More  for the buck

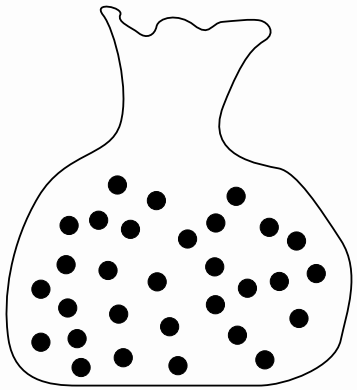
Distribute measurements for maximal information gain

Severin Bang

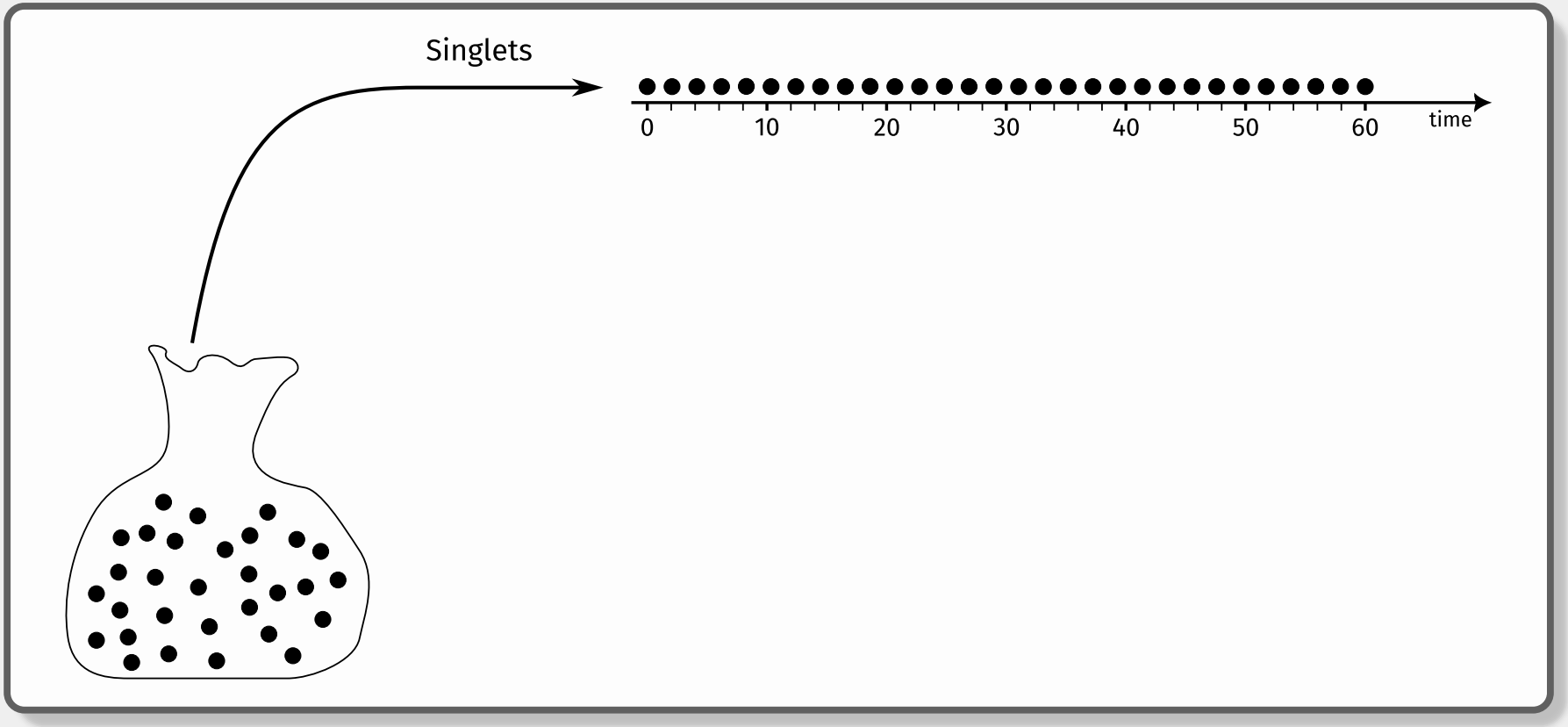
Institut of Physics, University of Freiburg

