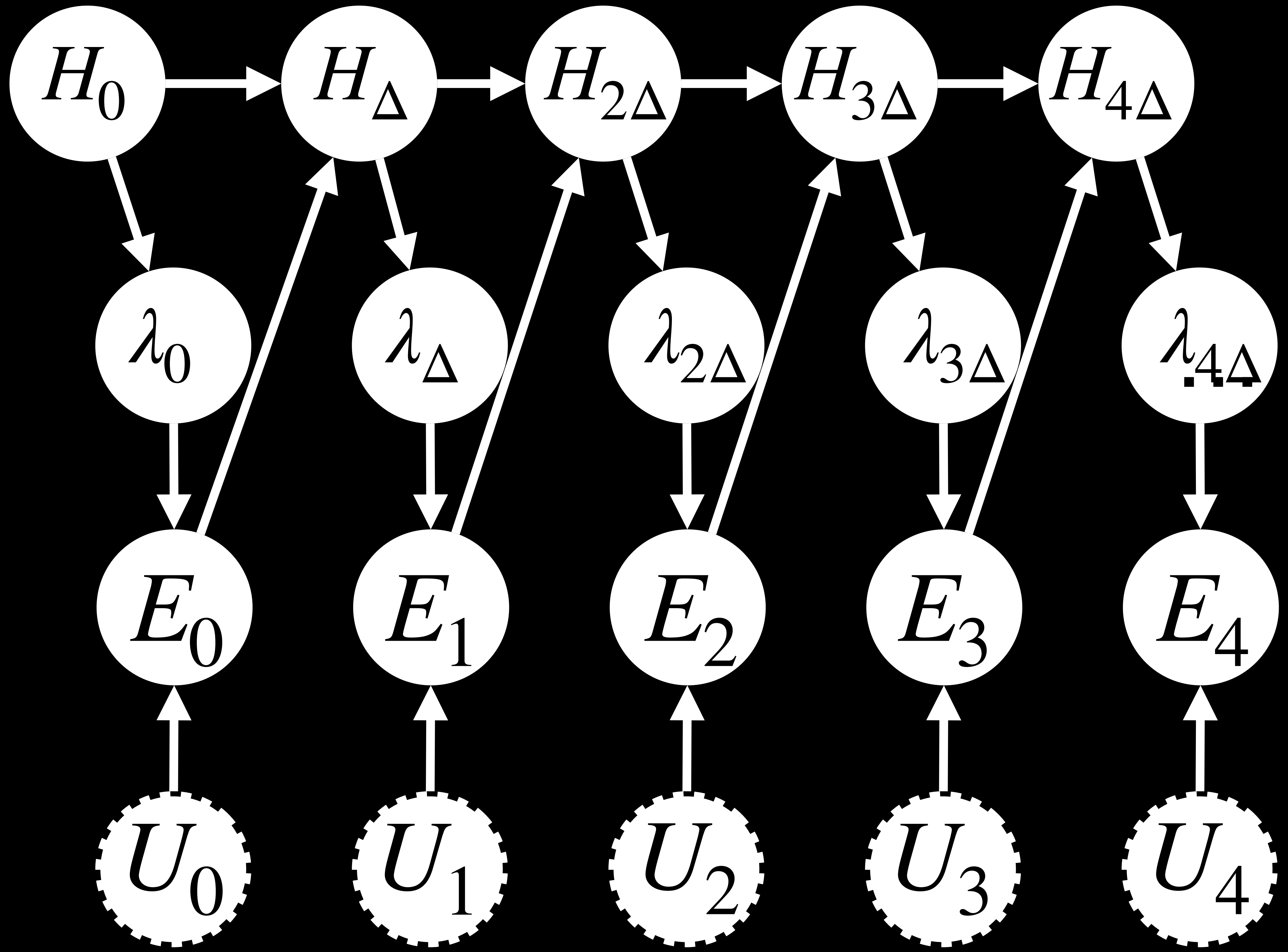
Instantaneous rate at time t

Event

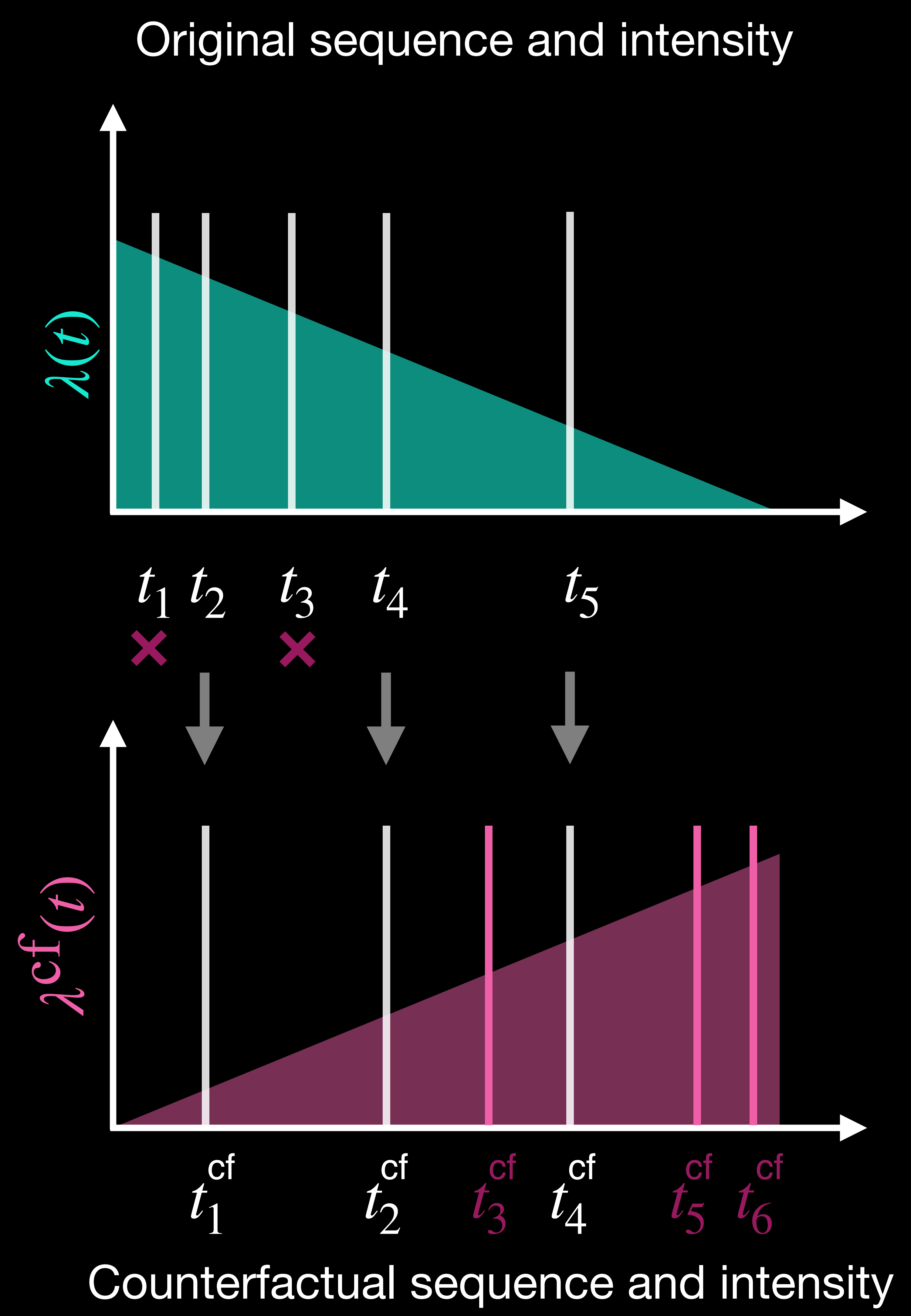
Binary variable indicates if event at t

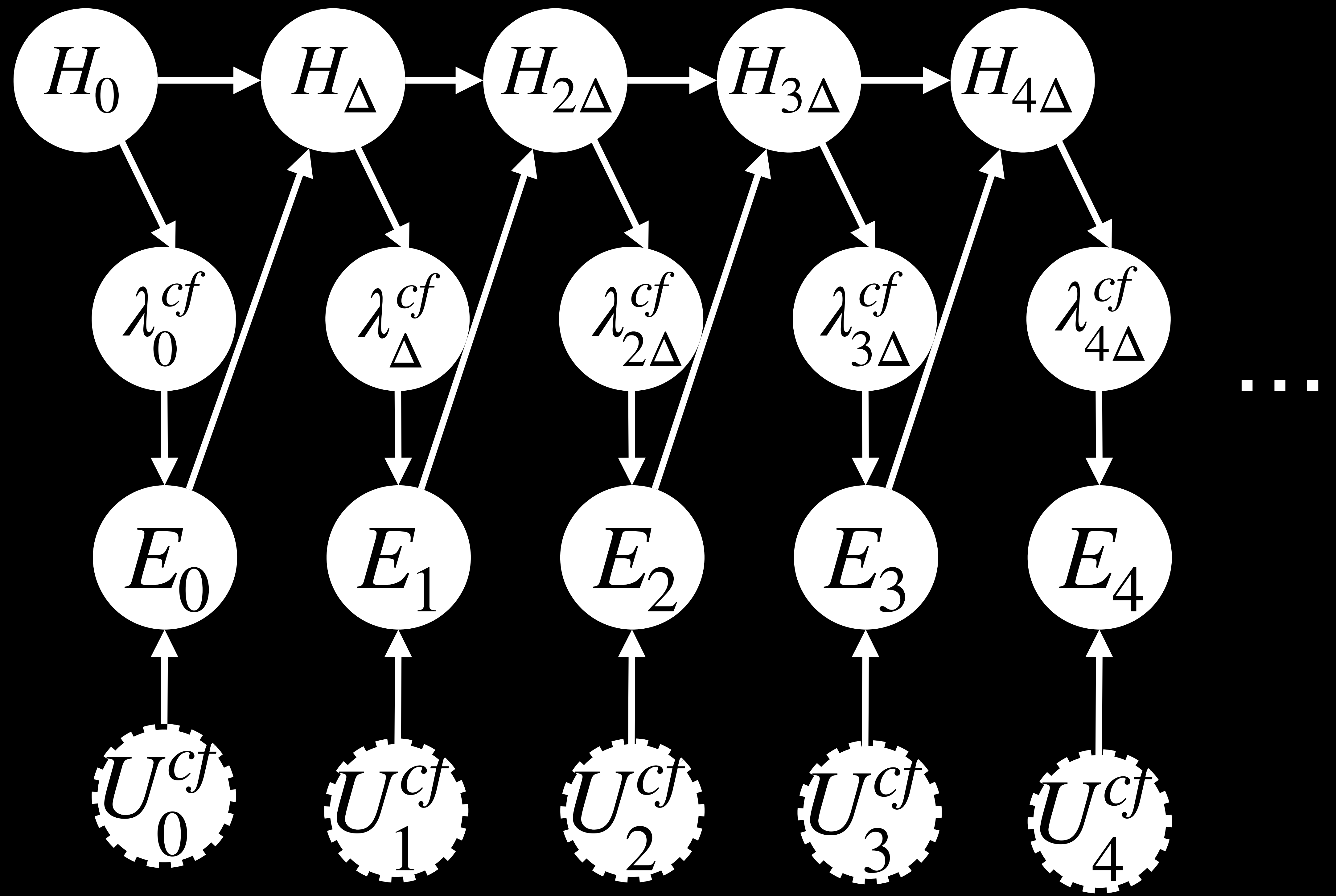
Noise

Uniformly at random in $[0,1]$

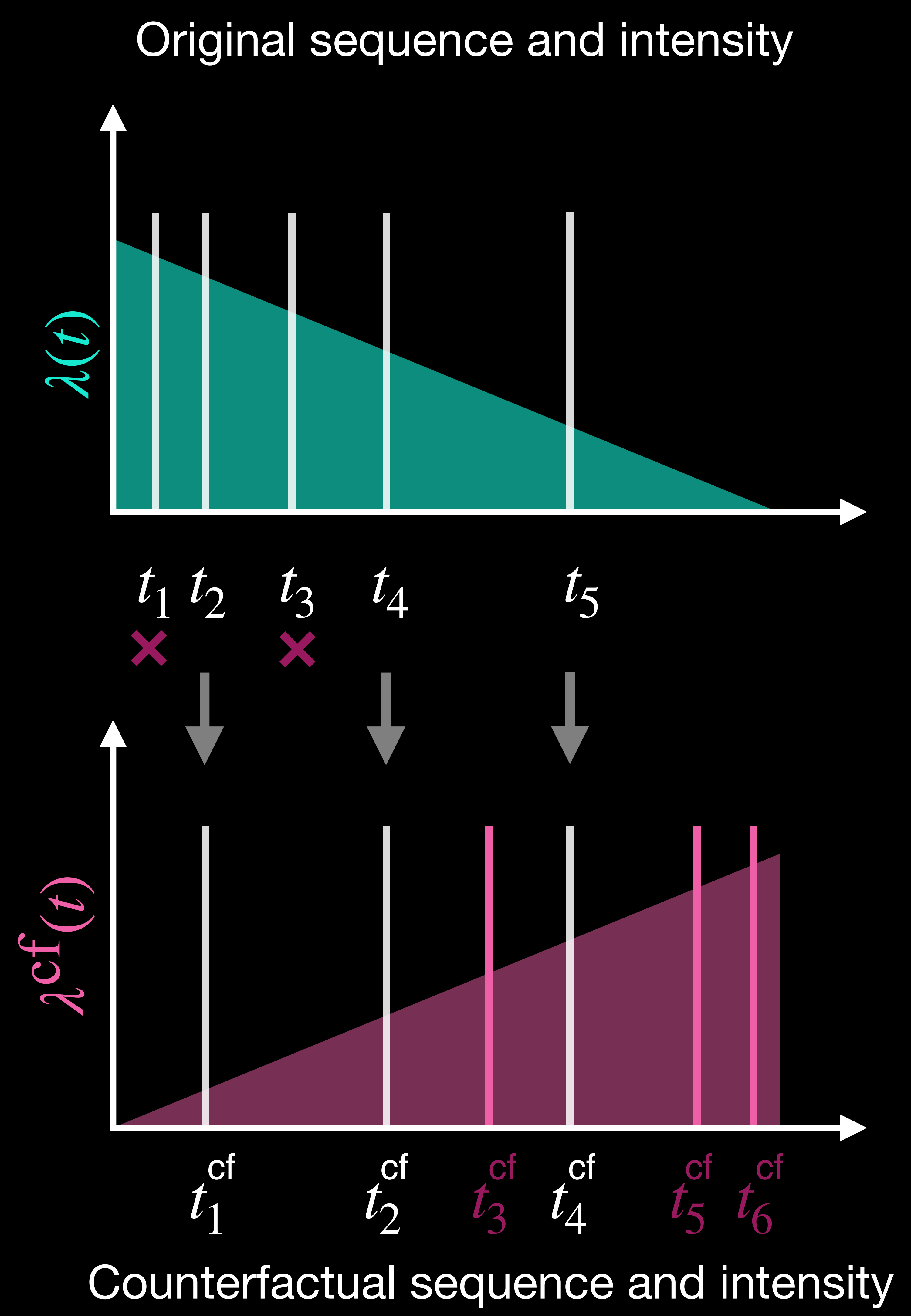


Naïve Method Counterfactual Generation

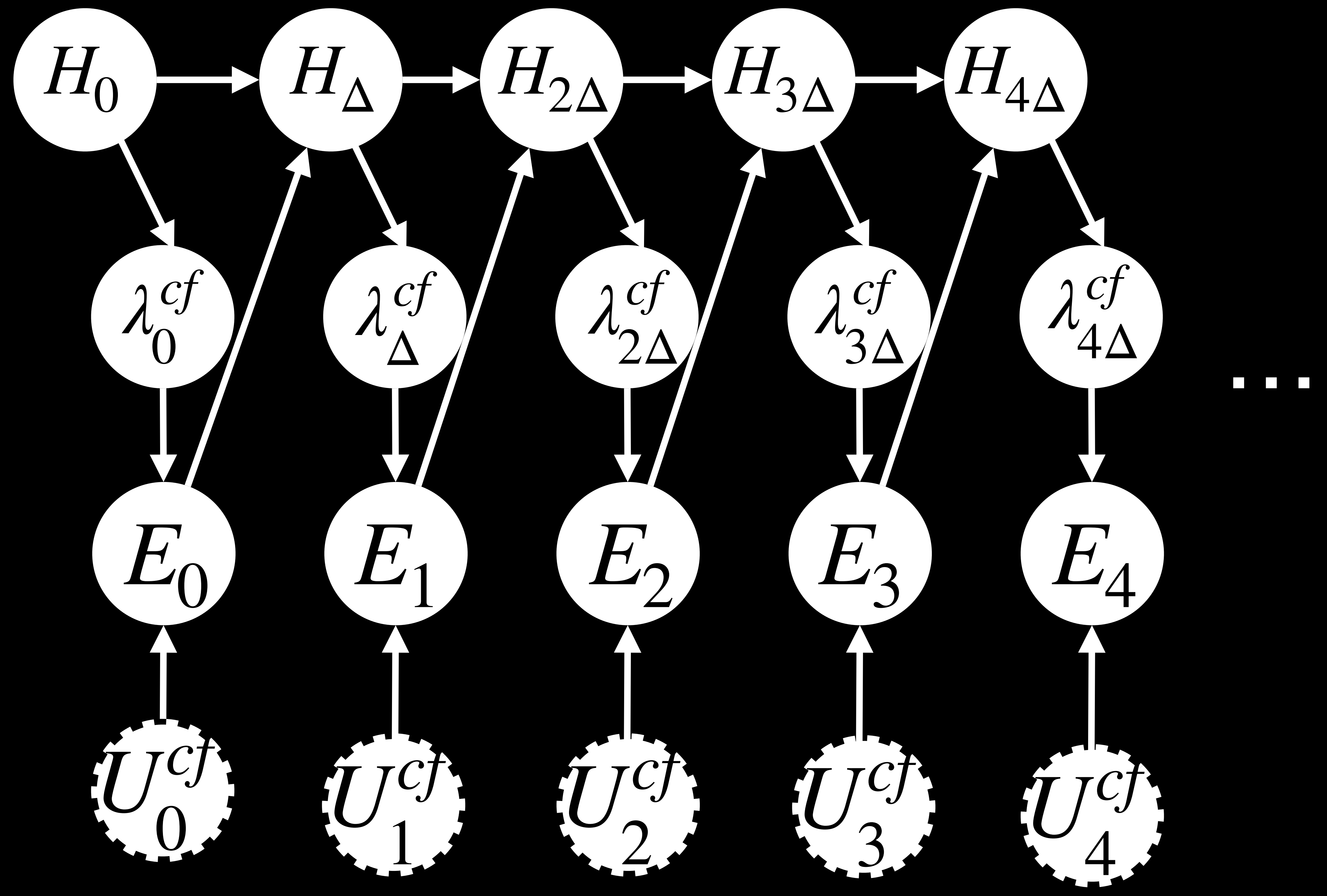




Naïve Method Counterfactual Generation

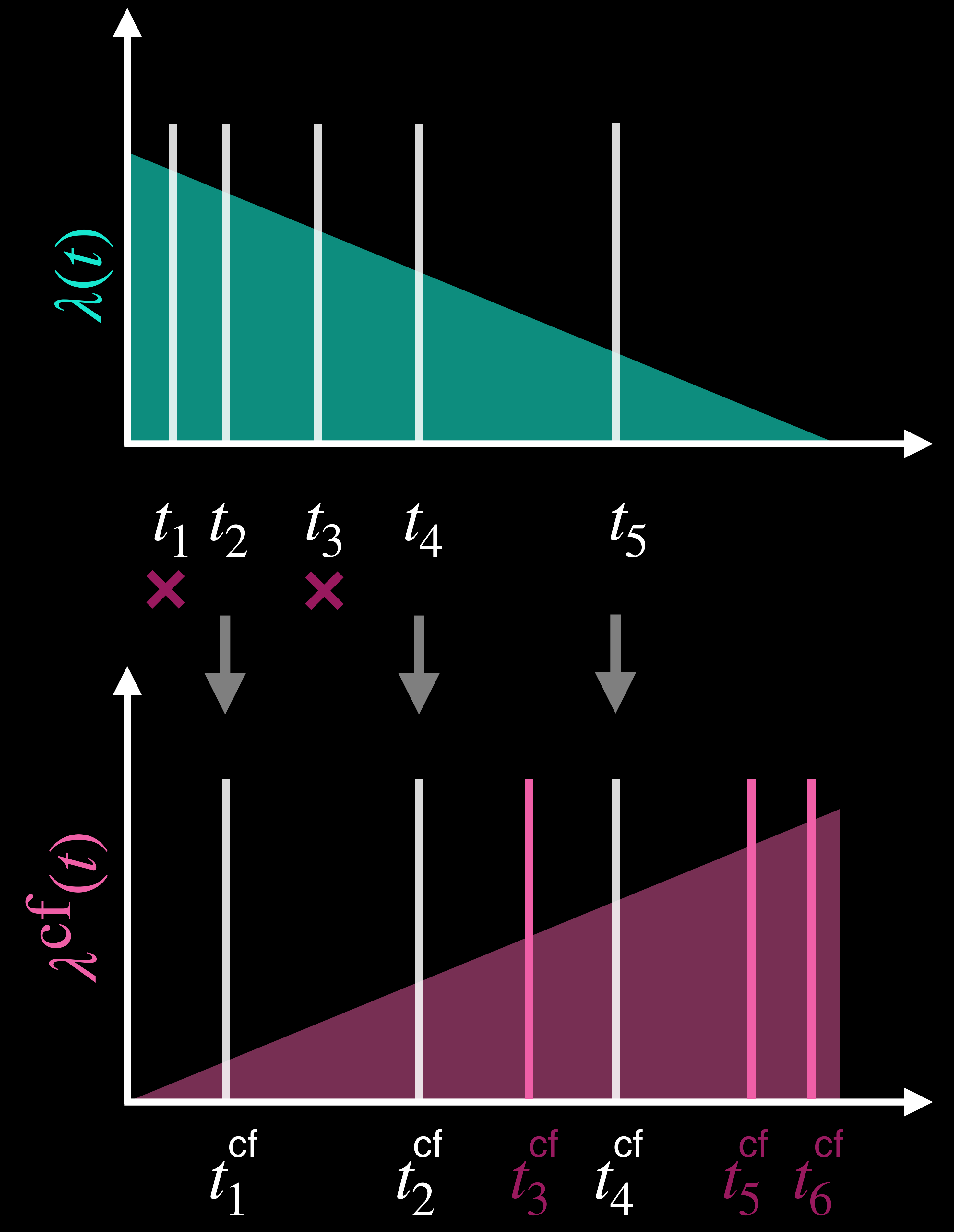


Noise posterior
based on the Observed Outcome



Naïve Method Counterfactual Generation

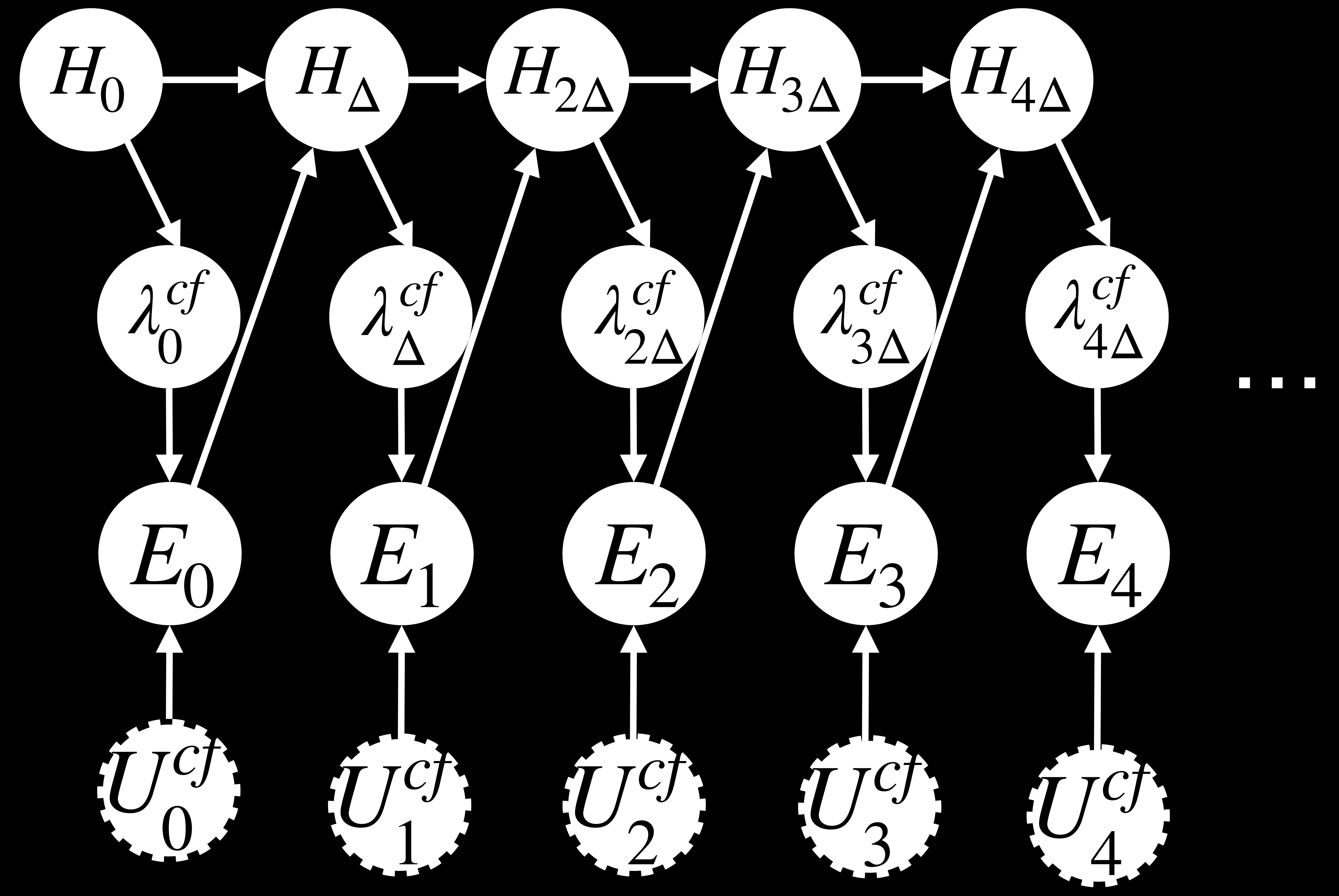
Original sequence and intensity



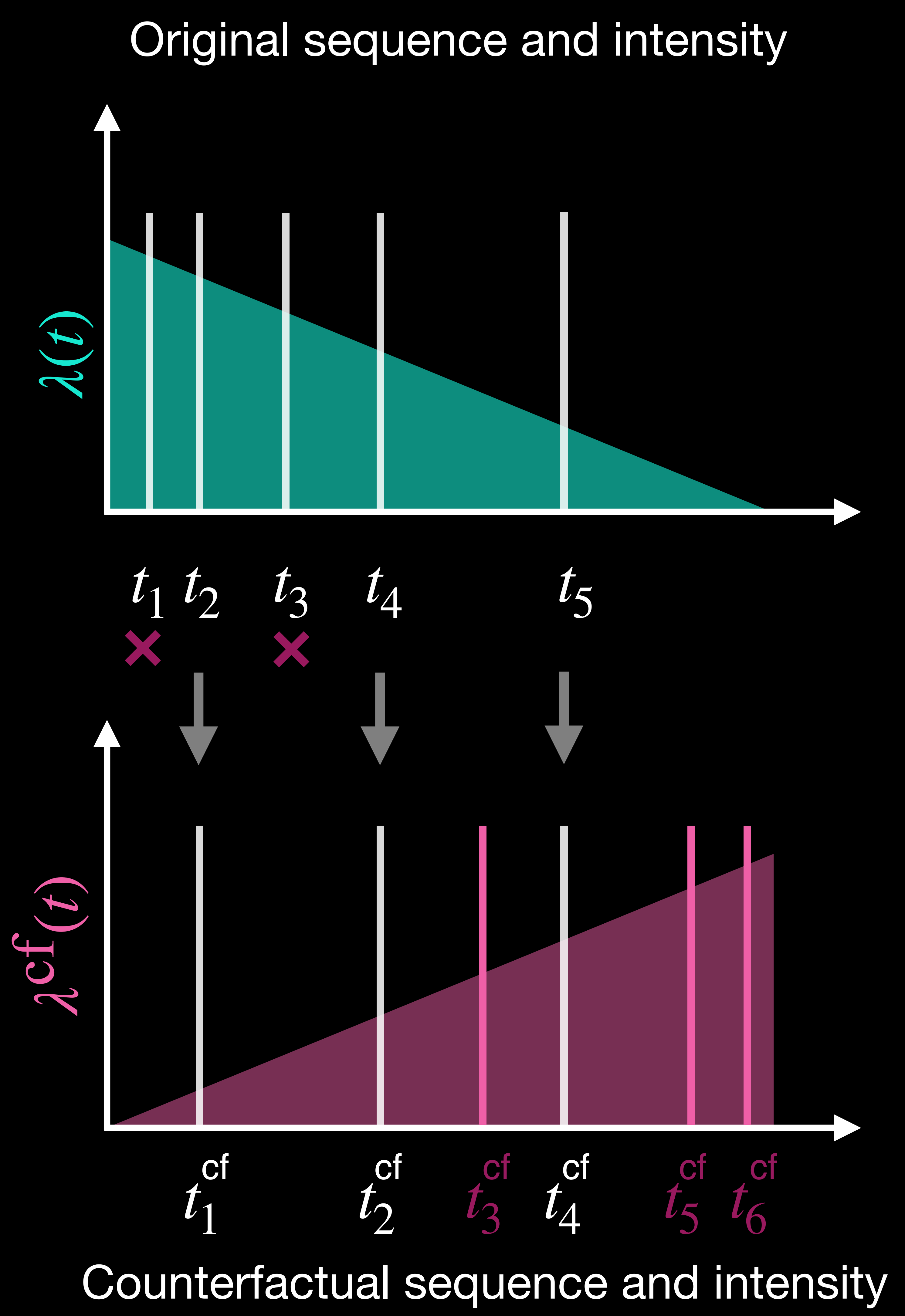
Counterfactual sequence and intensity

Hypothetical Intensity