October 11, 2022

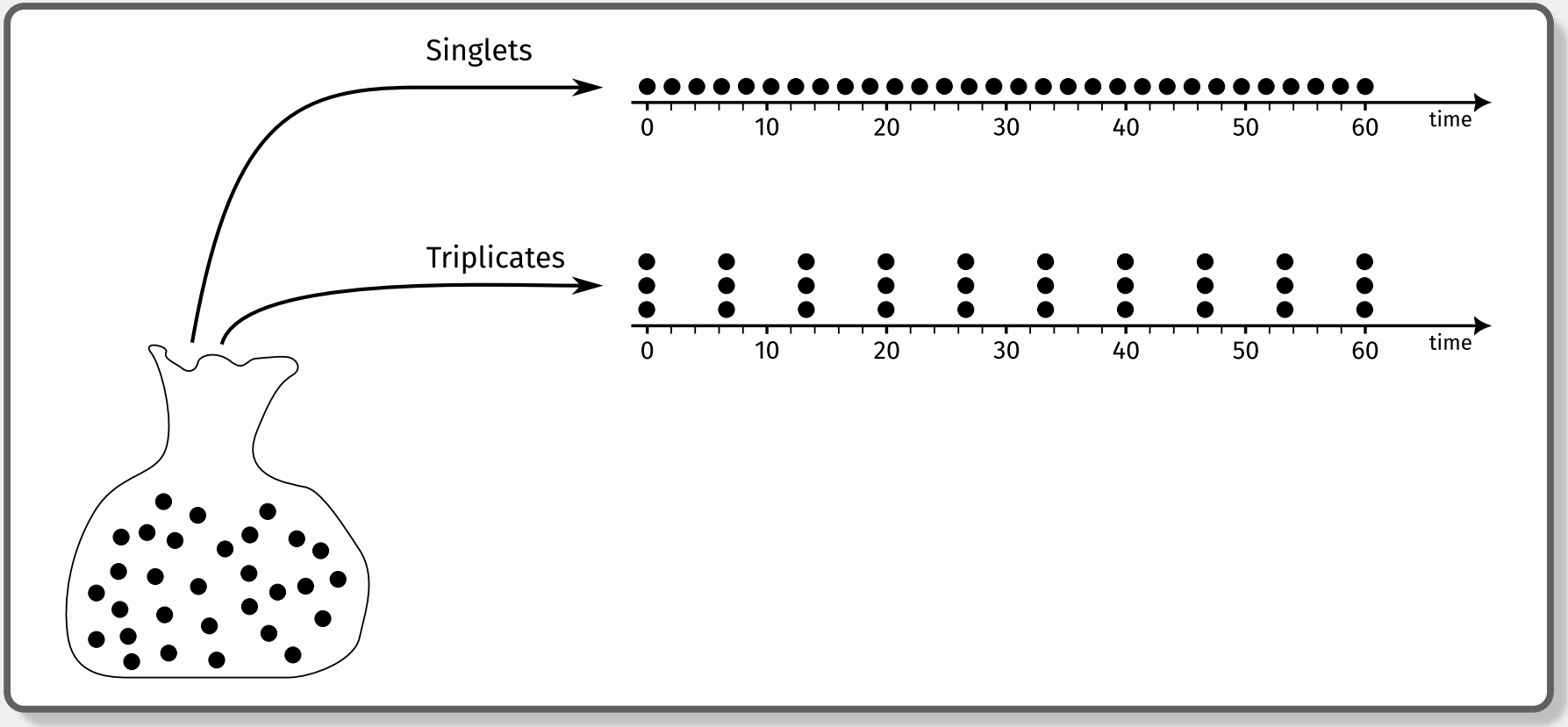
Premises



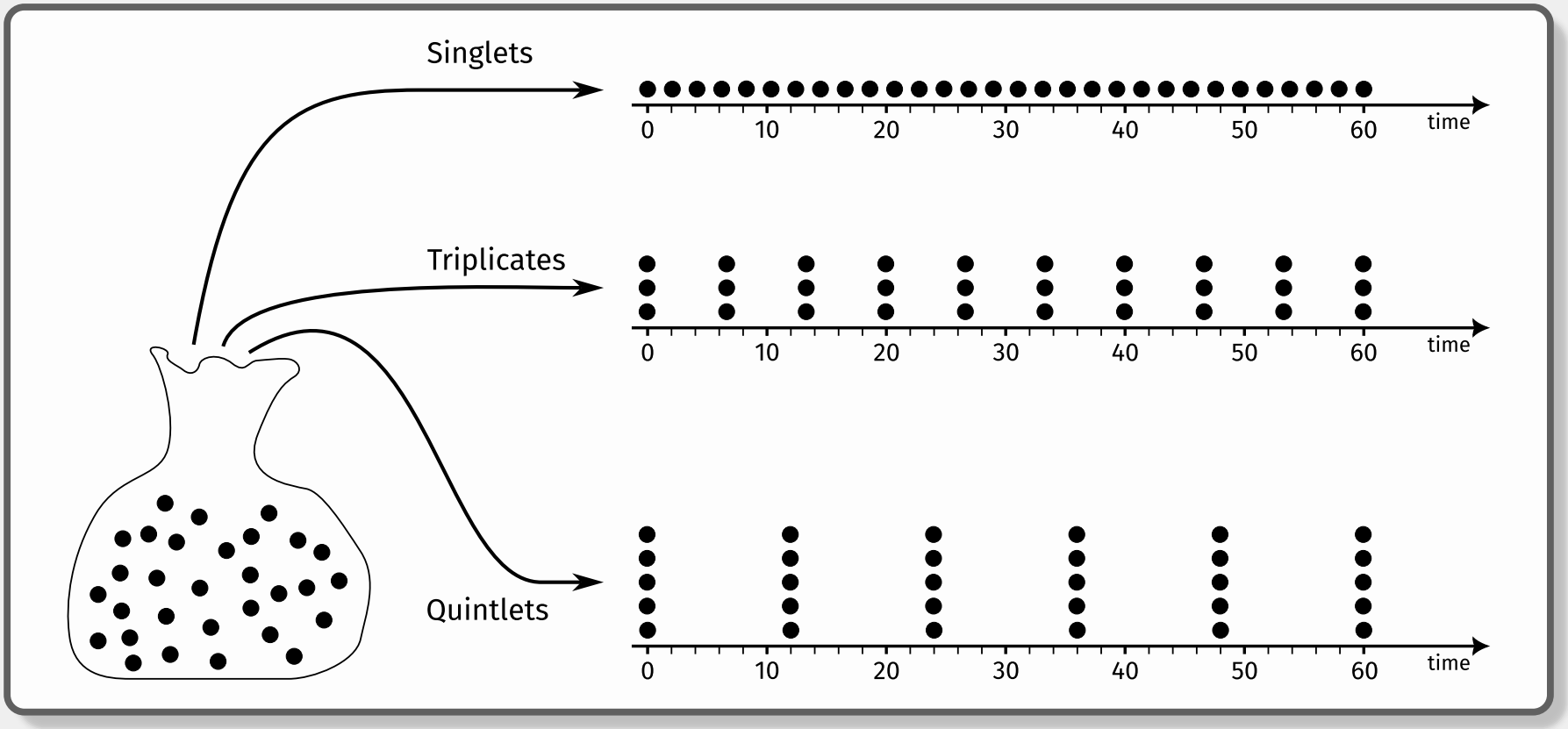
Premises



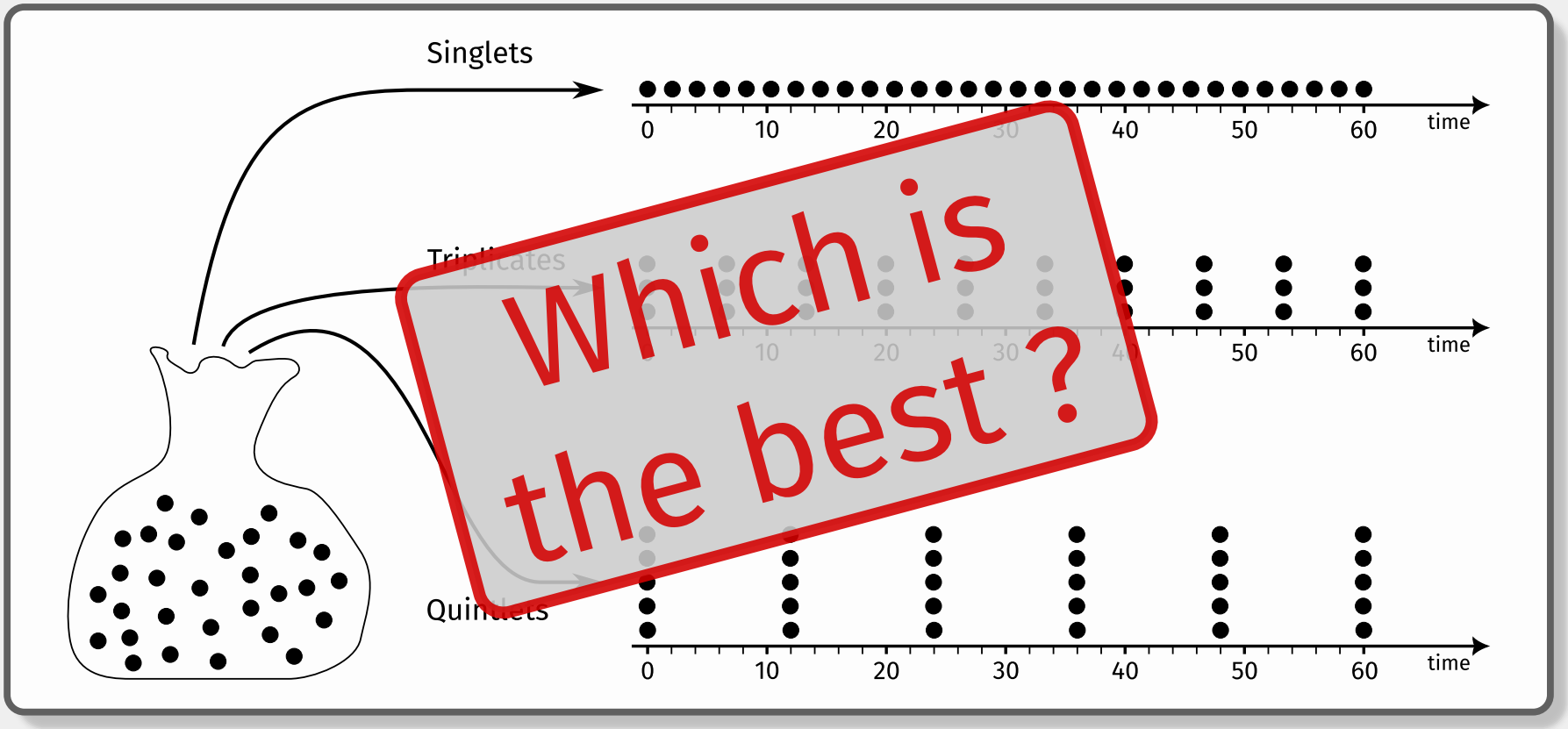
Premises


Singlets
0
10
20
30
40
50
60
time
Triplicates
0
10
20
30
40
50
60
time

Premises



Premises


Singlets
Triplicates
Quintlets
0
10
20
30
40
50
60
time
Which is the best ?

Idea

- Goal: model selection

Idea

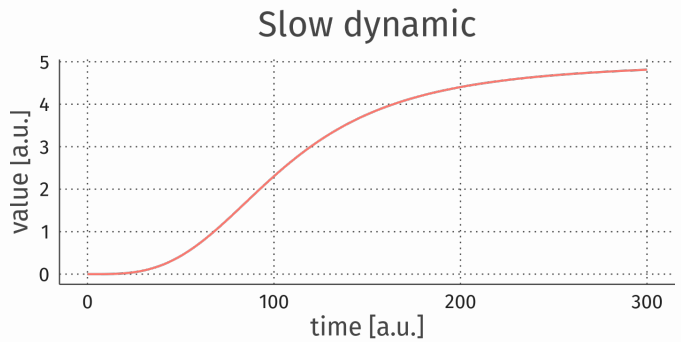
- Goal: model selection
- Fit full and reduced model to same dataset

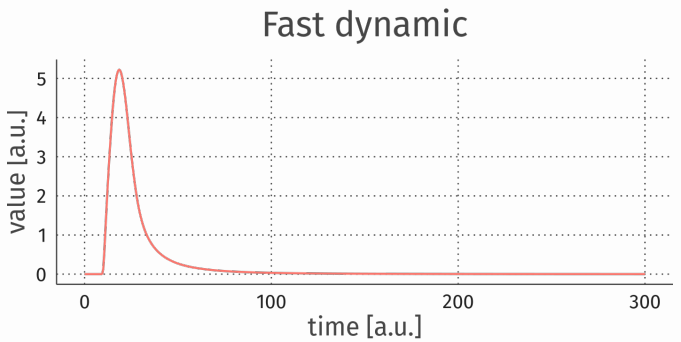
Idea

- Goal: model selection
- Fit full and reduced model to same dataset
- Is the smaller model sufficient to describe the data?