Instantaneous rate at time t

Noise posterior
based on the Observed Outcome



Naïve Method Counterfactual Generation

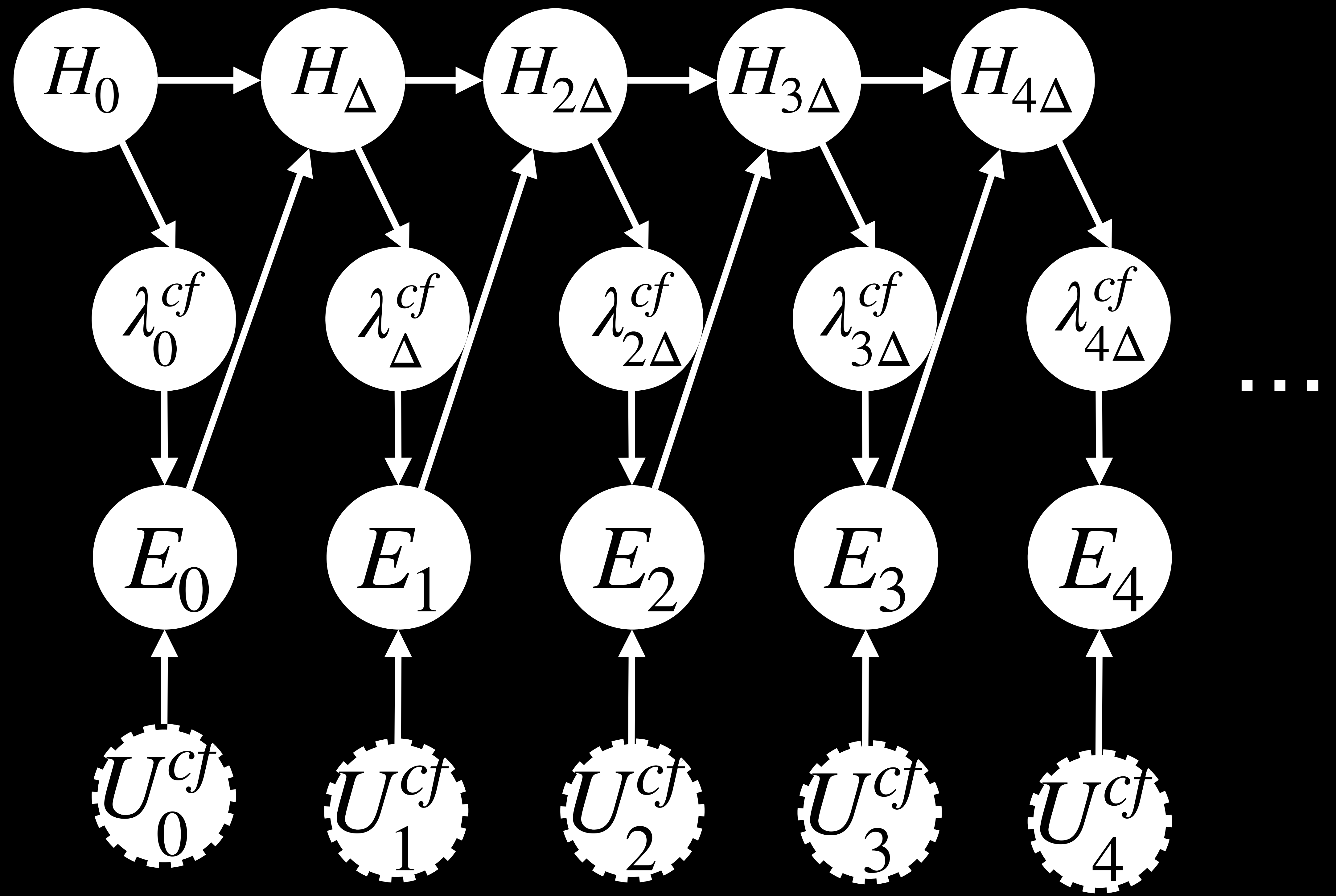


Counterfactual History
Past events up to t

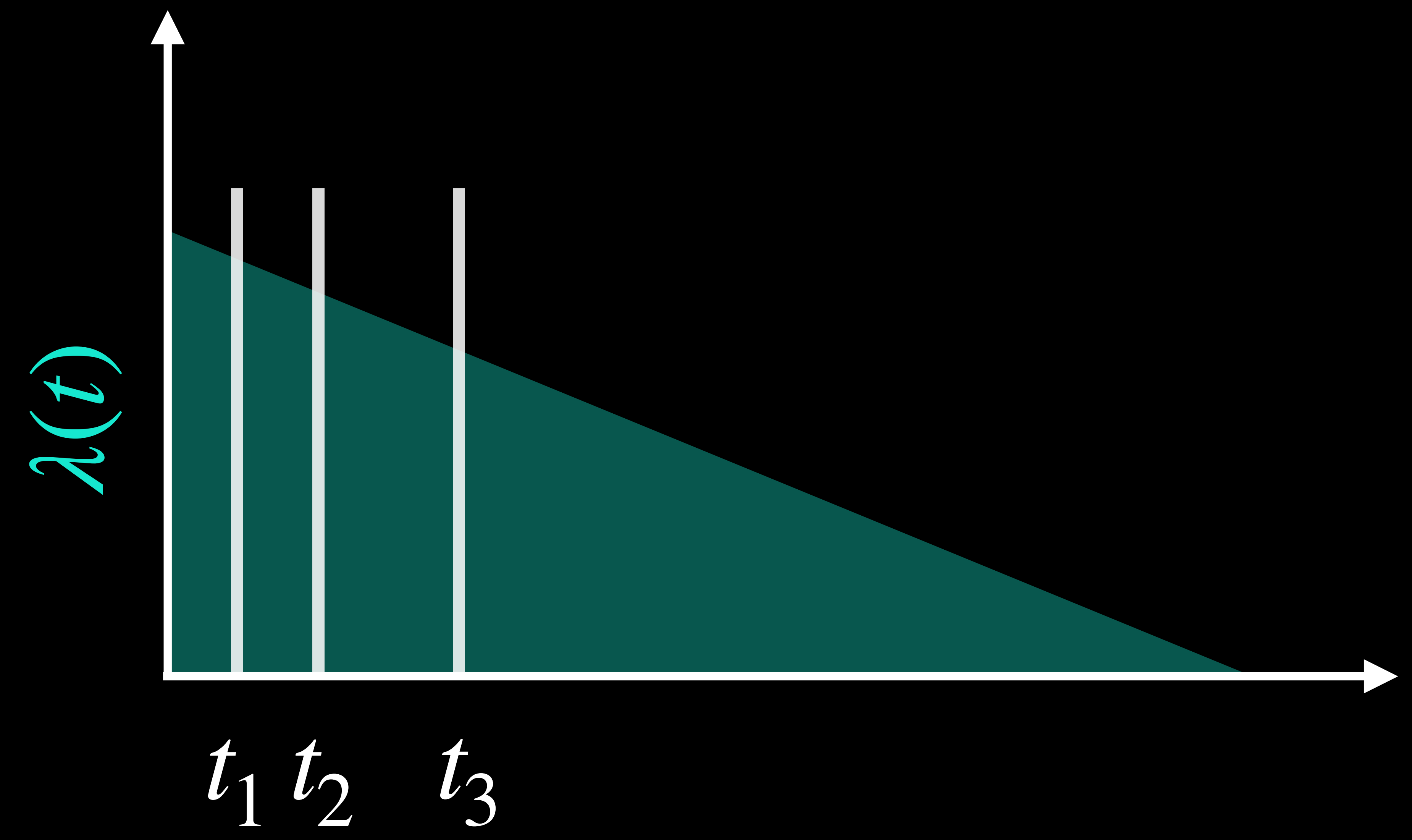
Hypothetical Intensity
Instantaneous rate at time t

Counterfactual Events
Binary variable indicates if event at t

Noise posterior
based on the Observed Outcome



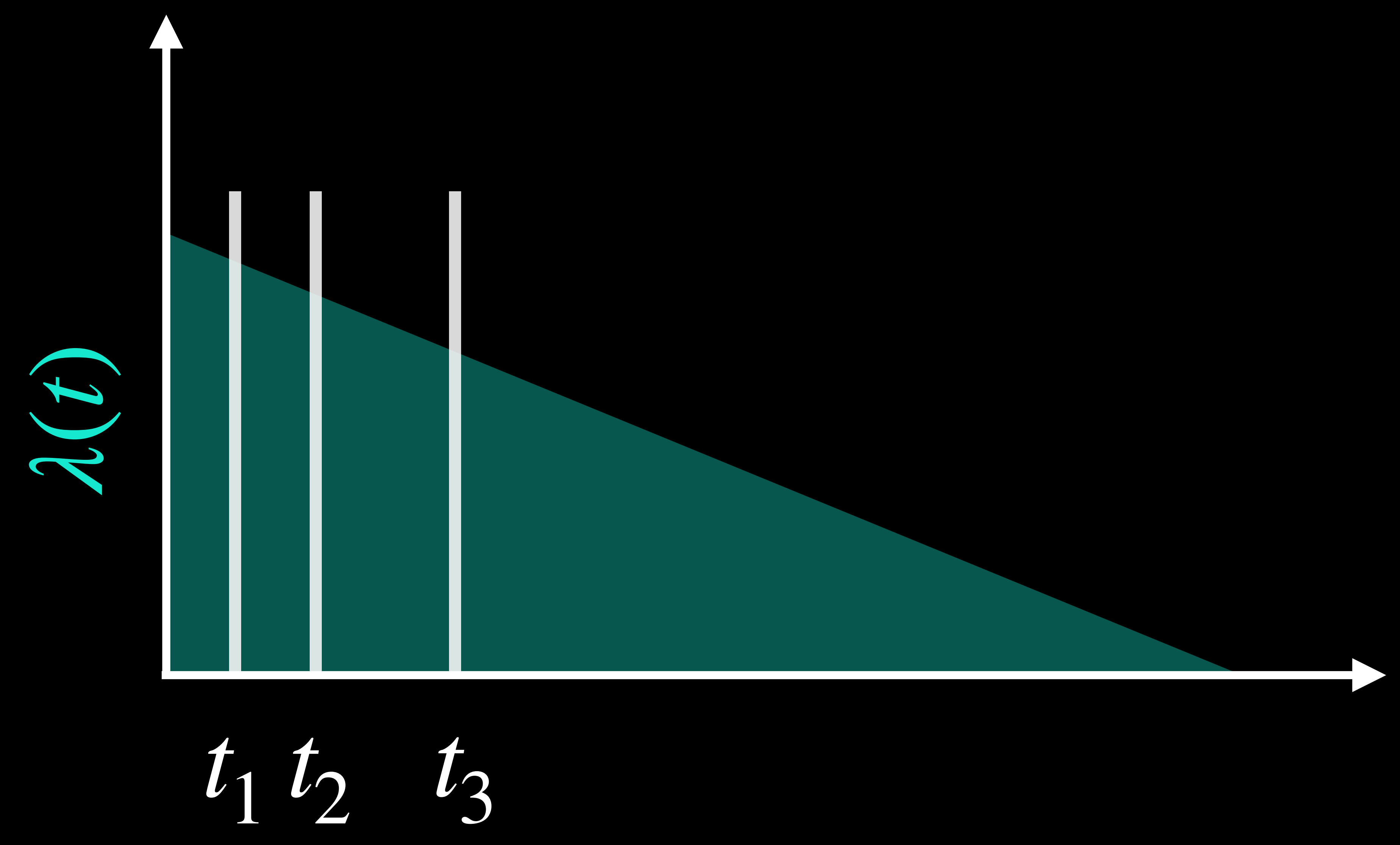
Numerical Integration



Probability of no event happening

$$u \sim U([0,1])$$

Numerical Integration



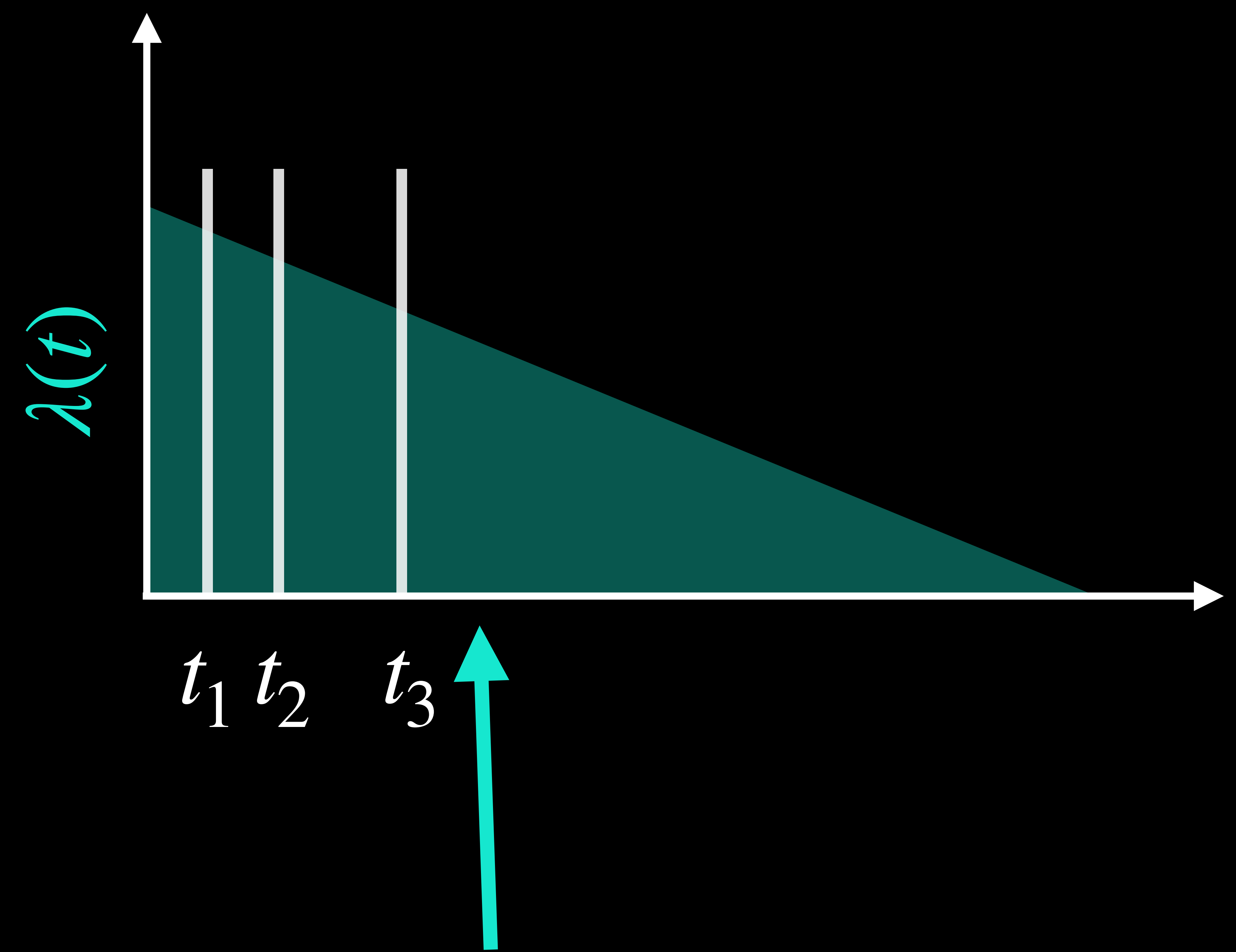
Probability of no event happening

$$u \sim U([0,1])$$

Numerically Integrate till probability of event happening is greater than u.