Idea

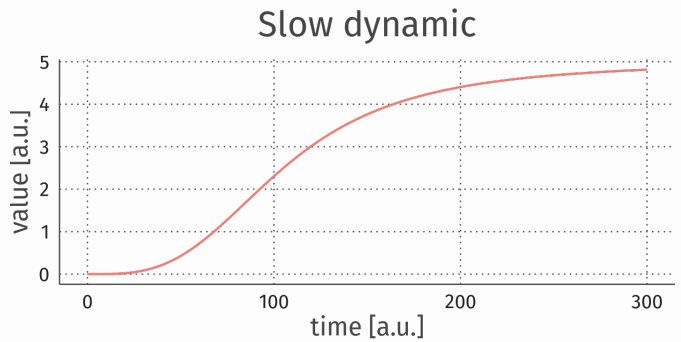
- Goal: model selection
- Fit full and reduced model to same dataset
- Is the smaller model sufficient to describe the data?
- Test on slow and fast dynamic

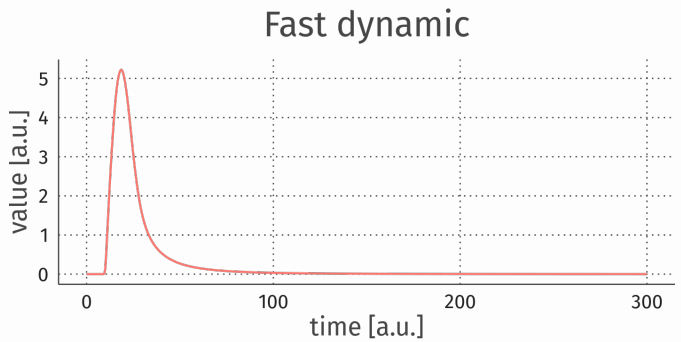




Idea

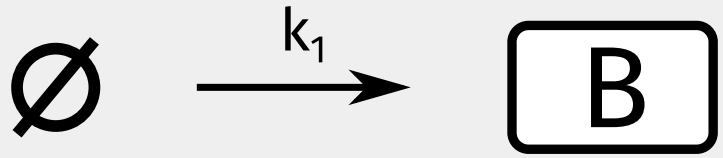
- Goal: model selection
- Fit full and reduced model to same dataset
- Is the smaller model sufficient to describe the data?
- Test on slow and fast dynamic