$$u \geq e^{-\int_{t'}^{t'+x} \lambda(t)dt}$$

Numerical Integration



$$-\ln(u) \leq \int_{t'}^{t'+x} \lambda(t)dt$$

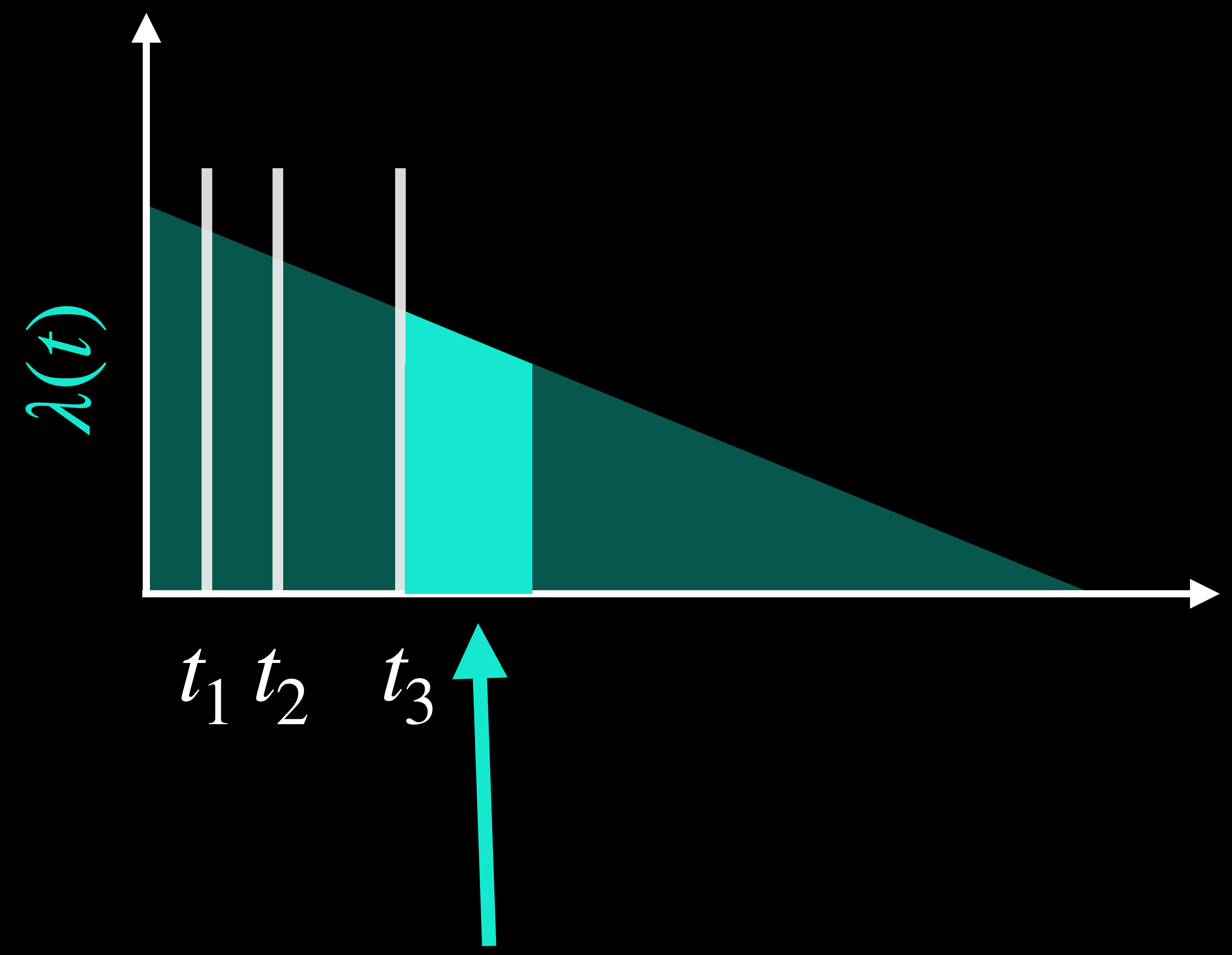
Probability of no event happening

$$u \sim U([0,1])$$

Numerically Integrate till probability of event happening is greater than u.

$$u \geq e^{-\int_{t'}^{t'+x} \lambda(t)dt}$$

Numerical Integration



$$-\ln(u) \leq \int_{t'}^{t'+x} \lambda(t)dt$$

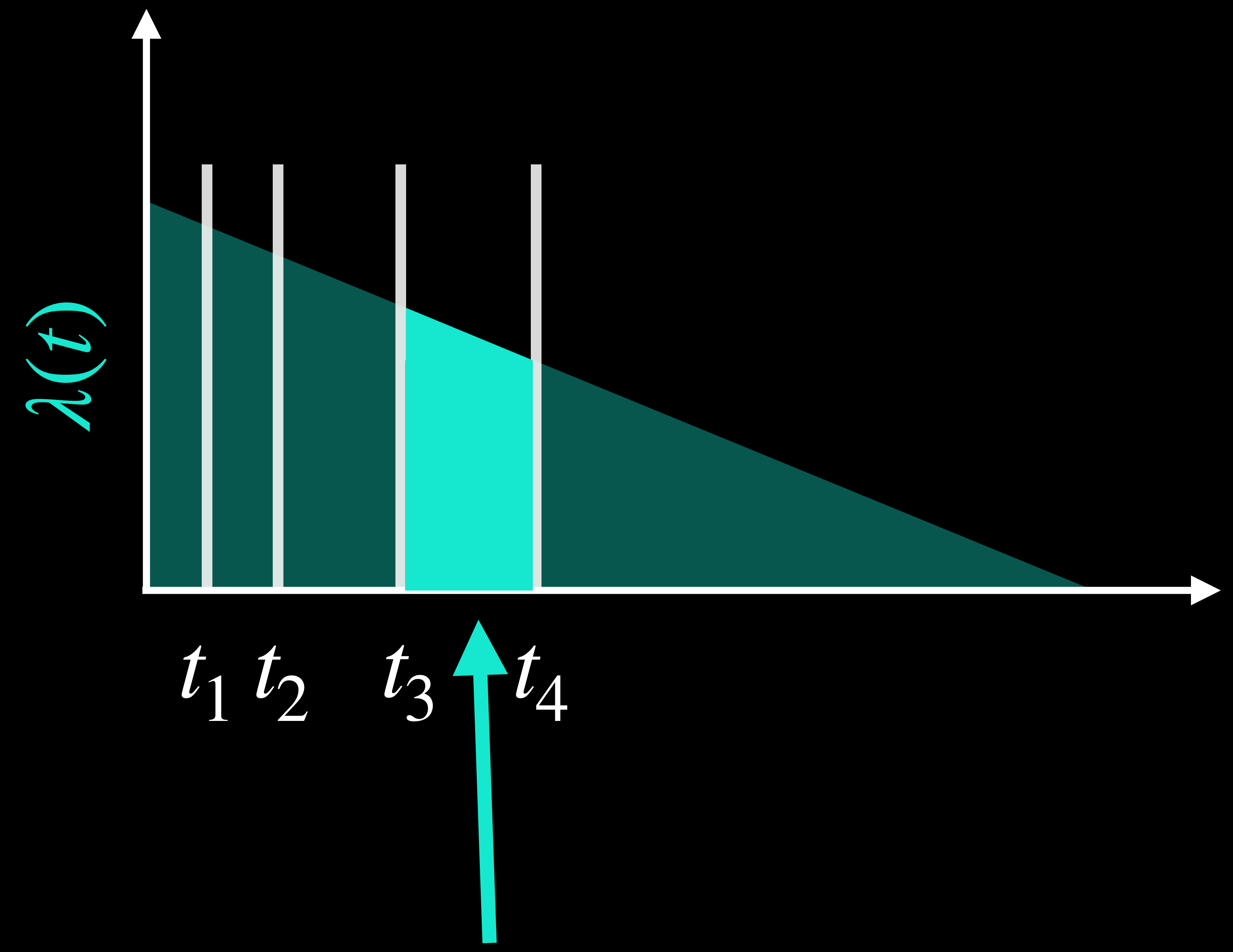
Probability of no event happening

$$u \sim U([0,1])$$

Numerically Integrate till probability of event happening is greater than u.

$$u \geq e^{-\int_{t'}^{t'+x} \lambda(t)dt}$$

Numerical Integration



$$-\ln(u) \leq \int_{t'}^{t'+x} \lambda(t)dt$$

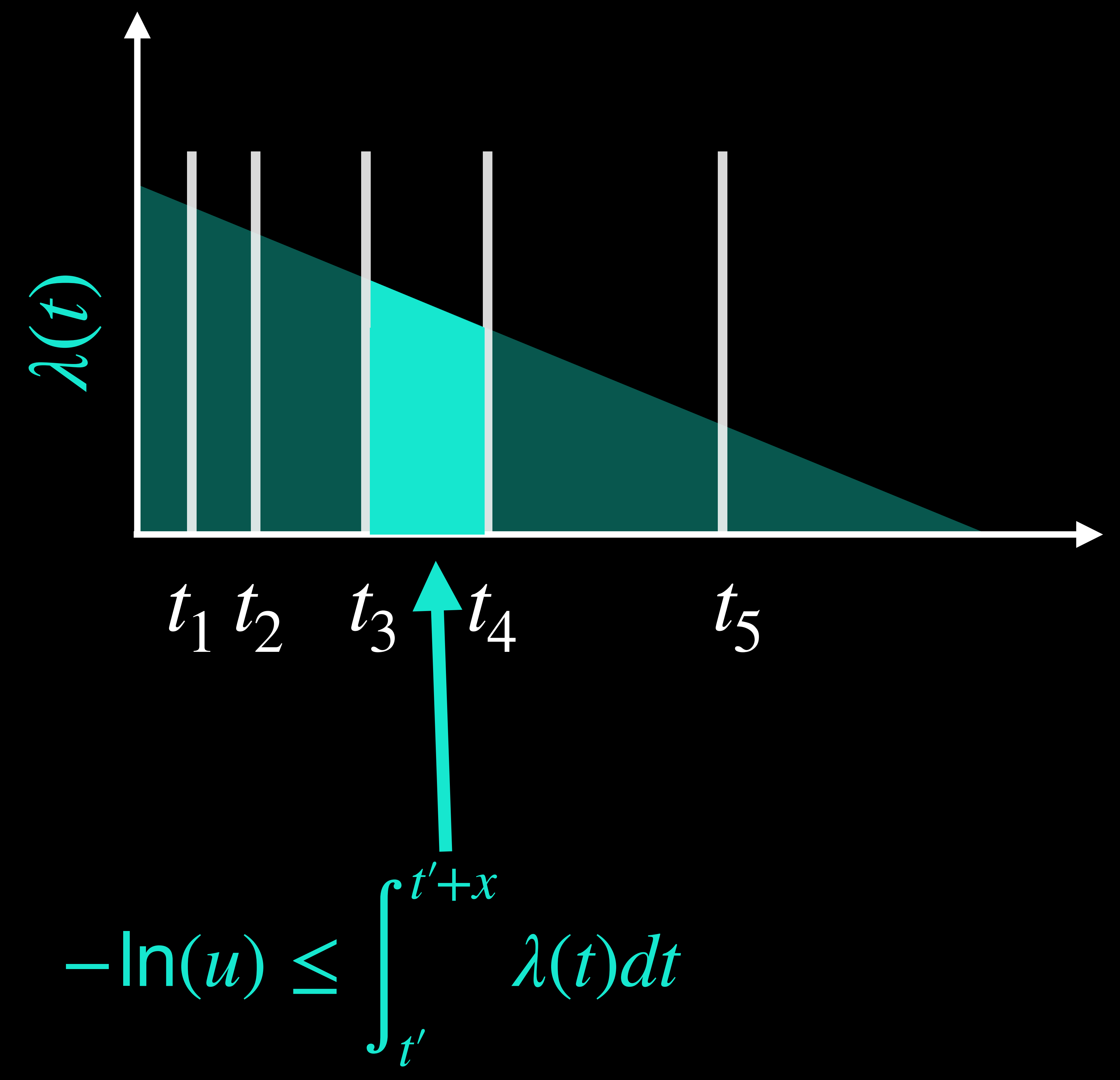
Probability of no event happening

$$u \sim U([0,1])$$

Numerically Integrate till probability of event happening is greater than u.

$$u \geq e^{-\int_{t'}^{t'+x} \lambda(t)dt}$$

Numerical Integration



Probability of no event happening

$$u \sim U([0,1])$$

Numerically Integrate till probability of event happening is greater than u.

$$u \geq e^{-\int_{t'}^{t'+x} \lambda(t)dt}$$

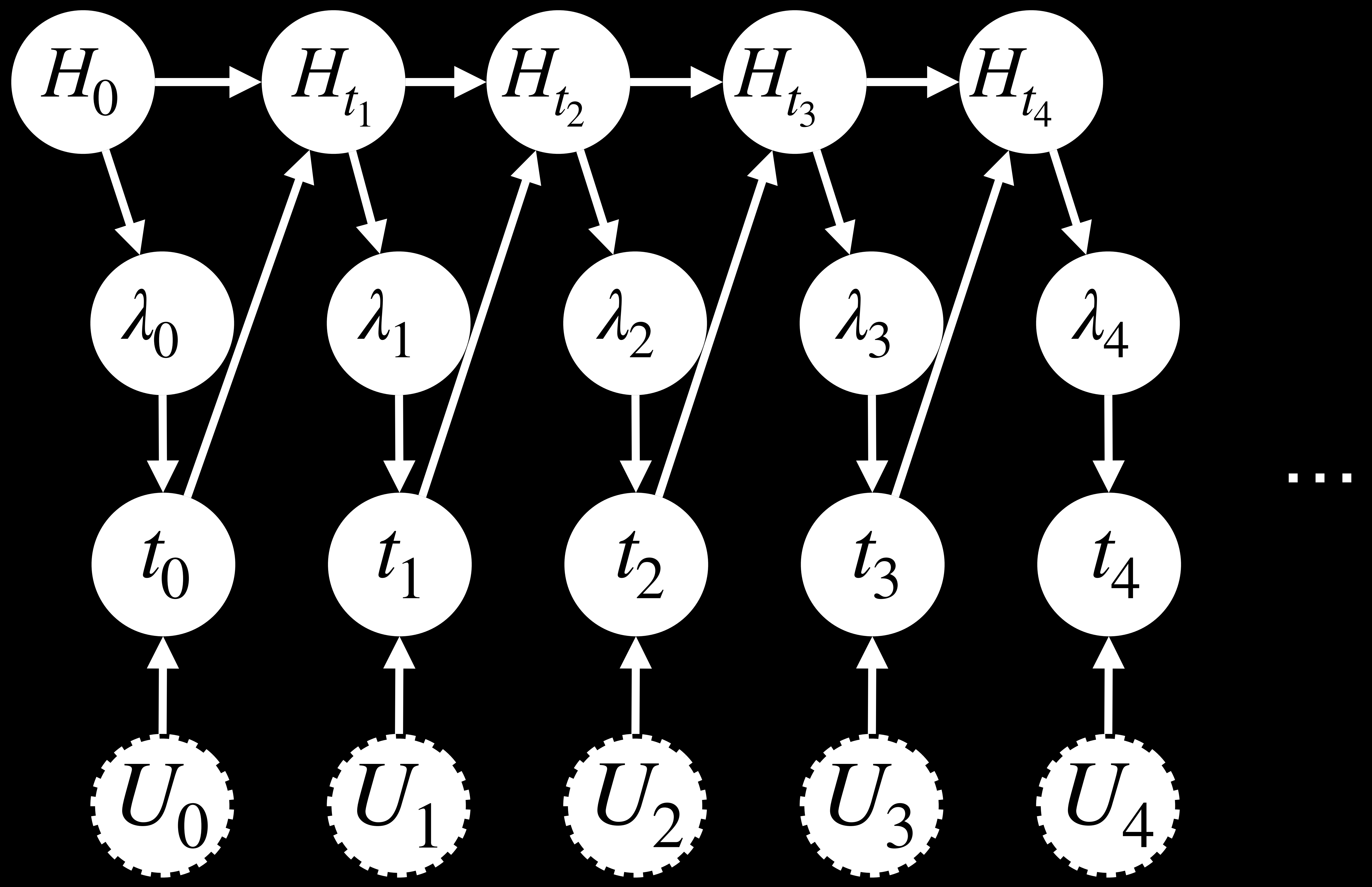
Numerical Integration SCM

History
Past events up event i

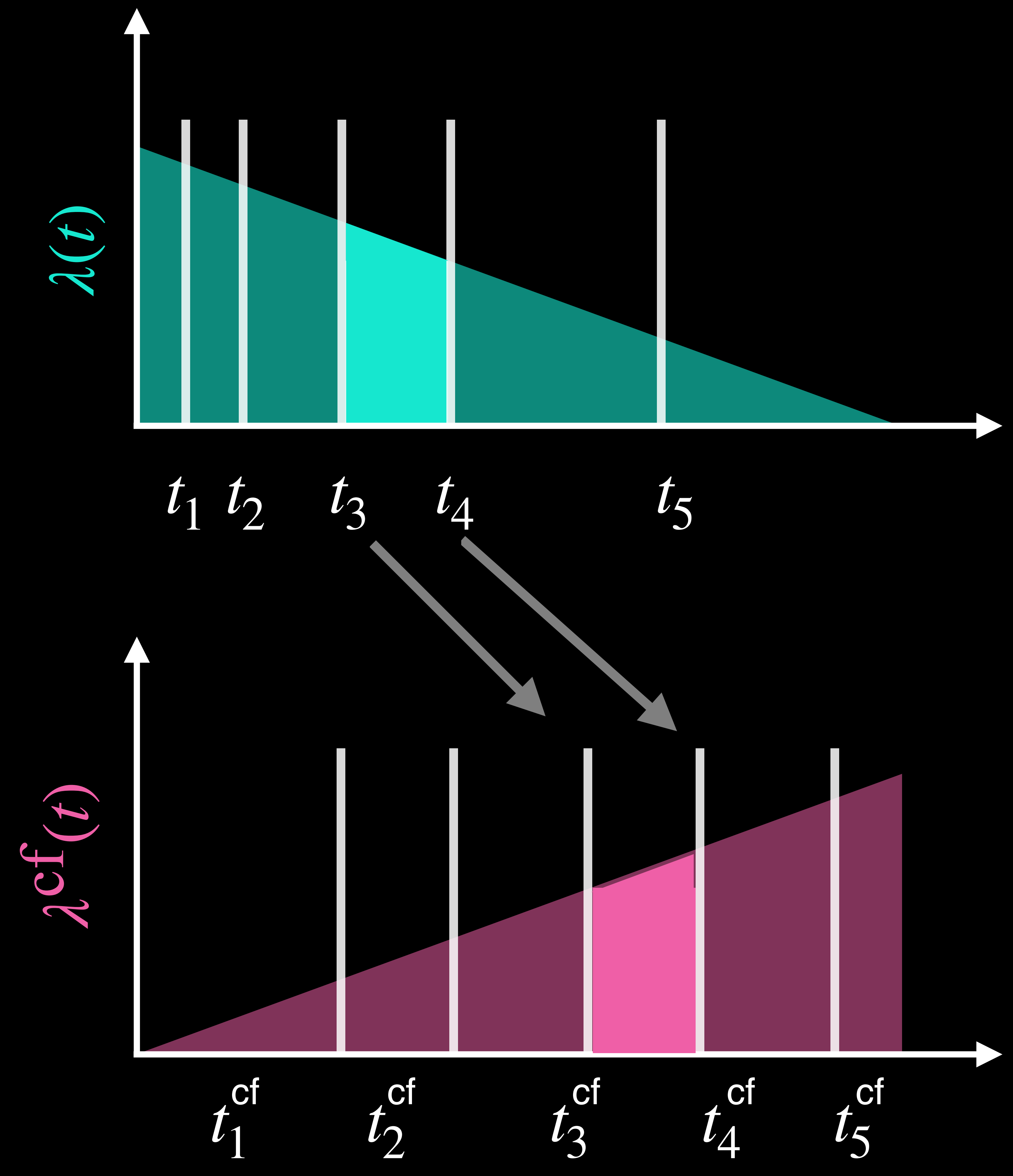
Intensity
Instantaneous rate after event i