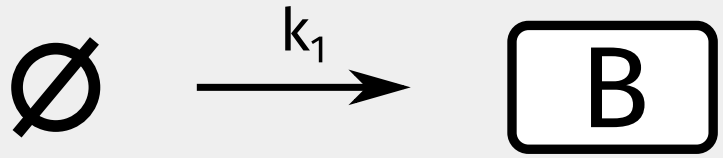


- Additionally: look at different time spans for data acquisition

Slow dynamic

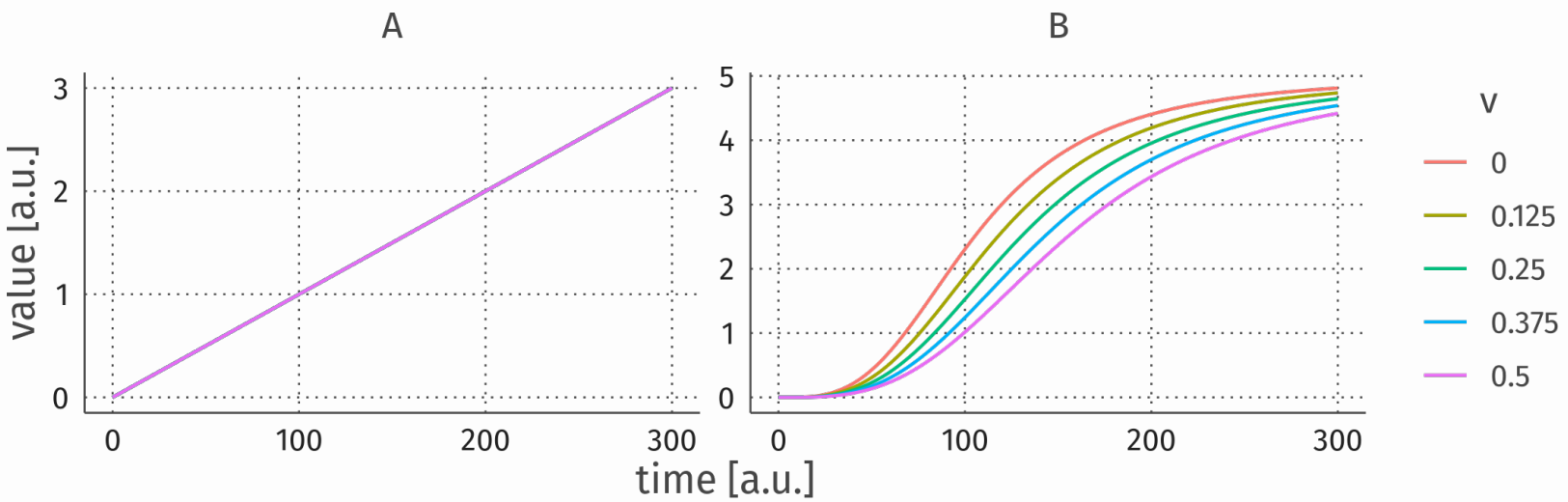


Slow dynamic



- Generation of *B* with Hill type kinetic, a linear *driver A* and Hill coefficient *n*

$$k_1 = \frac{A^n}{(k_M \cdot (1 + v))^n + A^n}$$



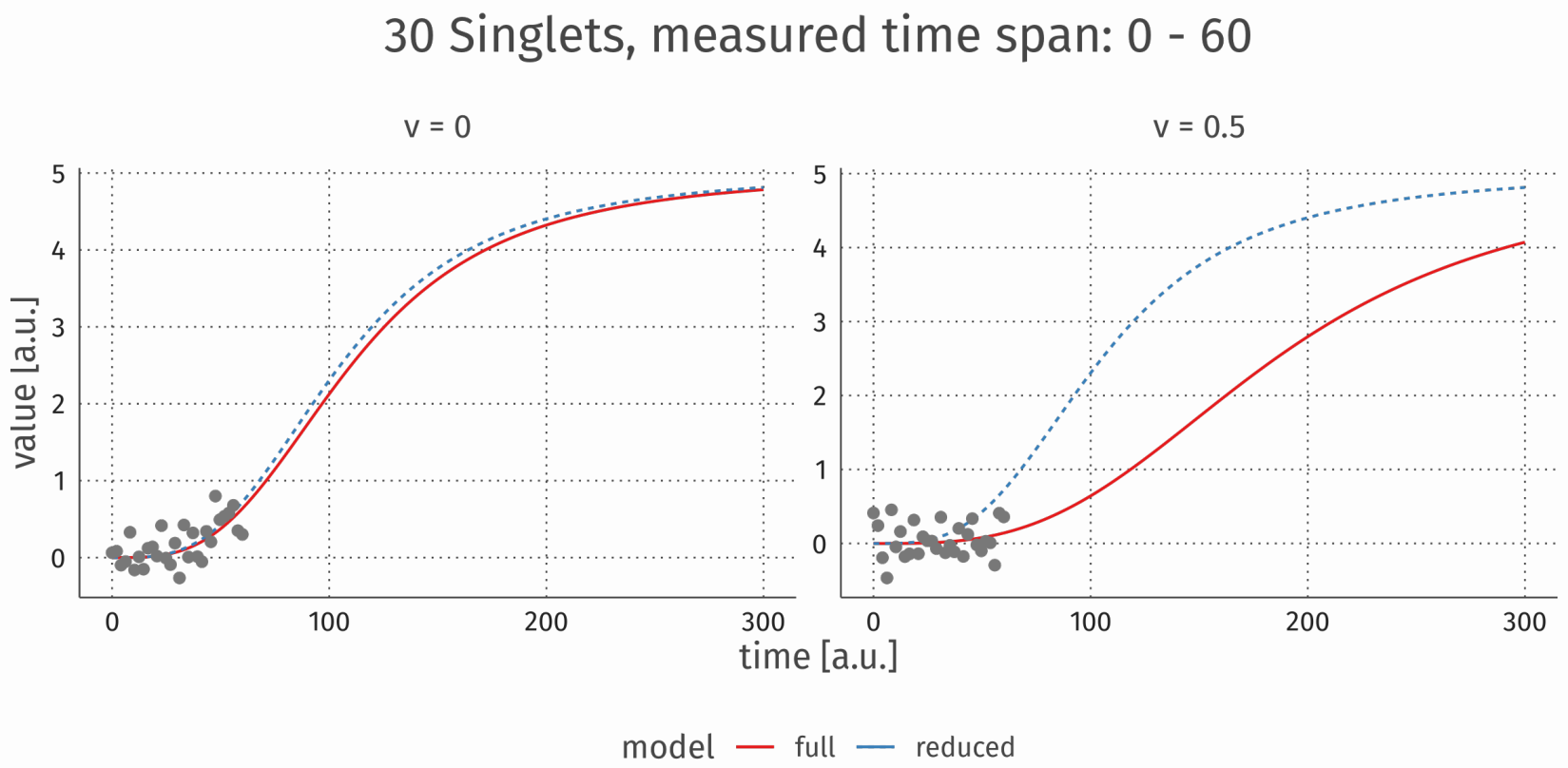
Simulation study

- Full model: k_M free; reduced model: k_M fixed
- Null hypothesis:

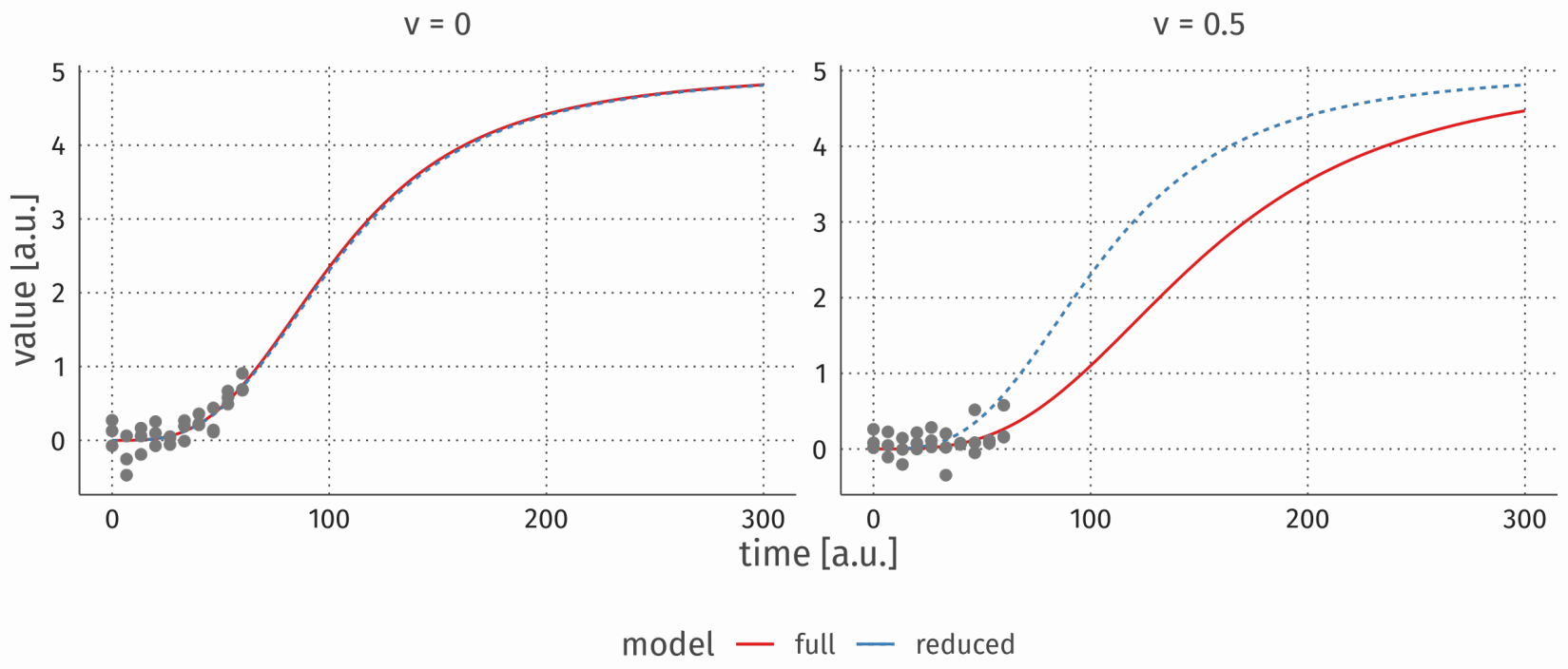
  > A *reduced* model with a fixed k_M describes the data.