Event Time
Event time of event i

Noise
Uniformly at random in $[0,1]$



Numerical Integration Counterfactual Generation

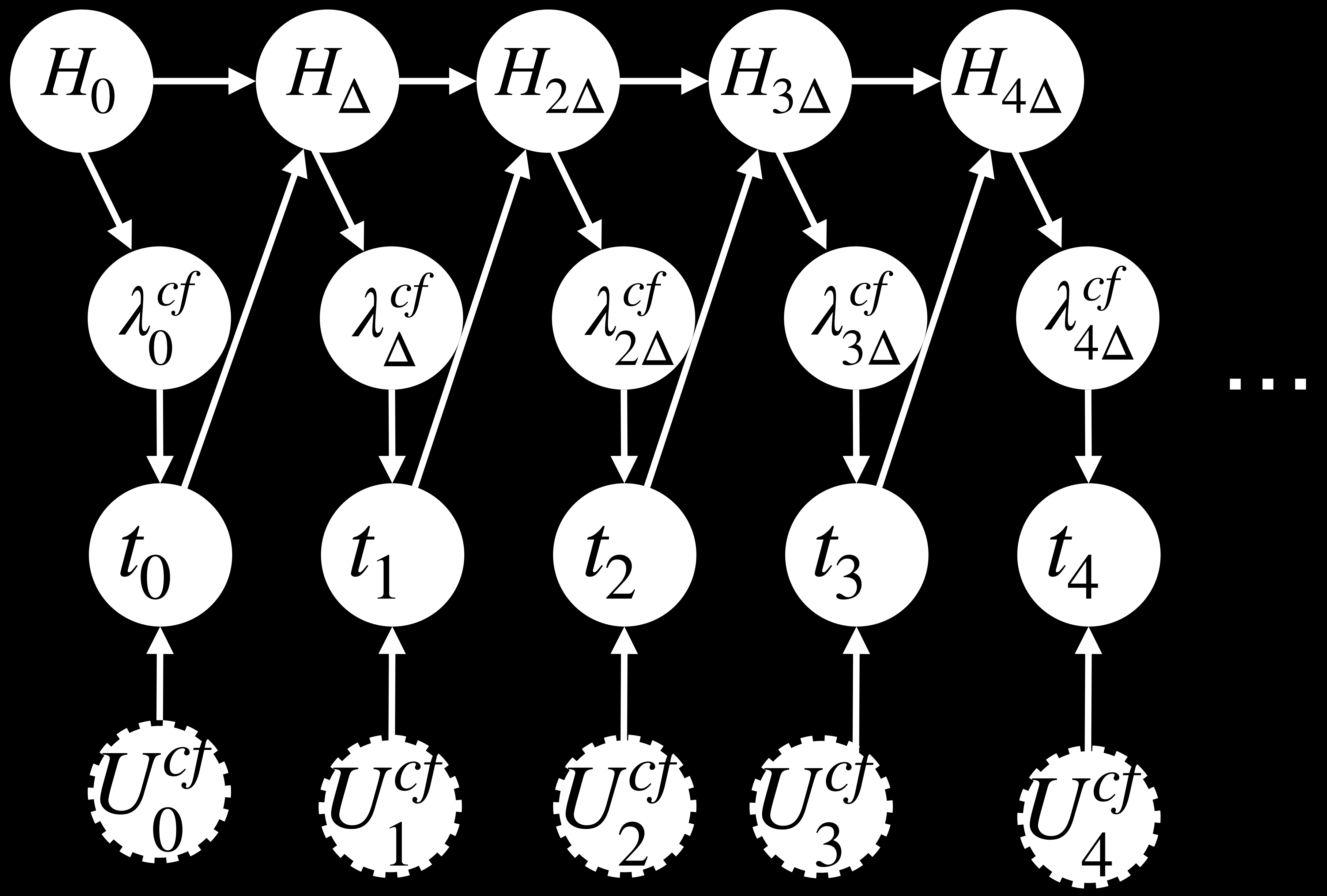


Counterfactual History
Past events up to event i

Counterfactual Intensity
Instantaneous rate at time i

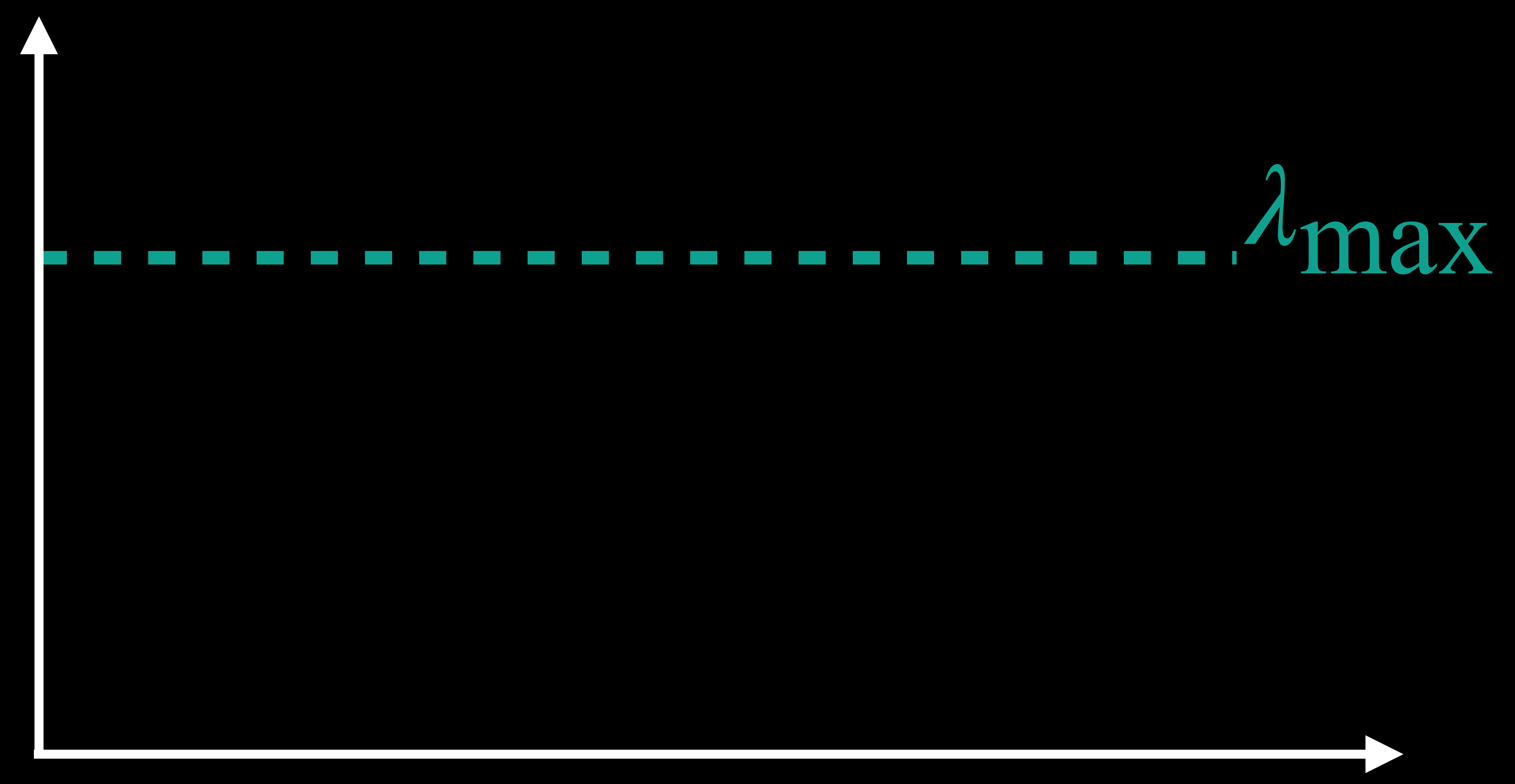
Counterfactual Event times
Event time of event i

Noise Conditioned on observed factual sequence

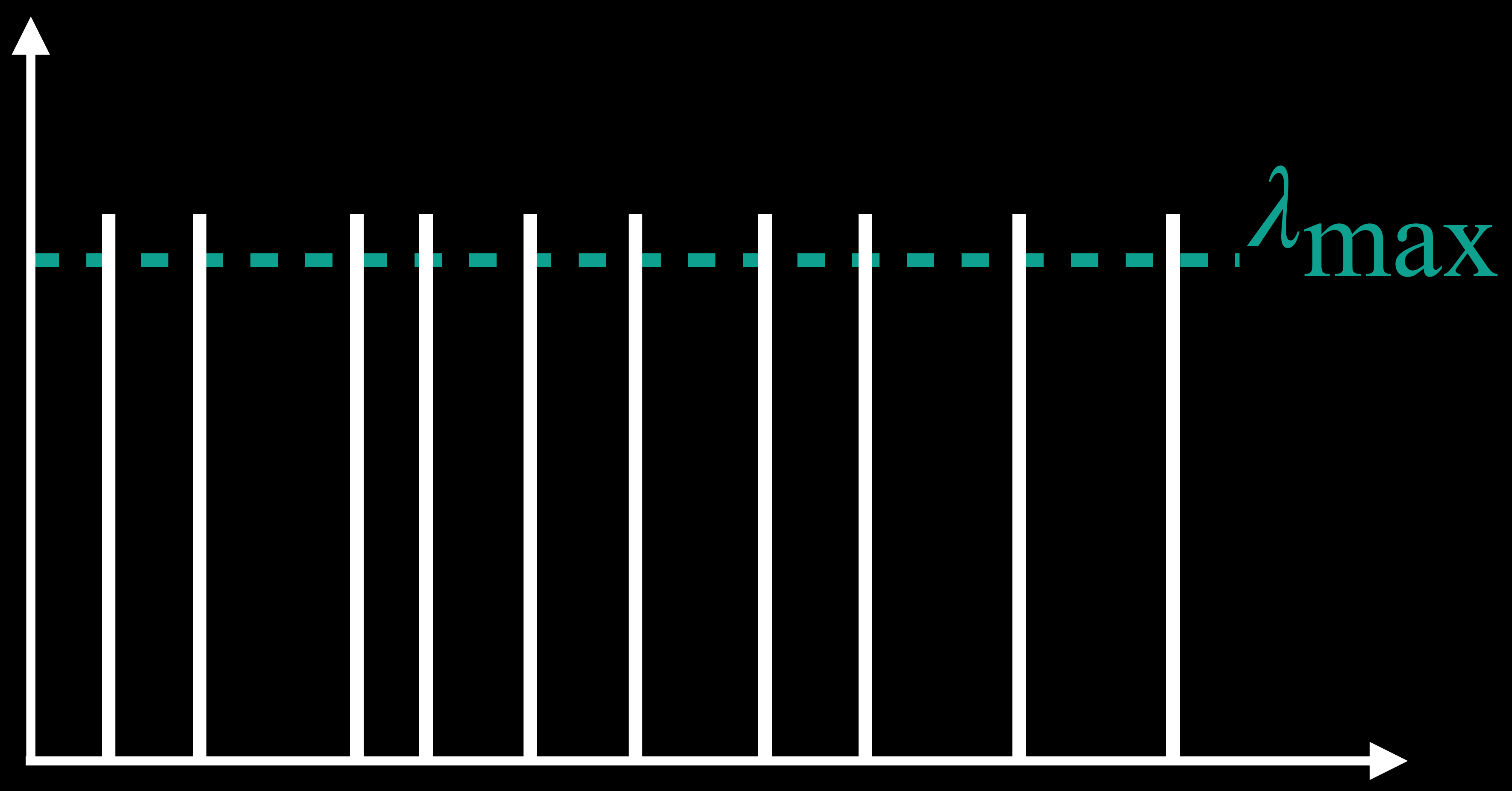


$$\int_{t_3}^{t_4} \lambda(t)dt = -\ln(u_3) = \int_{t_3^{cf}}^{t_4^{cf}} \lambda^{cf}(t)dt$$

Thinning Method



Thinning Method

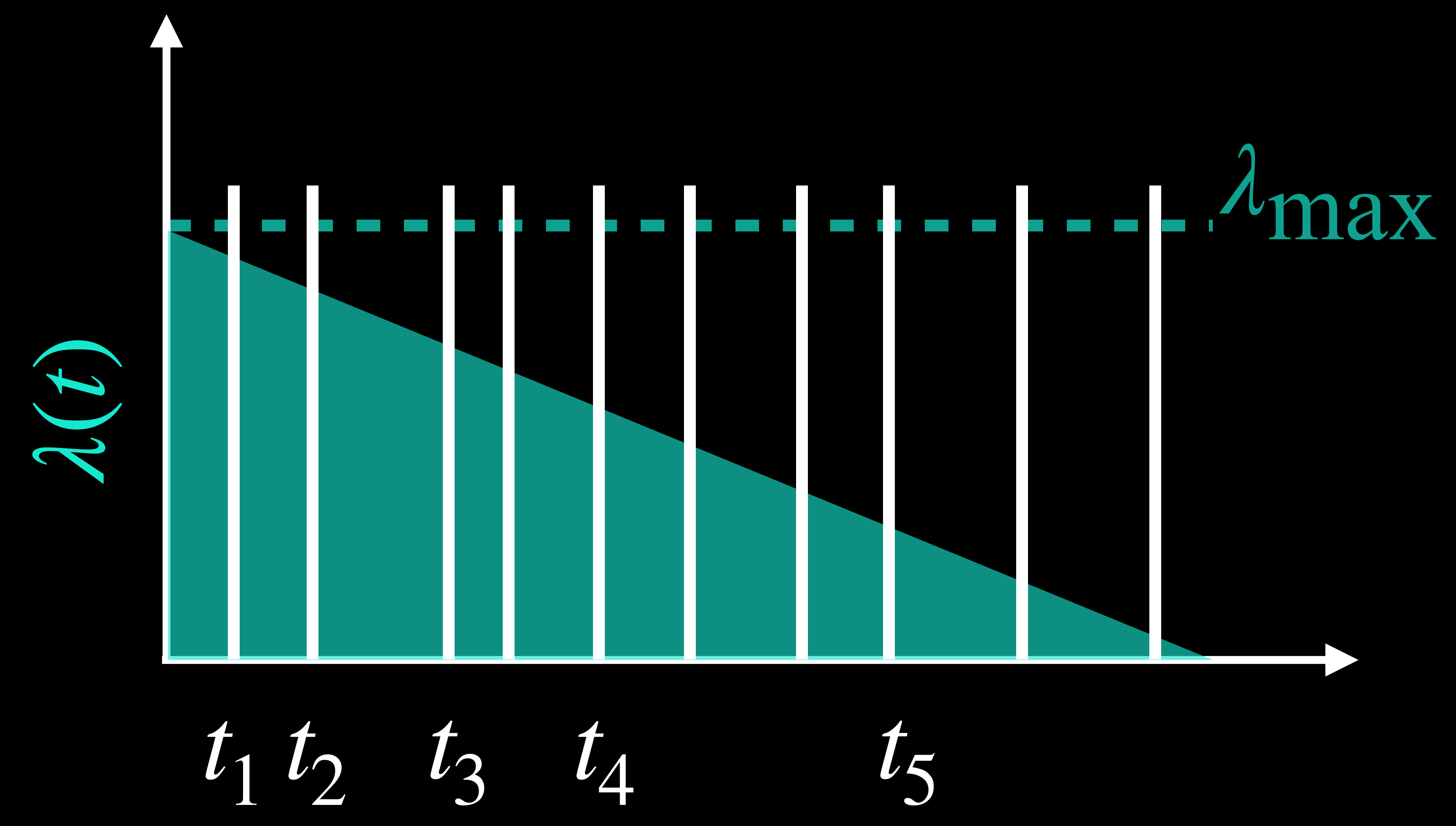


Generate candidate events

$$u \sim U[(0,1)]$$

$$t^c = \frac{-ln(u)}{\lambda_{max}}$$

Thinning Method



Generate candidate events

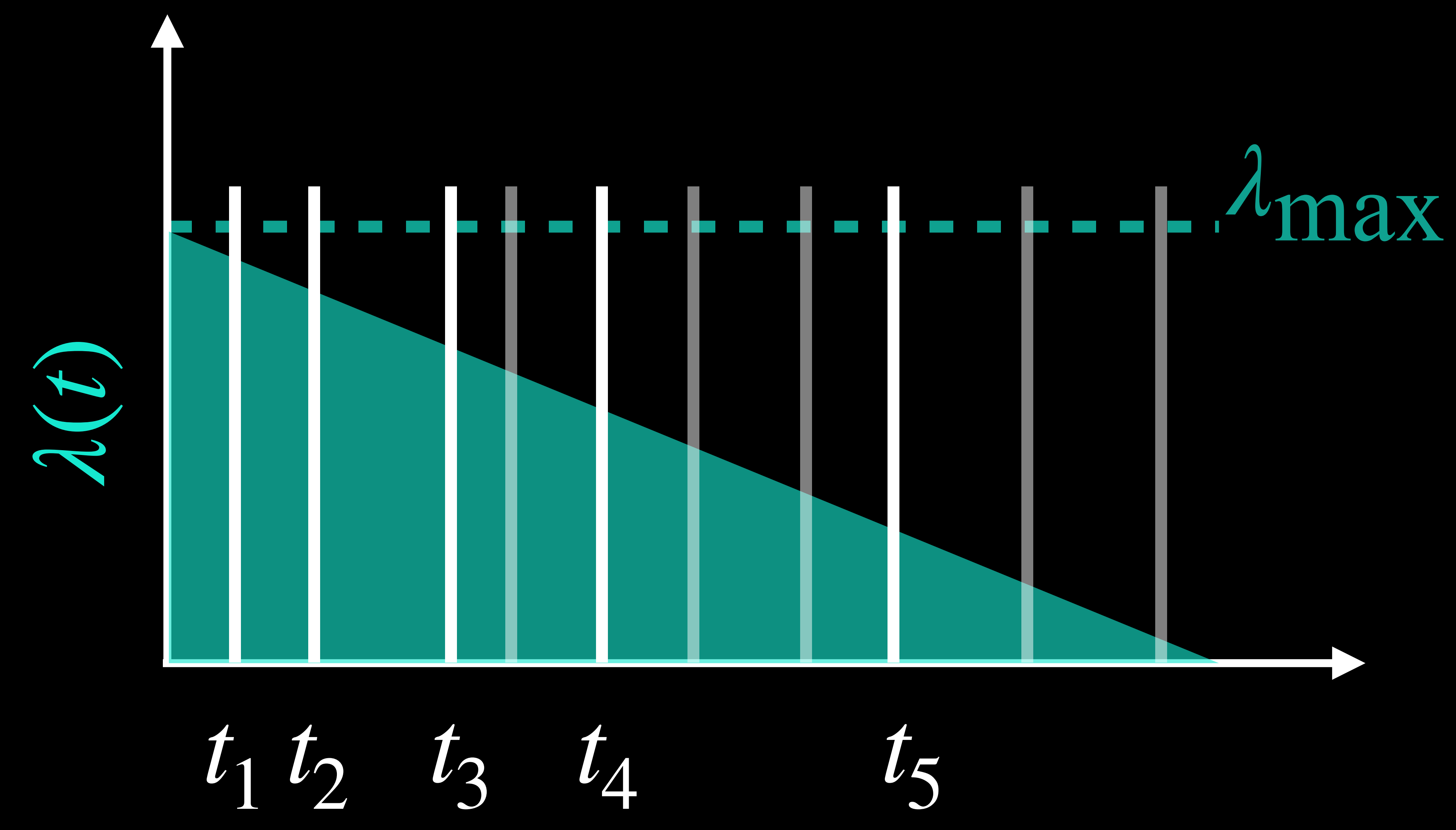
$$u \sim U[(0,1)]$$

$$t^c = \frac{-ln(u)}{\lambda_{max}}$$

Probability of acceptance for all candidate events

$$\frac{\lambda(t^c)}{\lambda_{max}}$$

Thinning Method



Generate candidate events

$$u \sim U[(0,1)]$$

$$t^c = \frac{-ln(u)}{\lambda_{max}}$$

Probability of acceptance for all candidate events

$$\frac{\lambda(t^c)}{\lambda_{max}}$$

We accept candidate event if

$$u' \sim U([0,1]) \qquad u' \leq \frac{\lambda(t^c)}{\lambda_{max}}$$

Thinning Method SCM

Noise
Uniformly at random in $[0,1]$

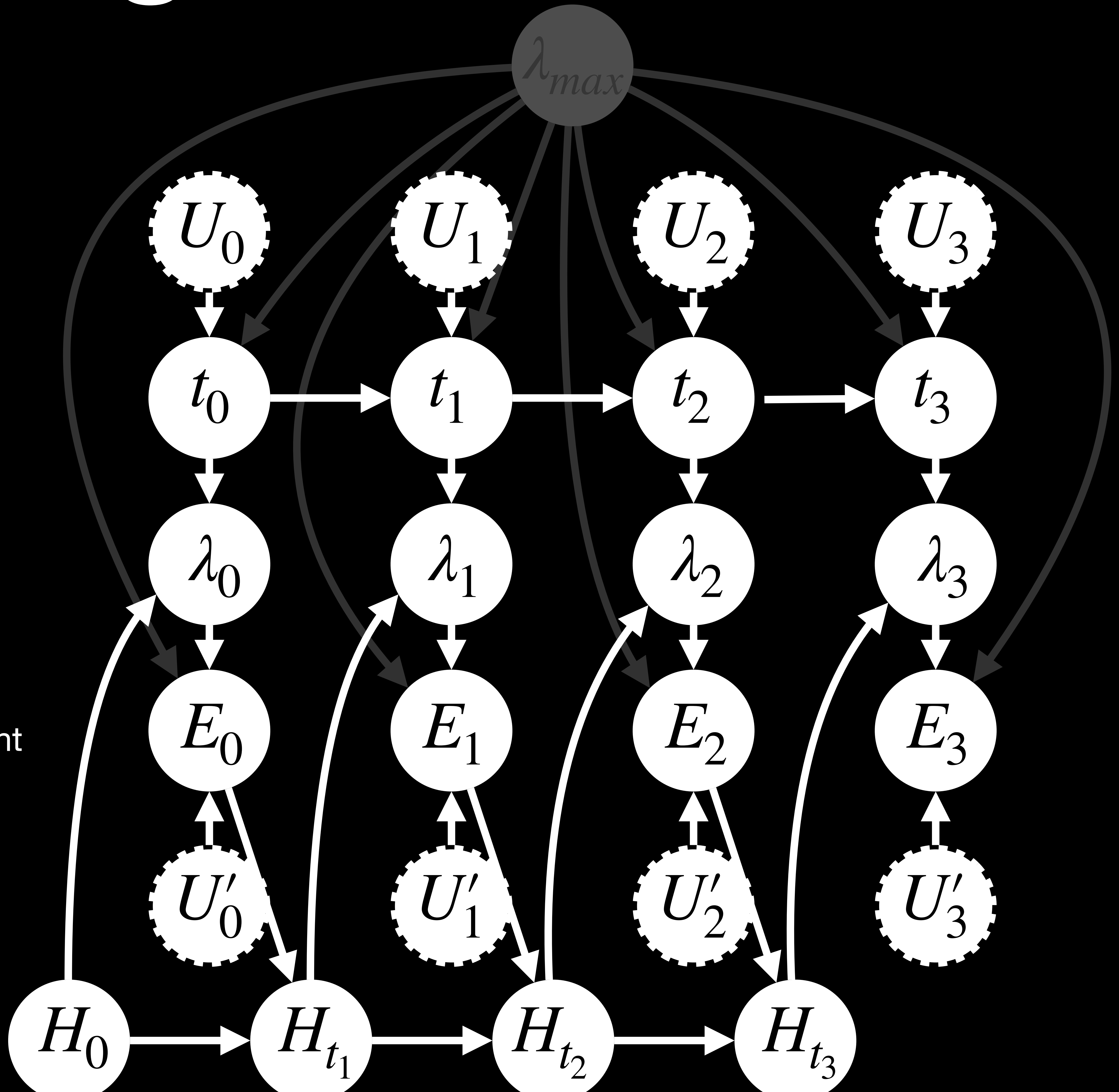
Event Time
Of the homogenous TPP

Intensity
Actual rate at time t

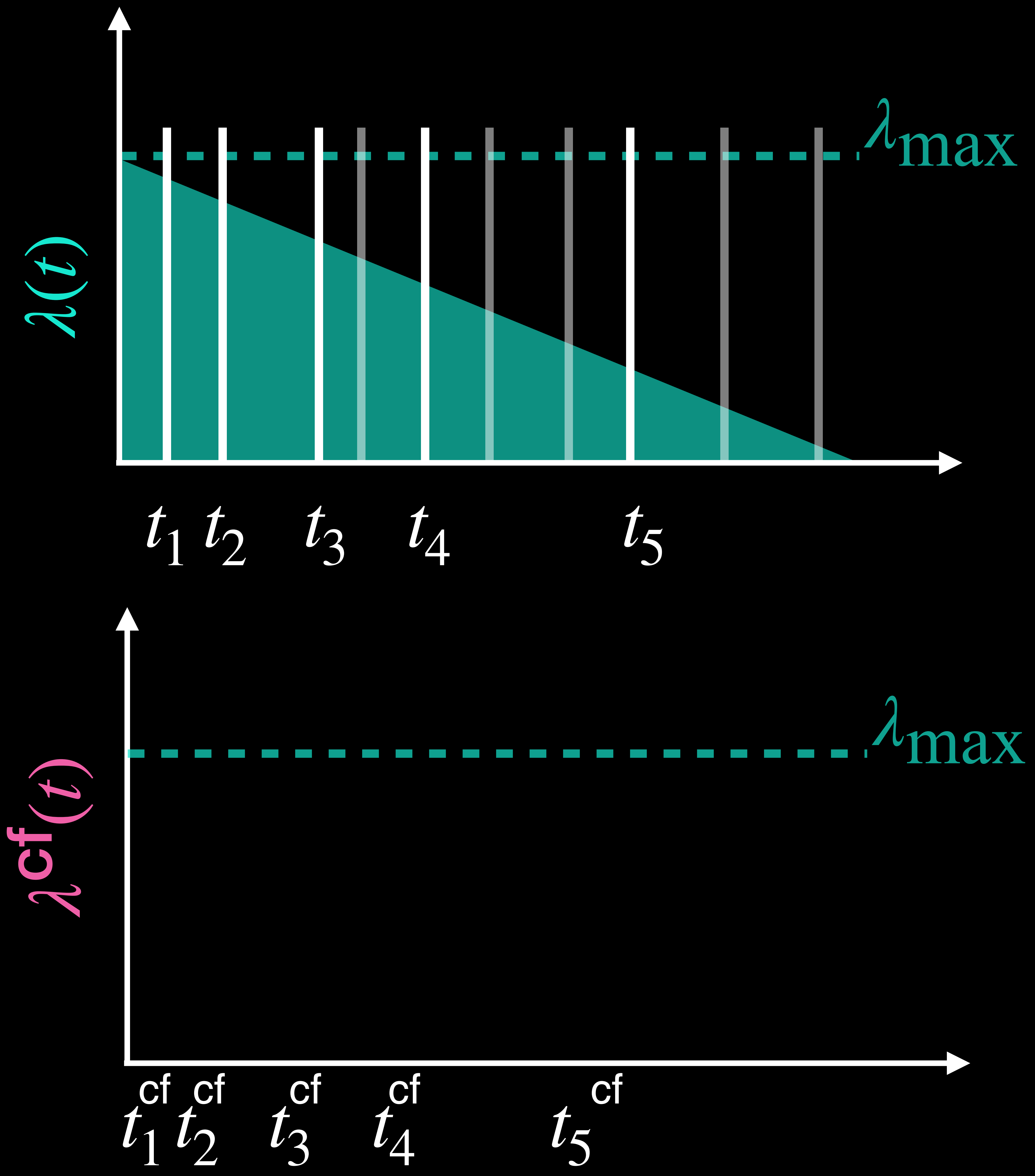
Event
Rejection or acceptance of event

Noise
Uniformly at random in $[0,1]$

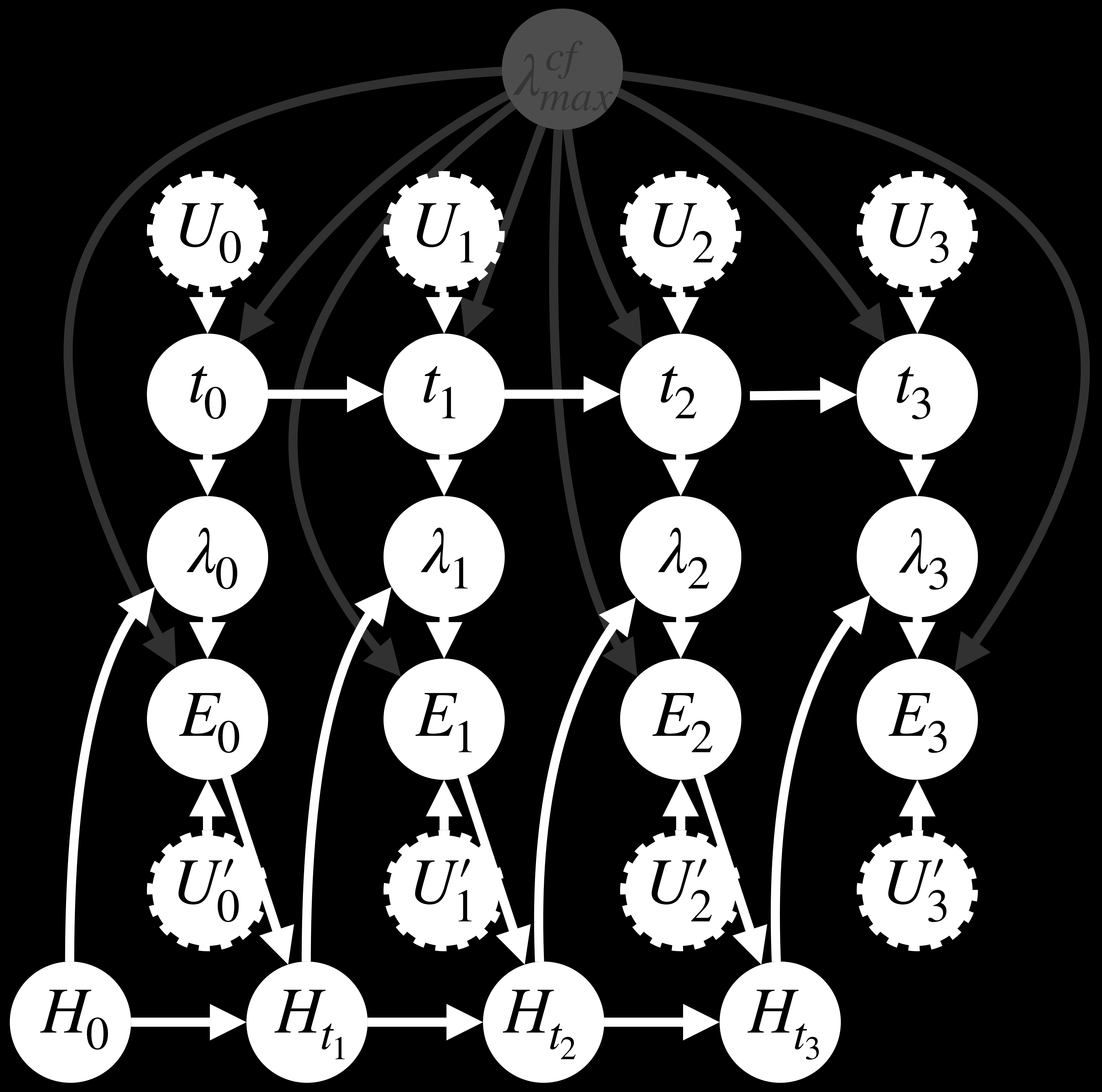
History
Past events up event i



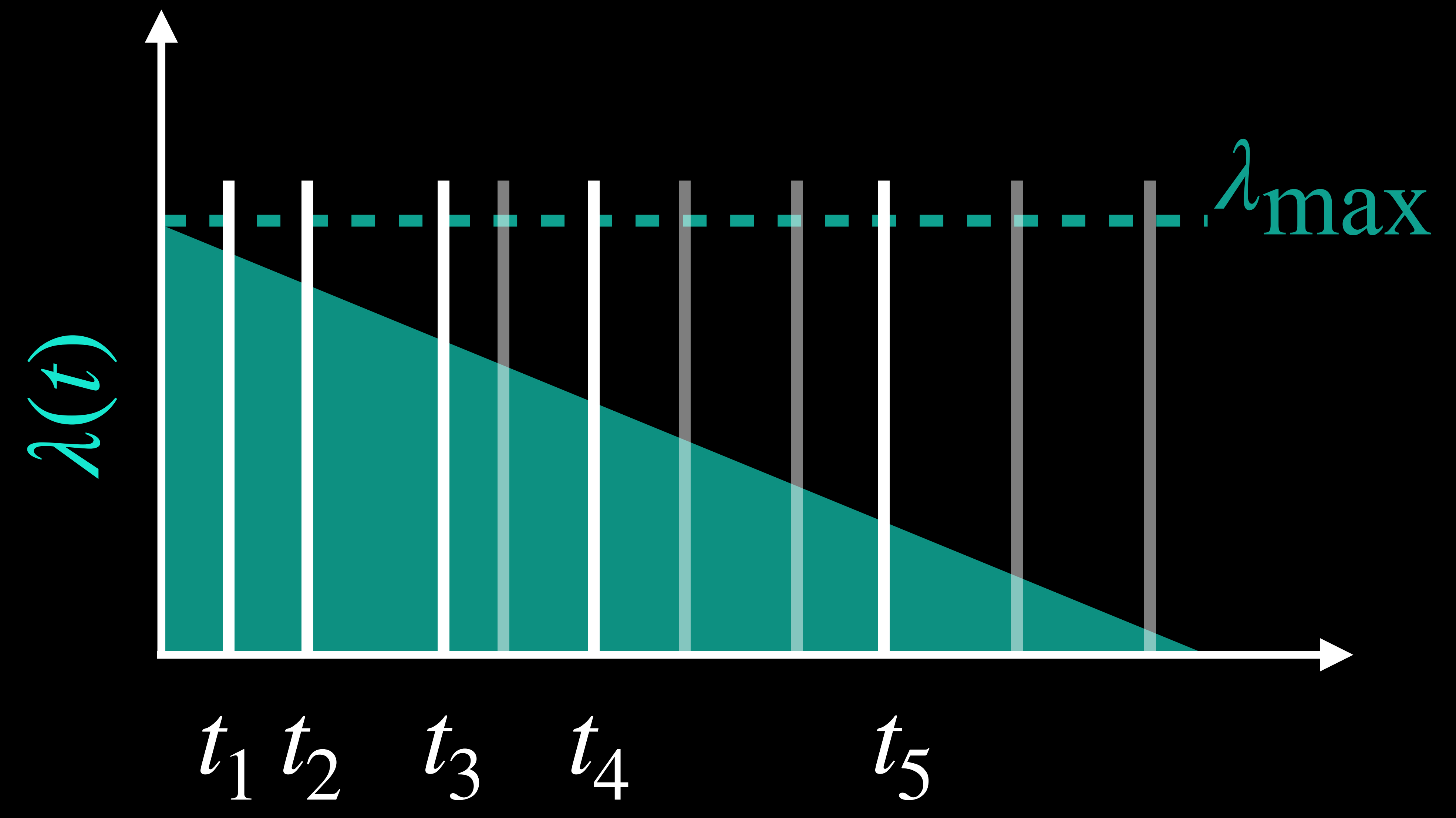
Thinning Method Counterfactual Generation

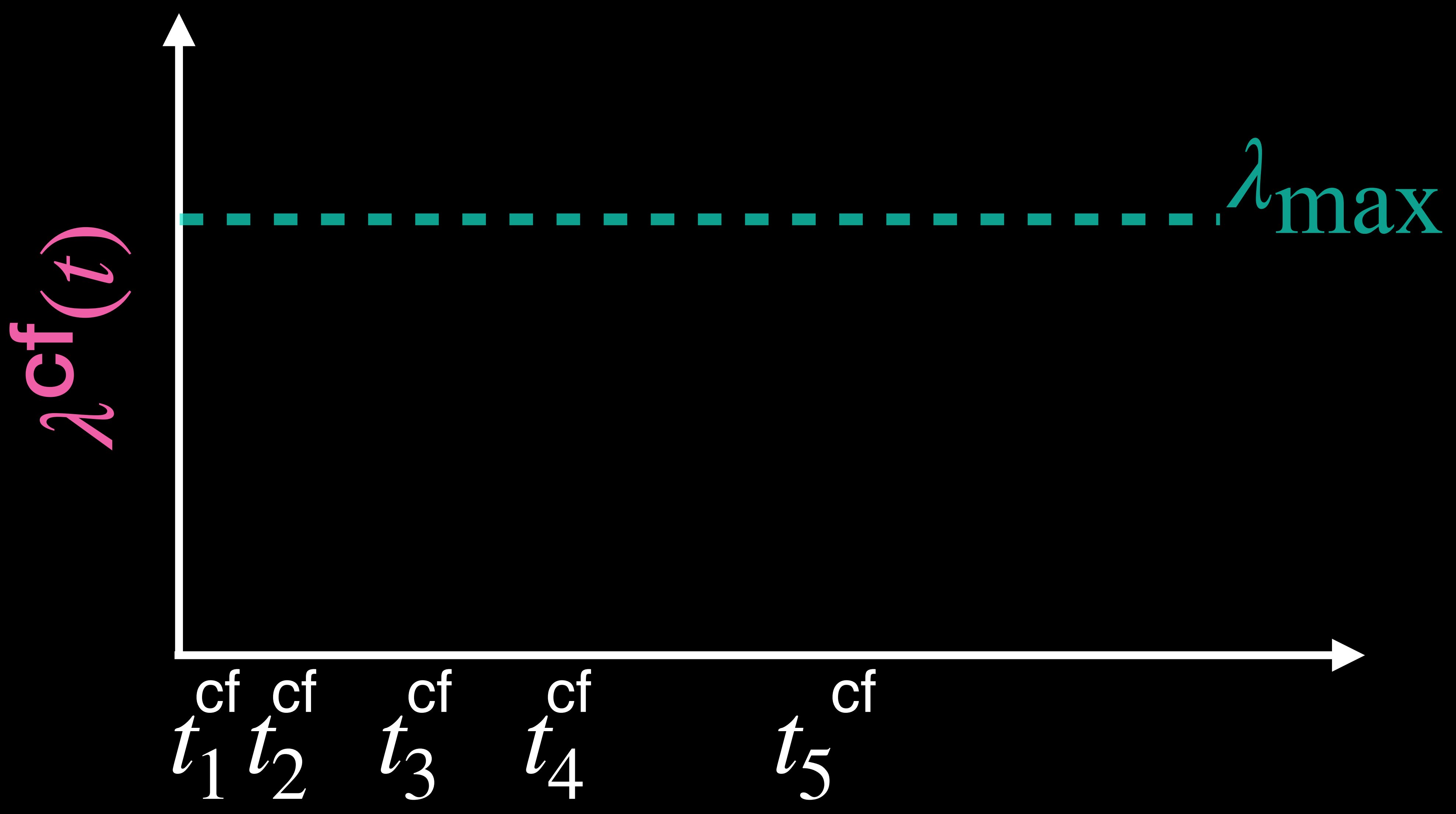


History
Past events up
event i



Thinning Method Counterfactual Generation

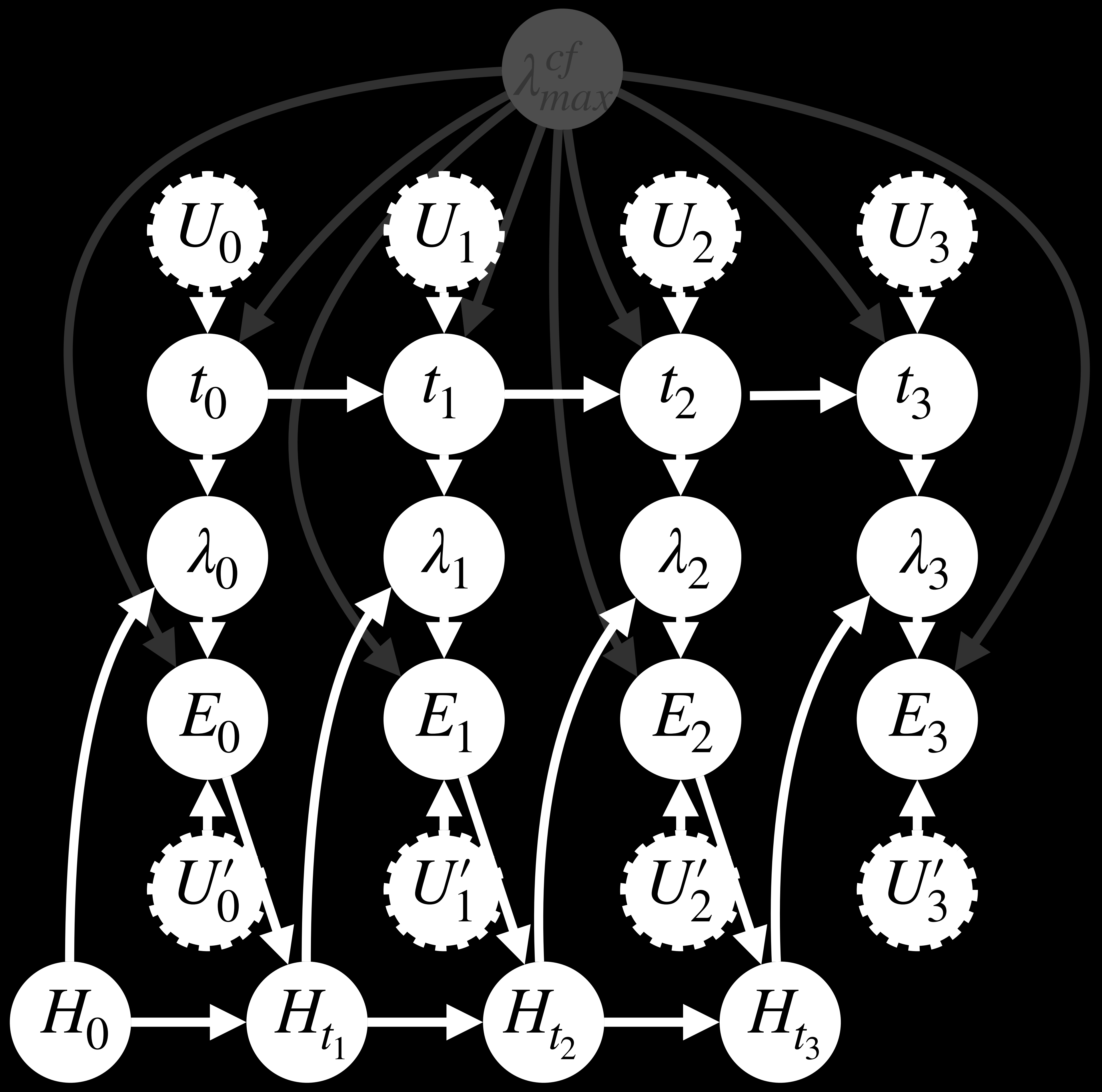




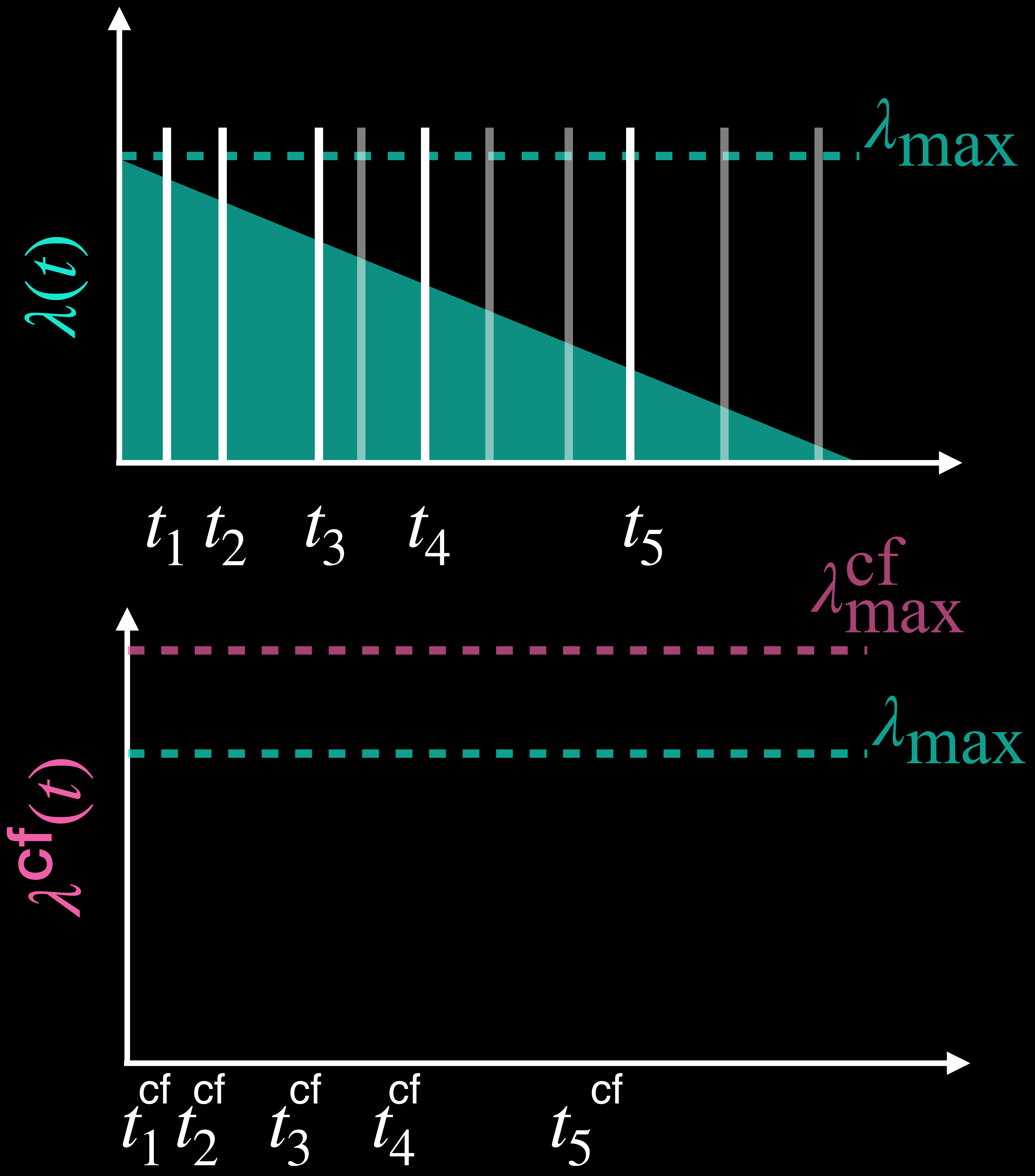
Noise posterior for candidate event generation

Noise posterior for accepting/rejecting candidate events

History
Past events up event i



Thinning Method Counterfactual Generation

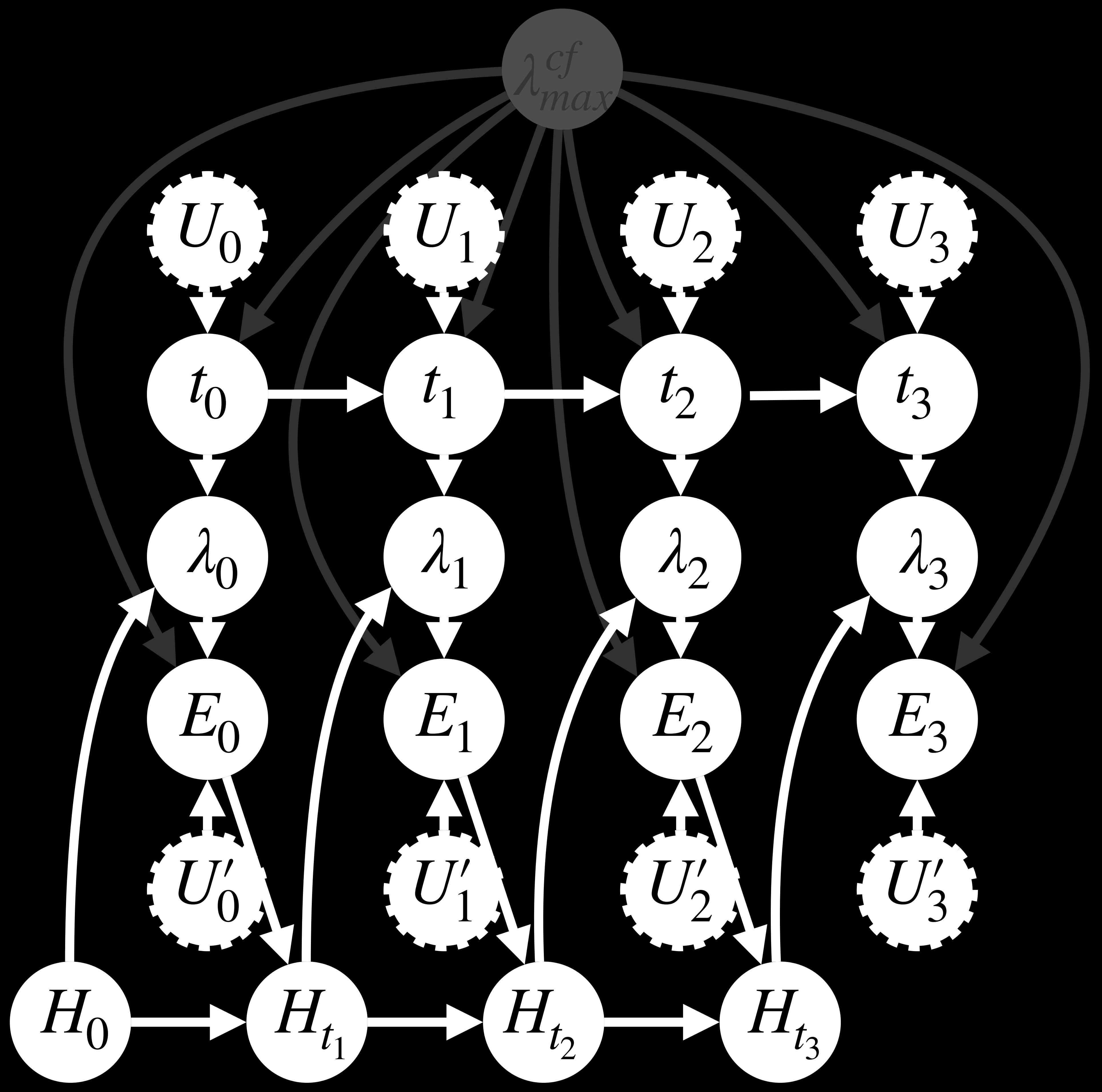


Hypothetical counterfactual λ^{cf}_{max}

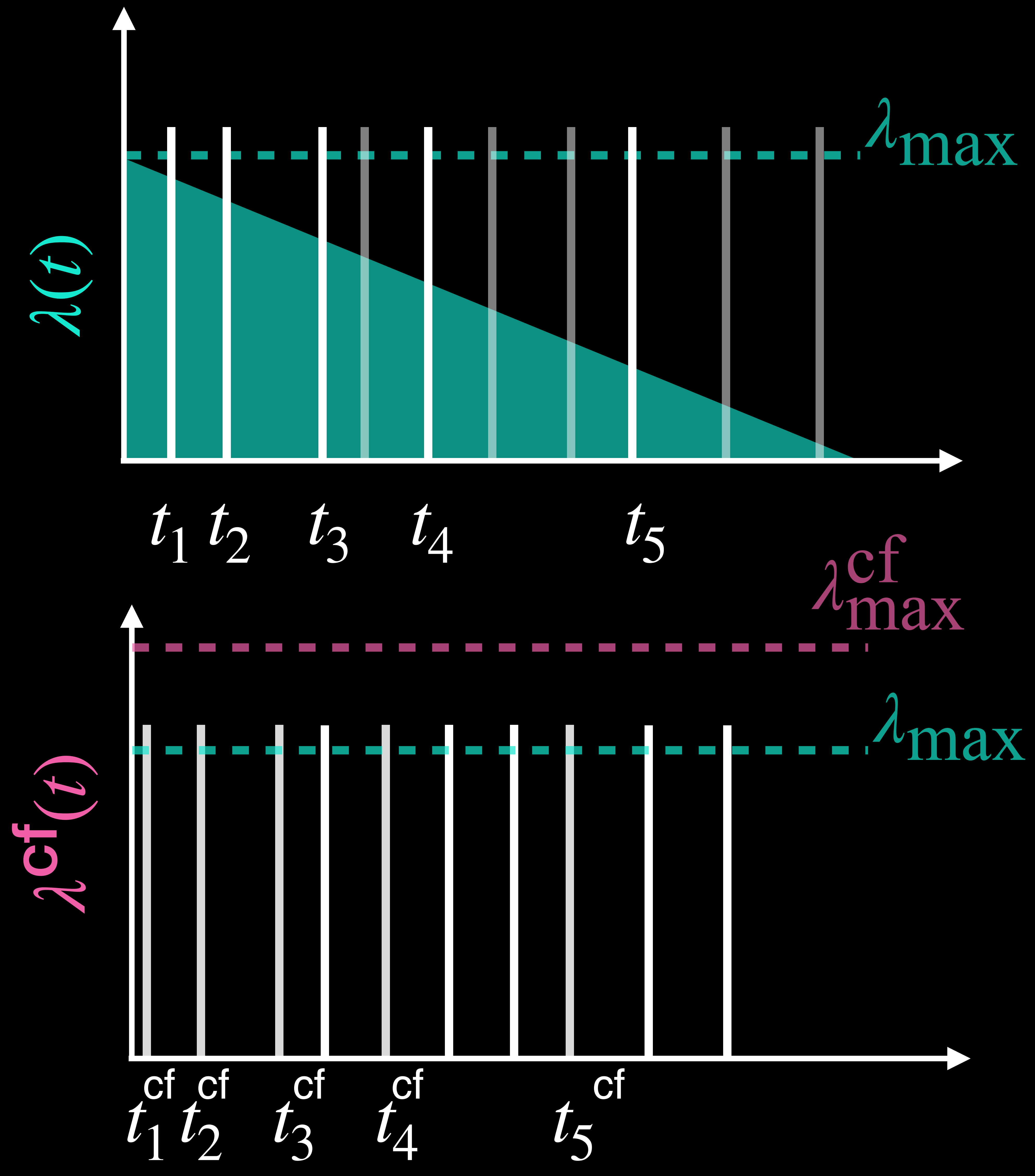
Noise posterior for candidate event generation

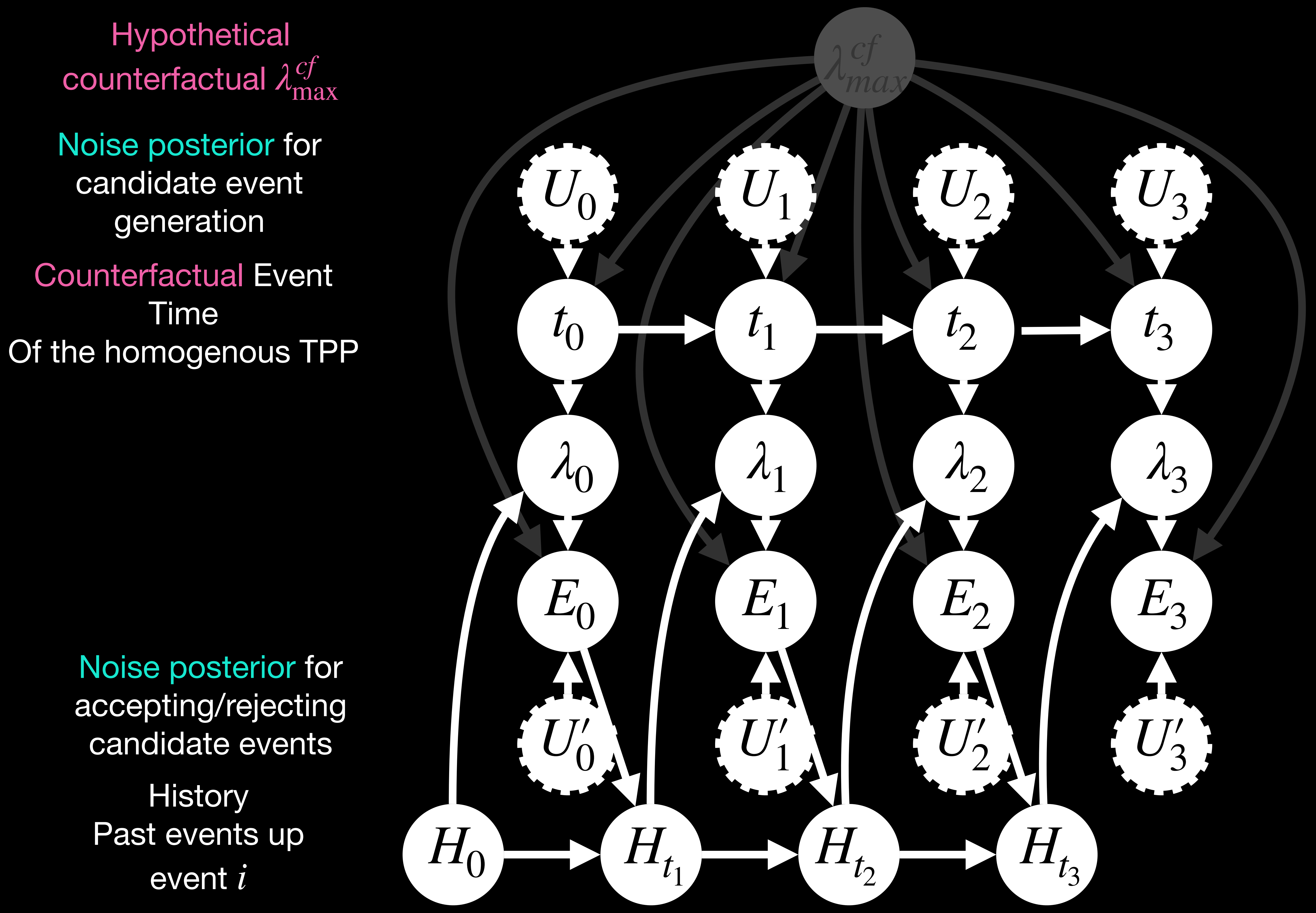
Noise posterior for accepting/rejecting candidate events

History
Past events up event i

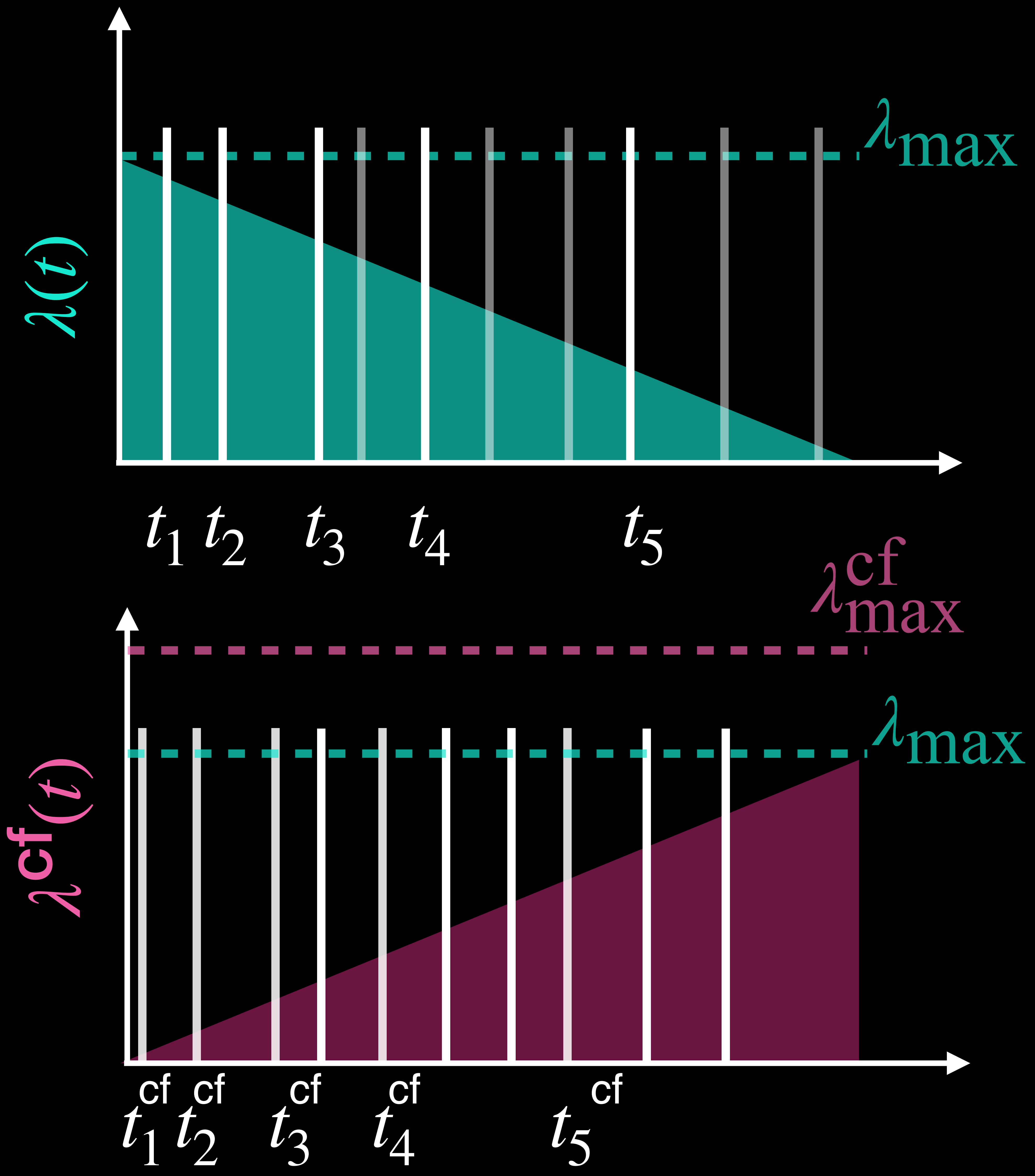


Thinning Method Counterfactual Generation





Thinning Method Counterfactual Generation



Hypothetical counterfactual λ^{cf}_{max}

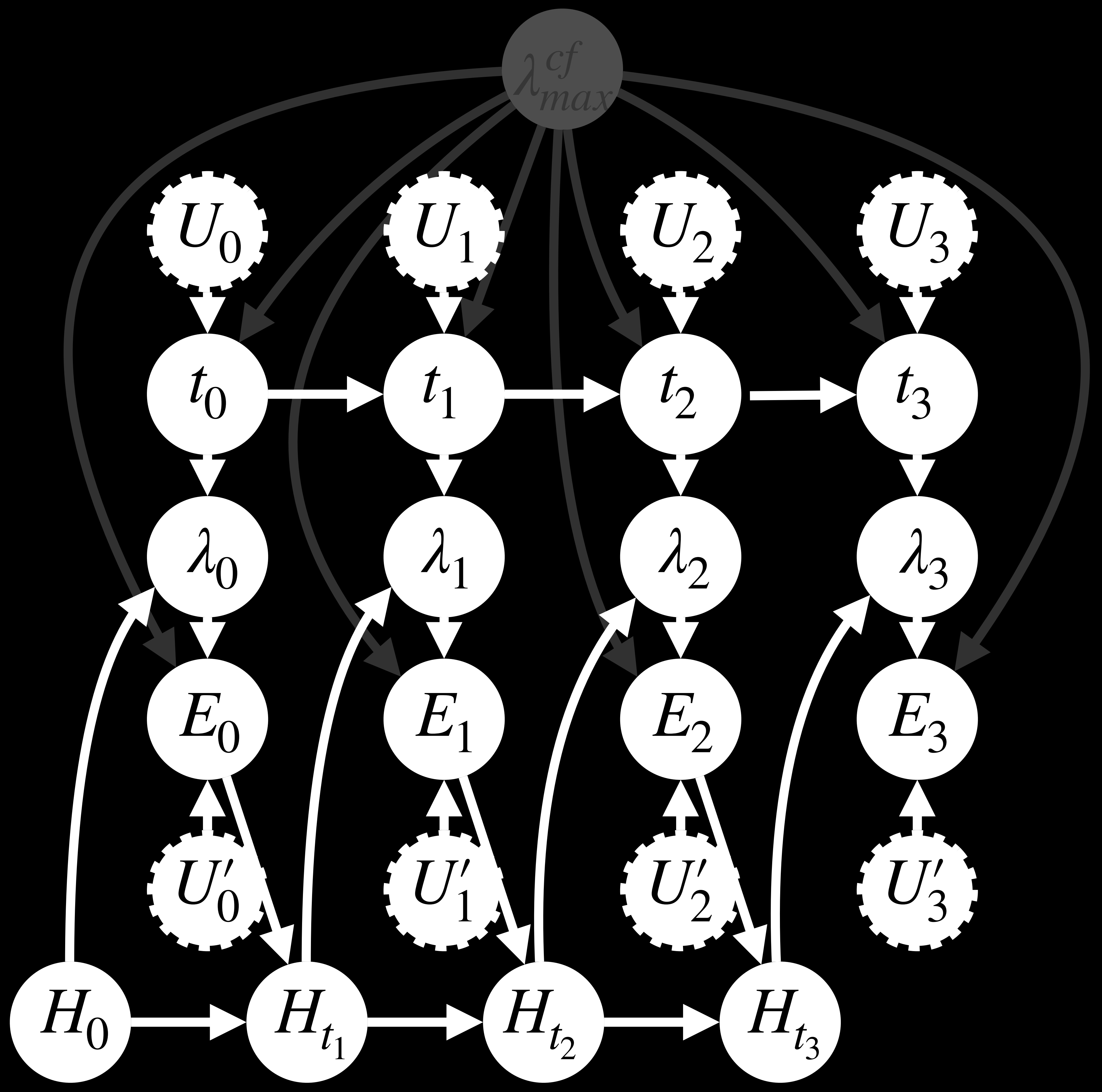
Noise posterior for candidate event generation

Counterfactual Event Time Of the homogenous TPP

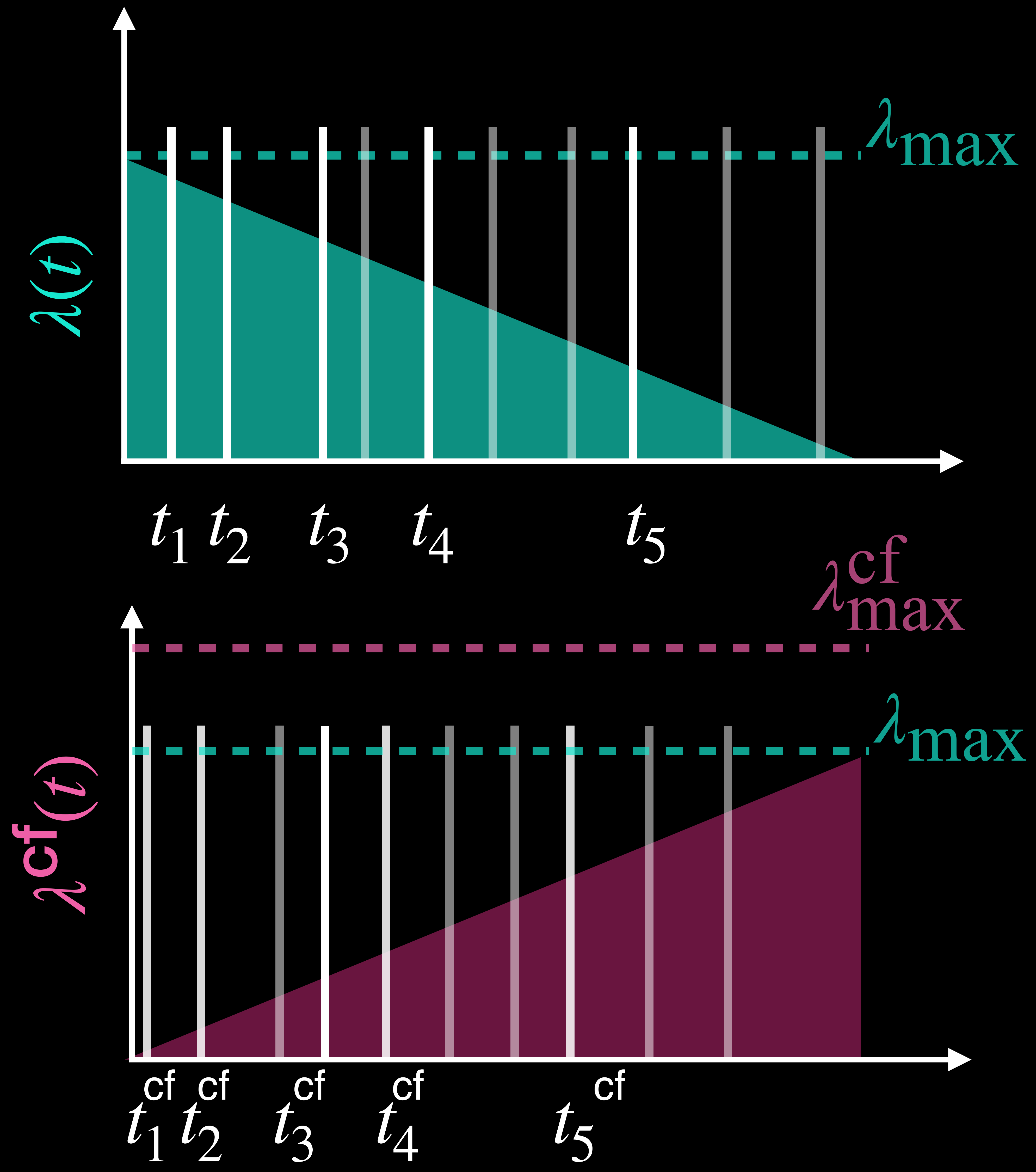
Intensity Hypothetical counterfactual rate at time t

Noise posterior for accepting/rejecting candidate events

History Past events up event i



Thinning Method Counterfactual Generation



Hypothetical counterfactual λ^{cf}_{max}

Noise posterior for candidate event generation

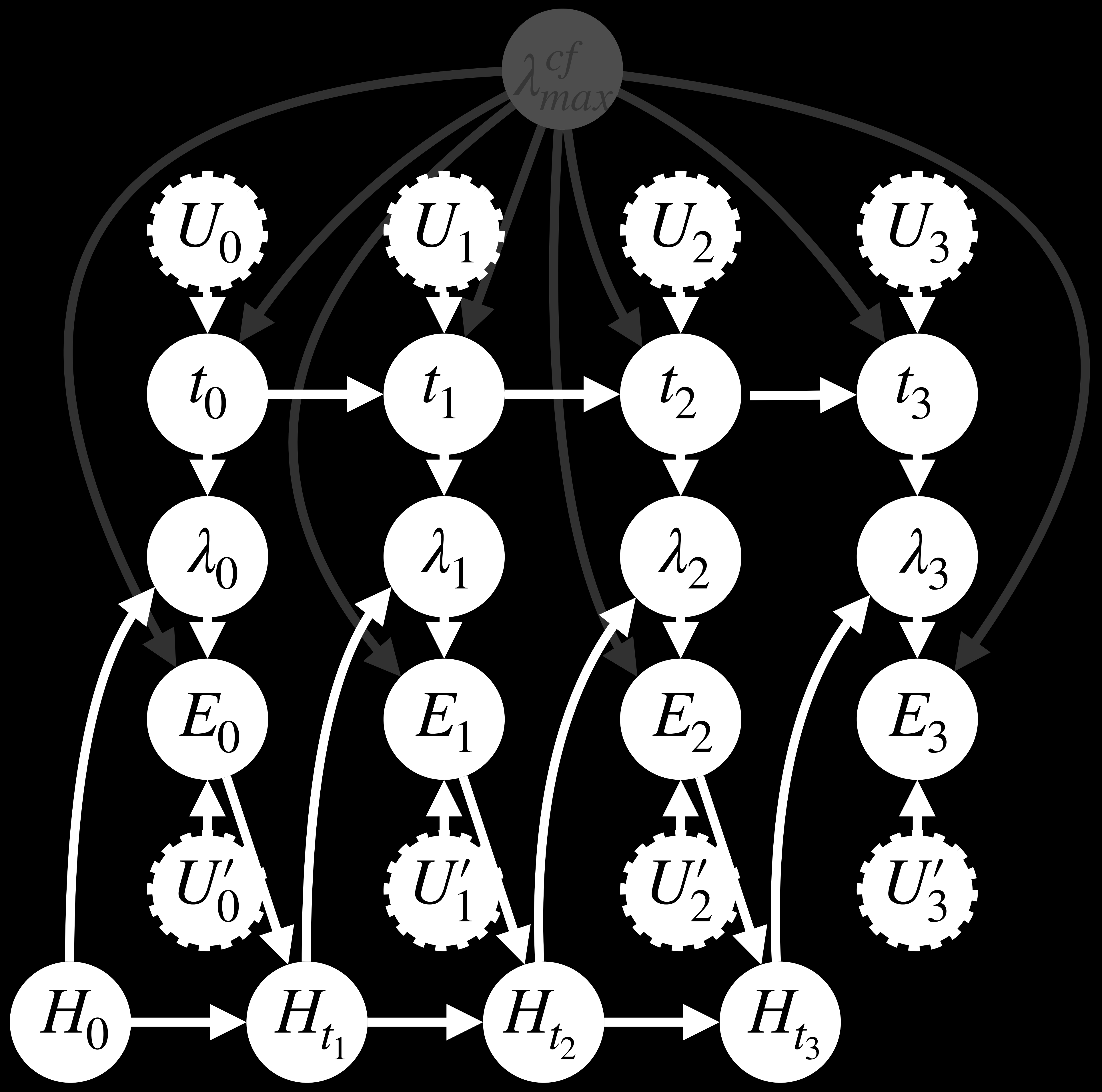
Counterfactual Event Time Of the homogenous TPP

Intensity Hypothetical counterfactual rate at time t

Counterfactual Event Rejection or acceptance

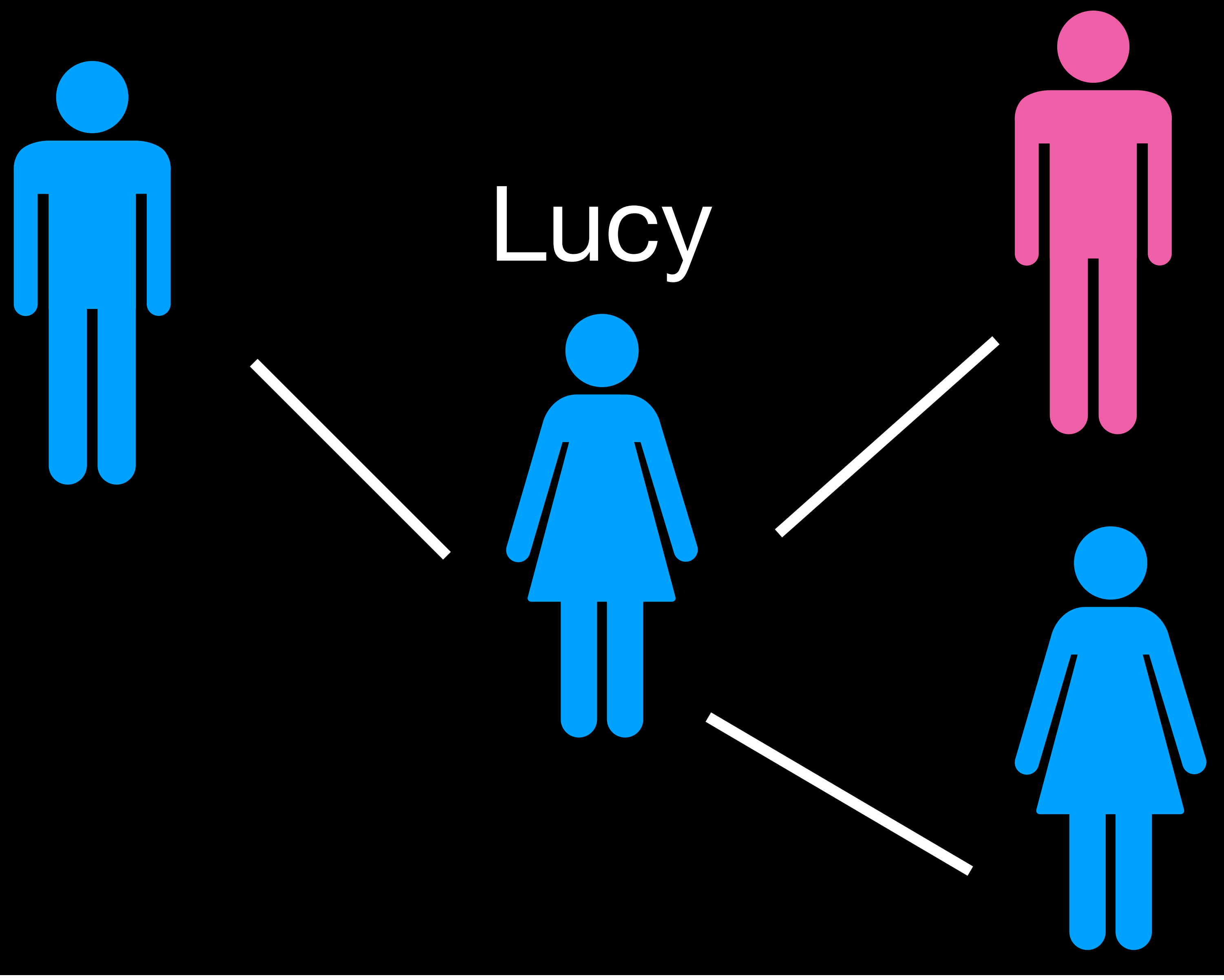
Noise posterior for accepting/rejecting candidate events

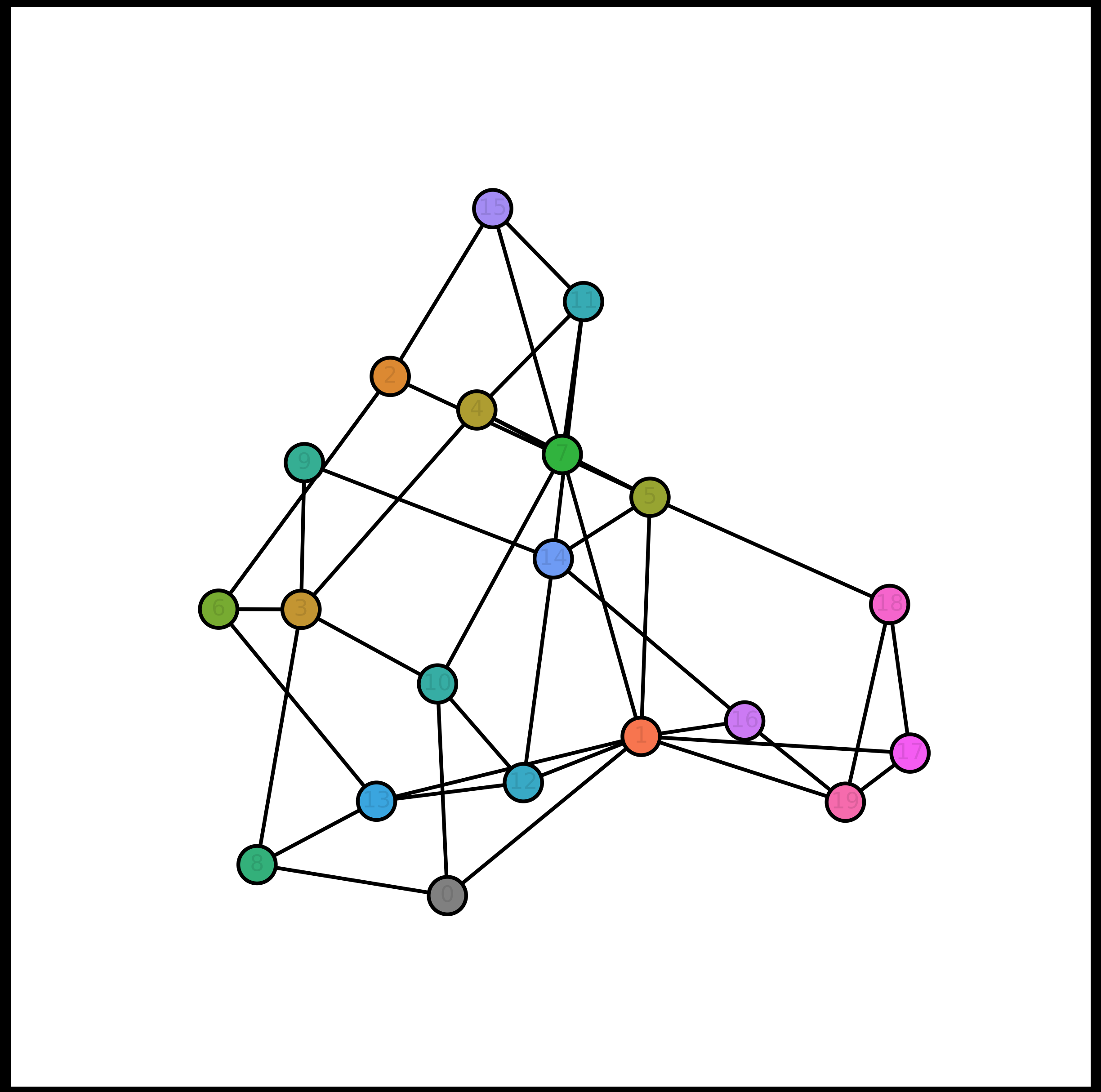
History Past events up event i

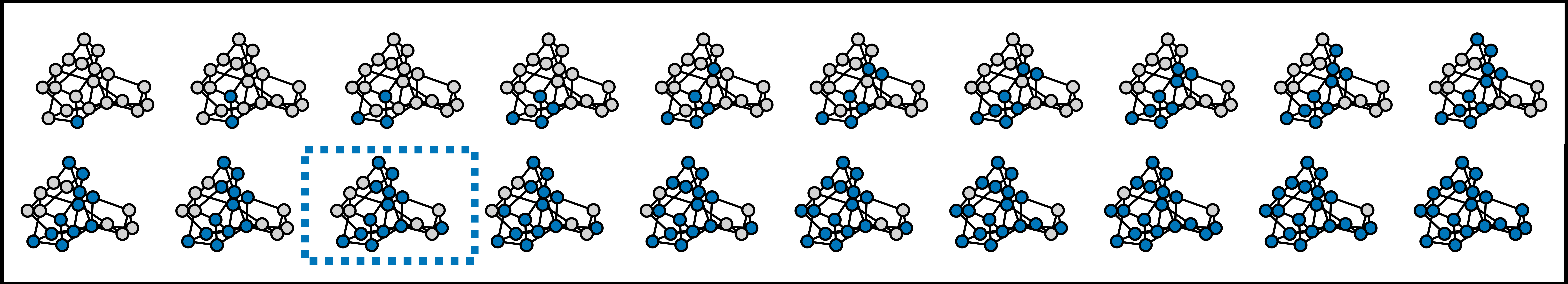


Revisiting our Case Study: A Mechanistic Model of Epidemic Spreading

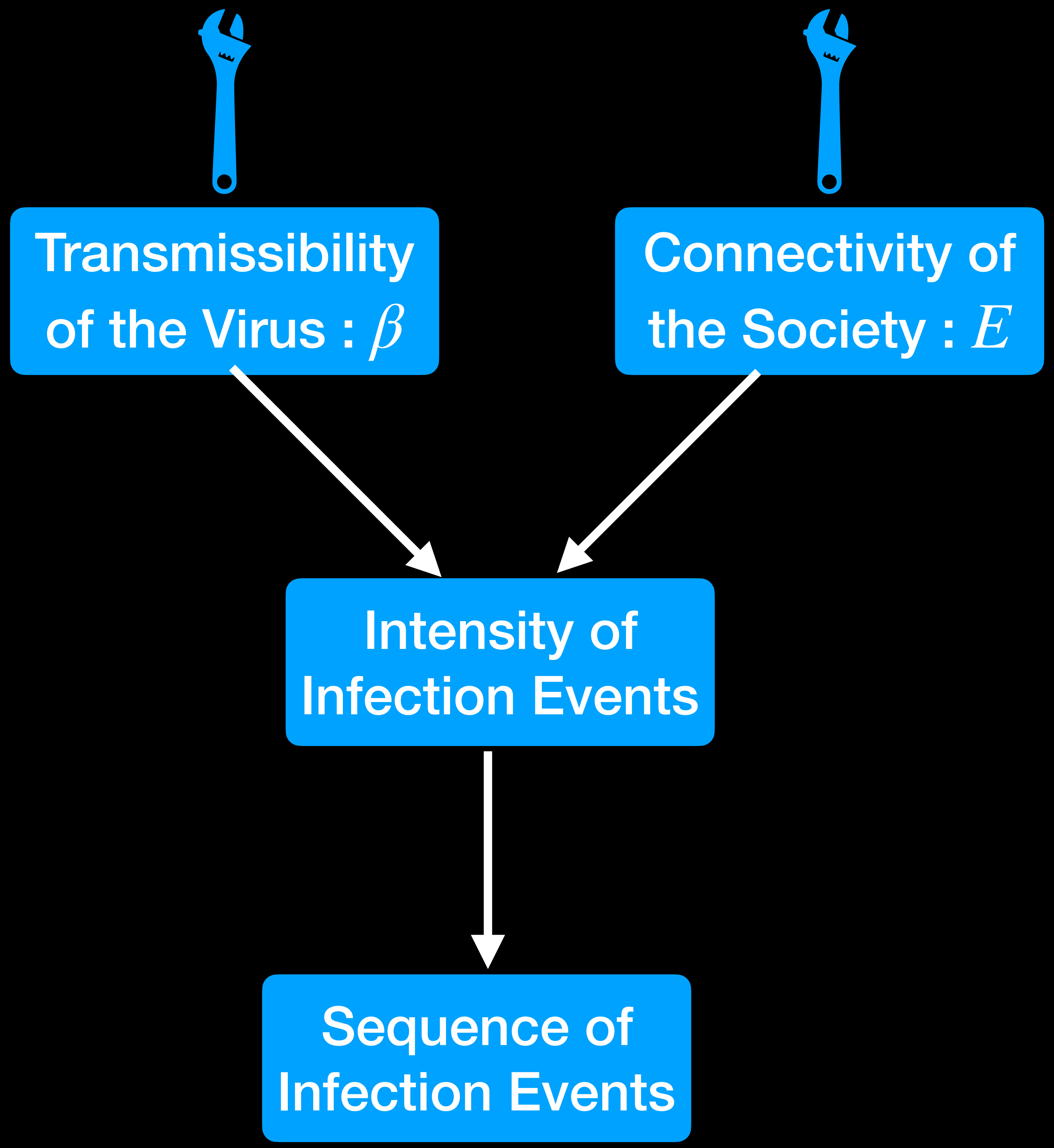
Simulation Case Study

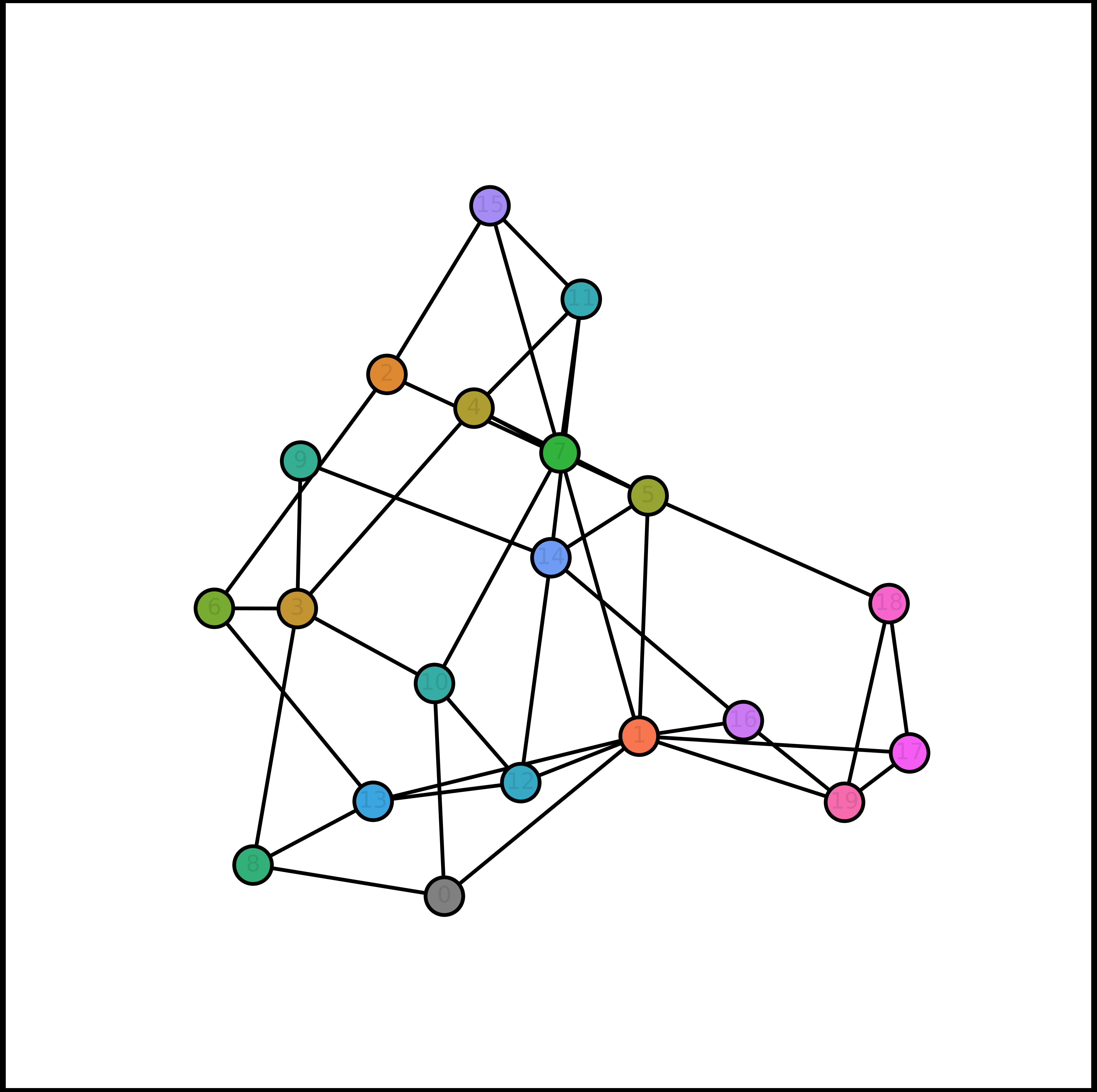




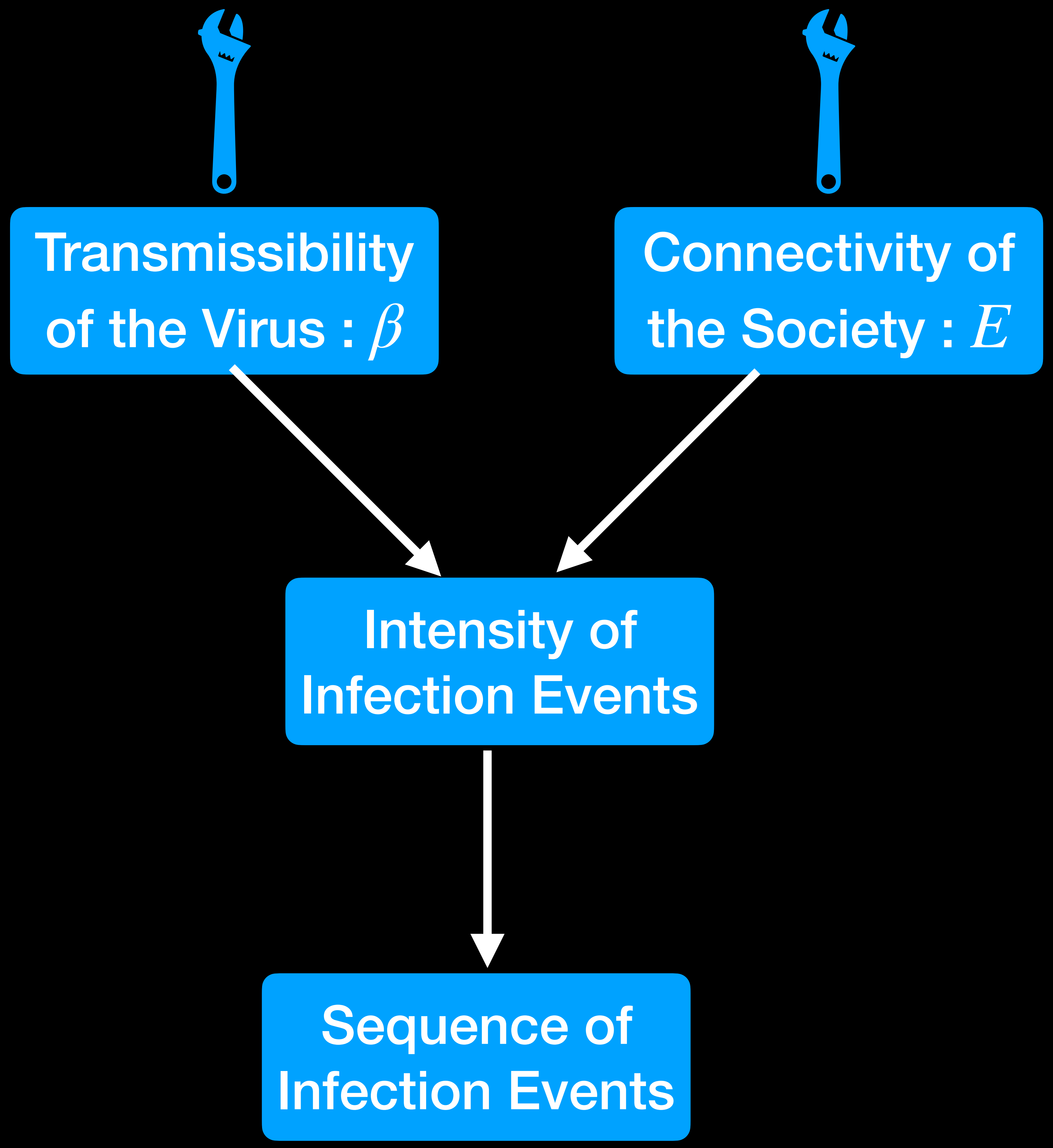


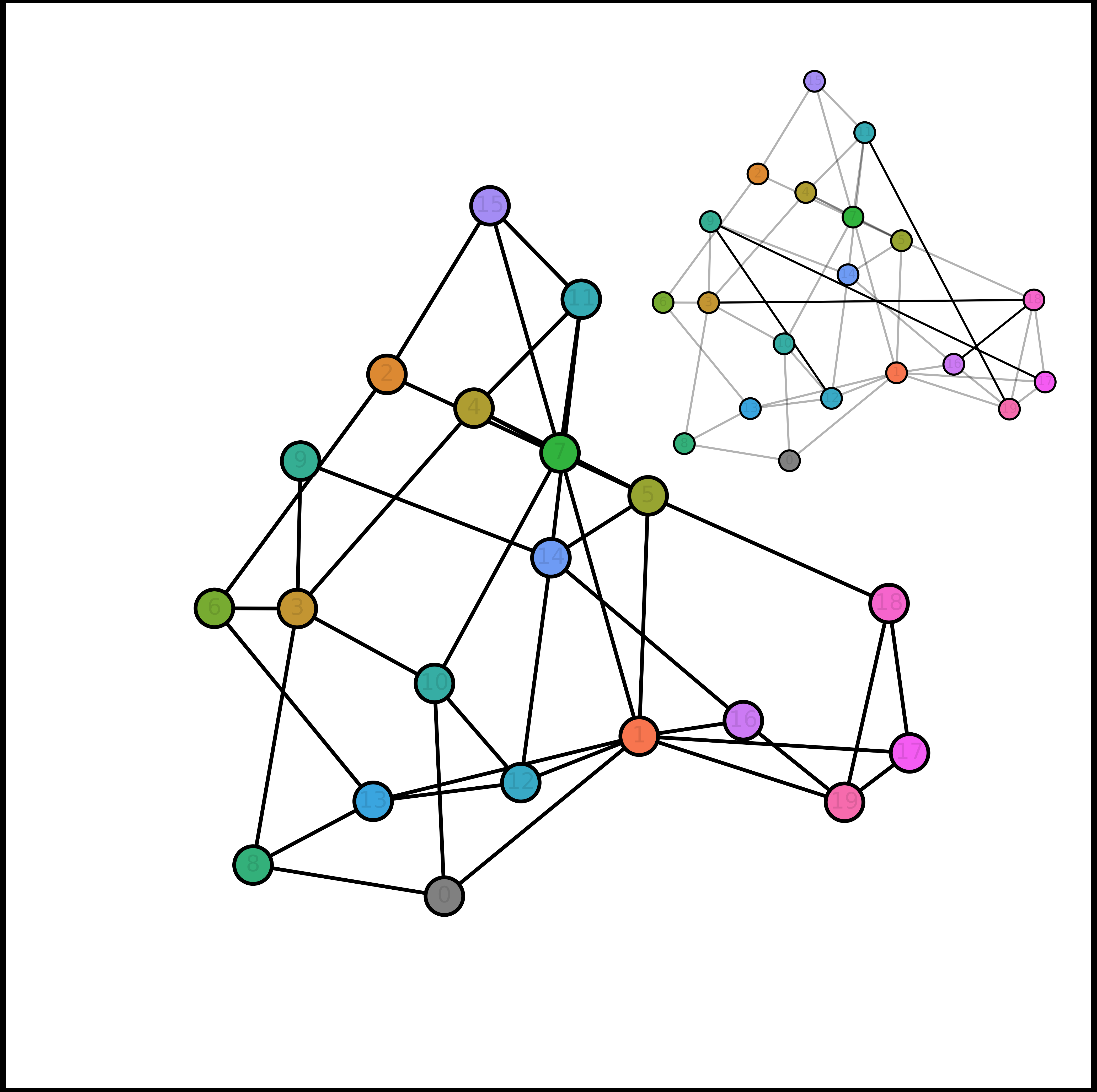
Computing Counterfactuals





Computing Counterfactuals

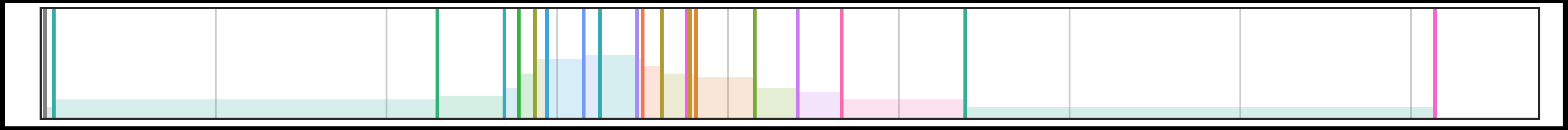




Counterfactual Simulations

Infections as Event Sequence

Original sequence

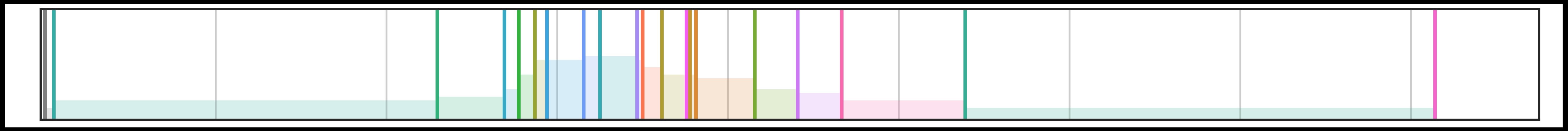


Time

Counterfactual Simulations

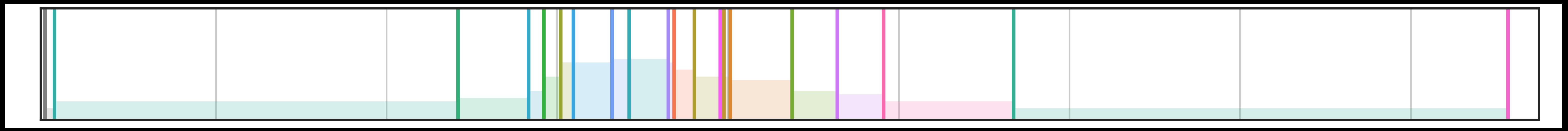
Infections as Event Sequence

Original sequence



Decreasing β (infectivity) decelerates the infection process

Counterfactual sequence with smaller infection rat

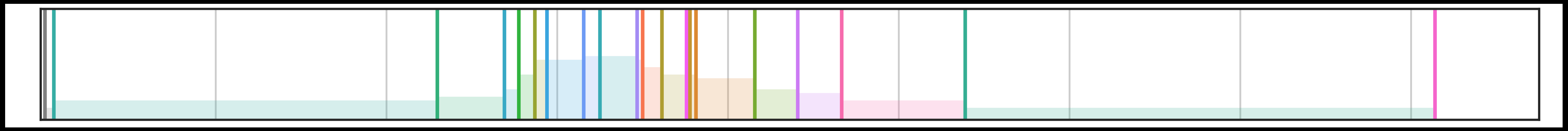


Time

Counterfactual Simulations

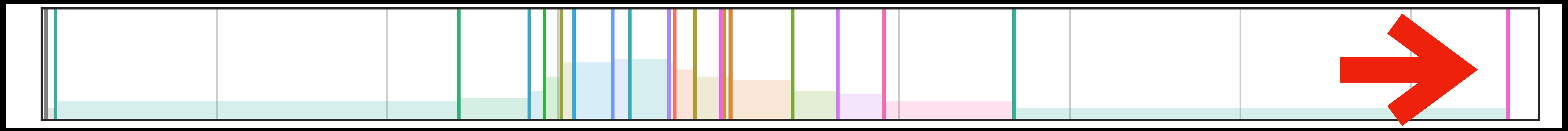
Infections as Event Sequence

Original sequence



Decreasing β (infectivity) decelerates the infection process

Counterfactual sequence with smaller infection rat

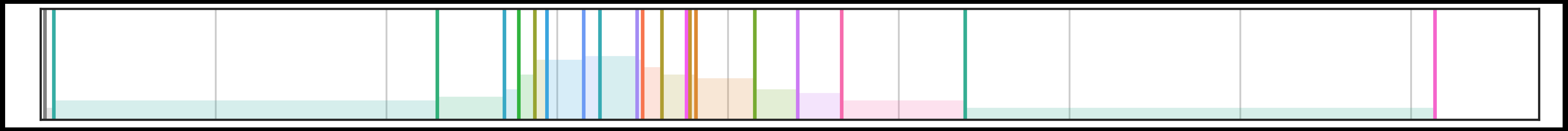


Time

Counterfactual Simulations

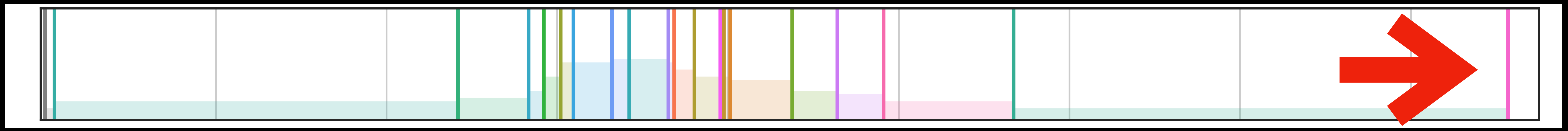
Infections as Event Sequence

Original sequence



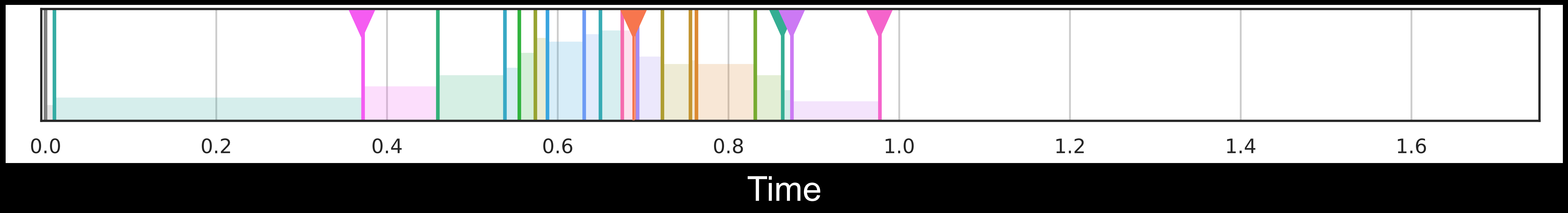
Decreasing β (infectivity) decelerates the infection process

Counterfactual sequence with smaller infection rat



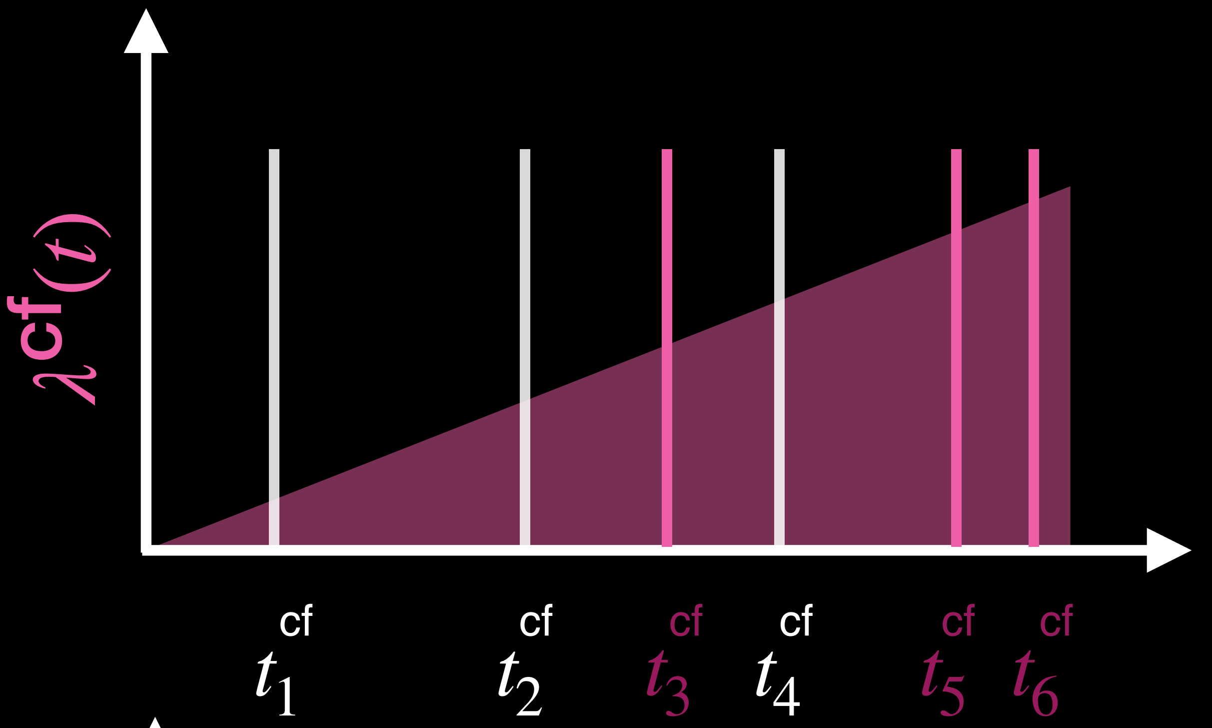
Introducing new edges (increasing connectivity) open new pathways

Counterfactual sequence with higher connectivity

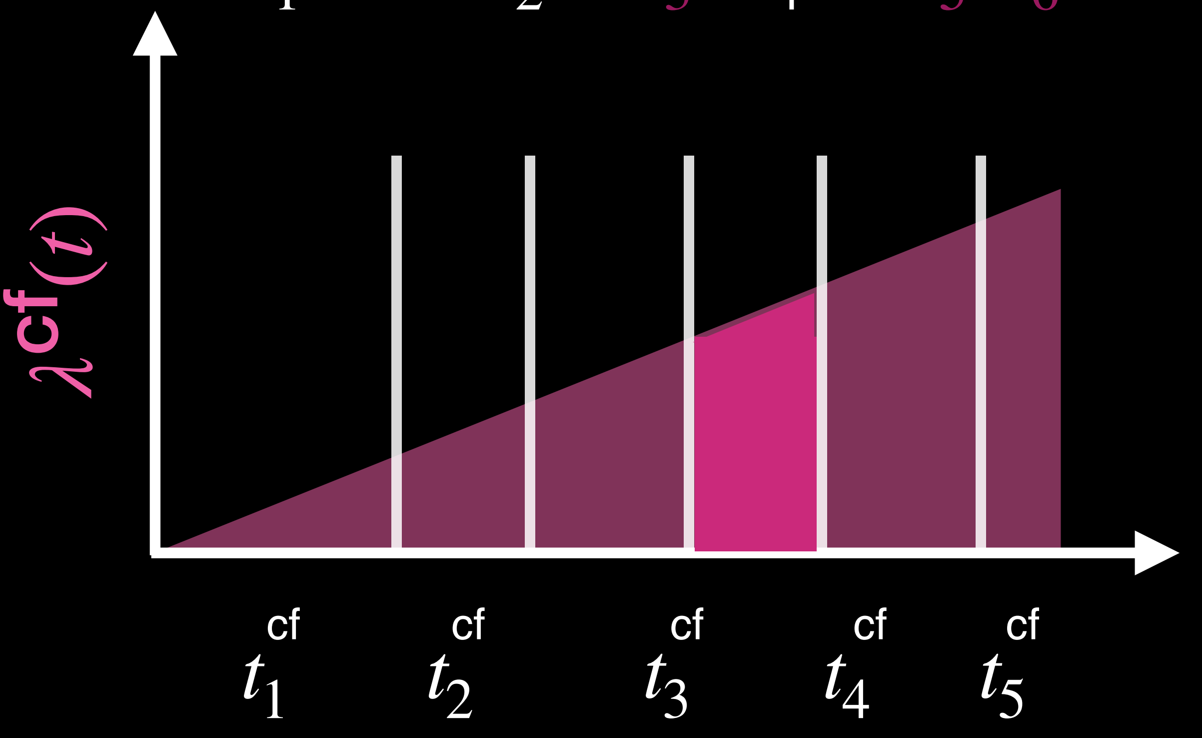


Recap

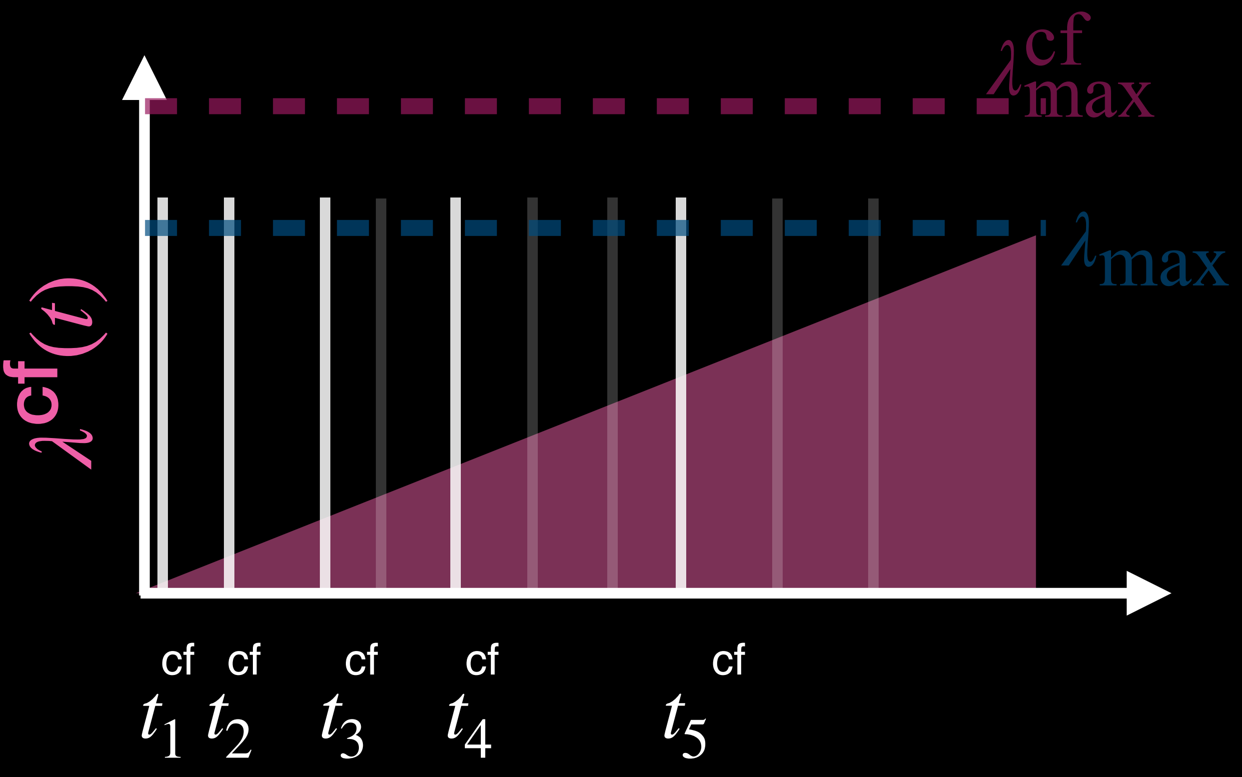
Naïve Method

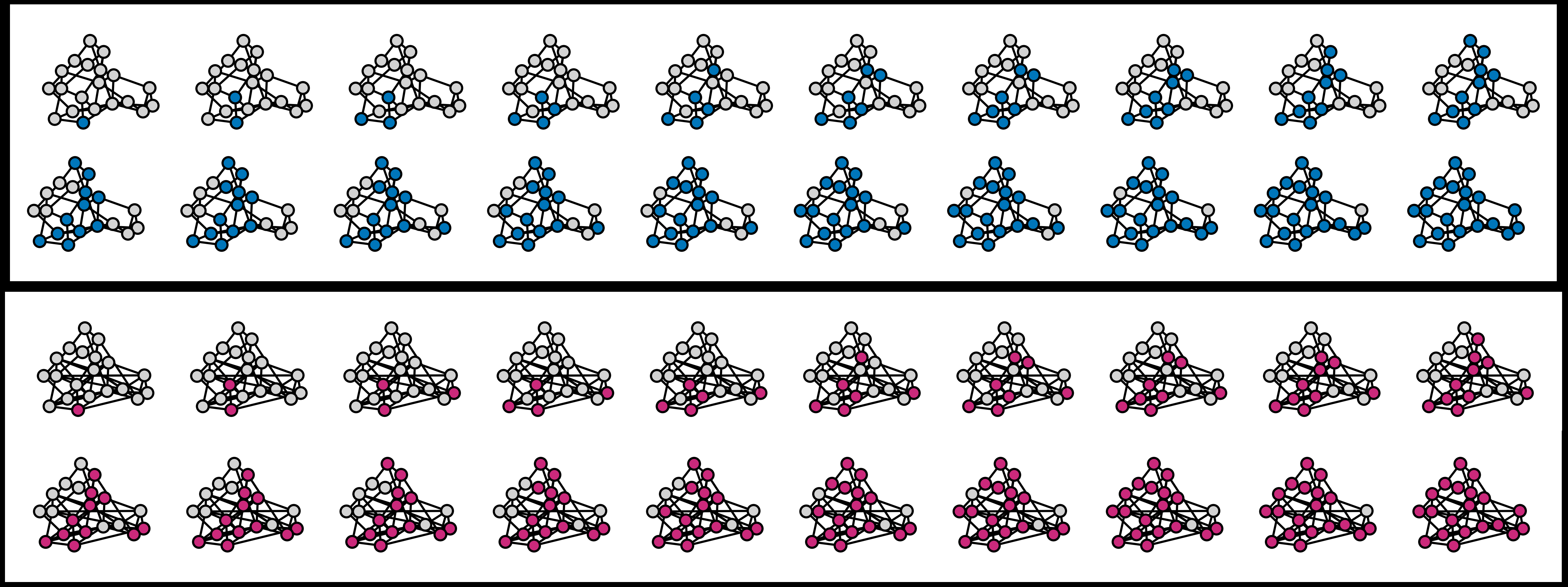


Numerical Integration Method



Thinning Method





Contact me at

sumantrak.mukherjee@dfki.de

Check out other cool projects

datasciapps.de

Thanks for your attention!

Monotonicity in Counterfactuals

P(H) = 0.4

P(T) = 0.6

P(H) = 0.3

P(T) = 0.7