- Data is generated under increasing violation of null hypothesis
- Gaussian noise is added to the data
- Fit full and reduced model
- Estimate the error via an error model

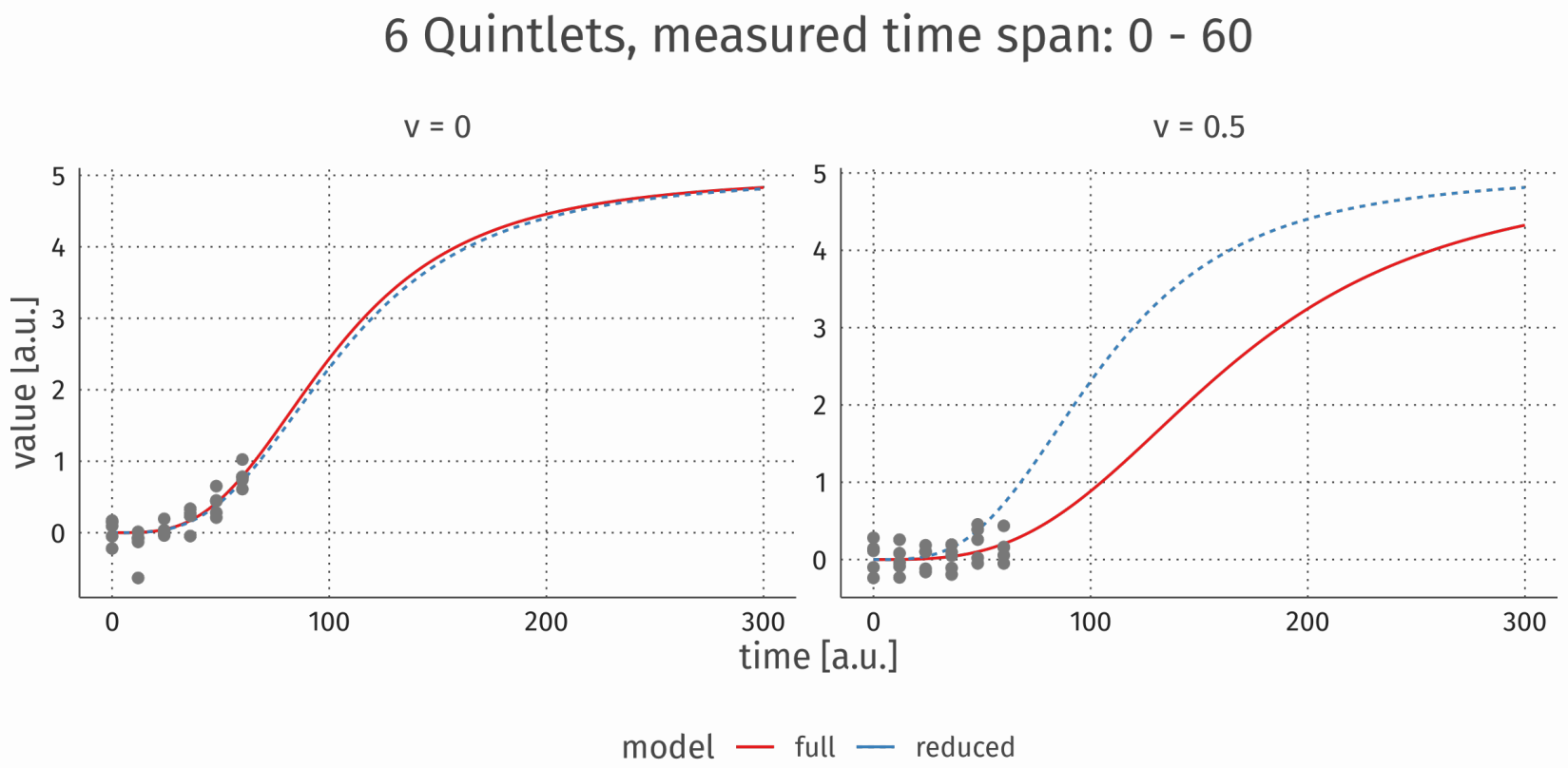
Example – slow dynamic



Example – slow dynamic



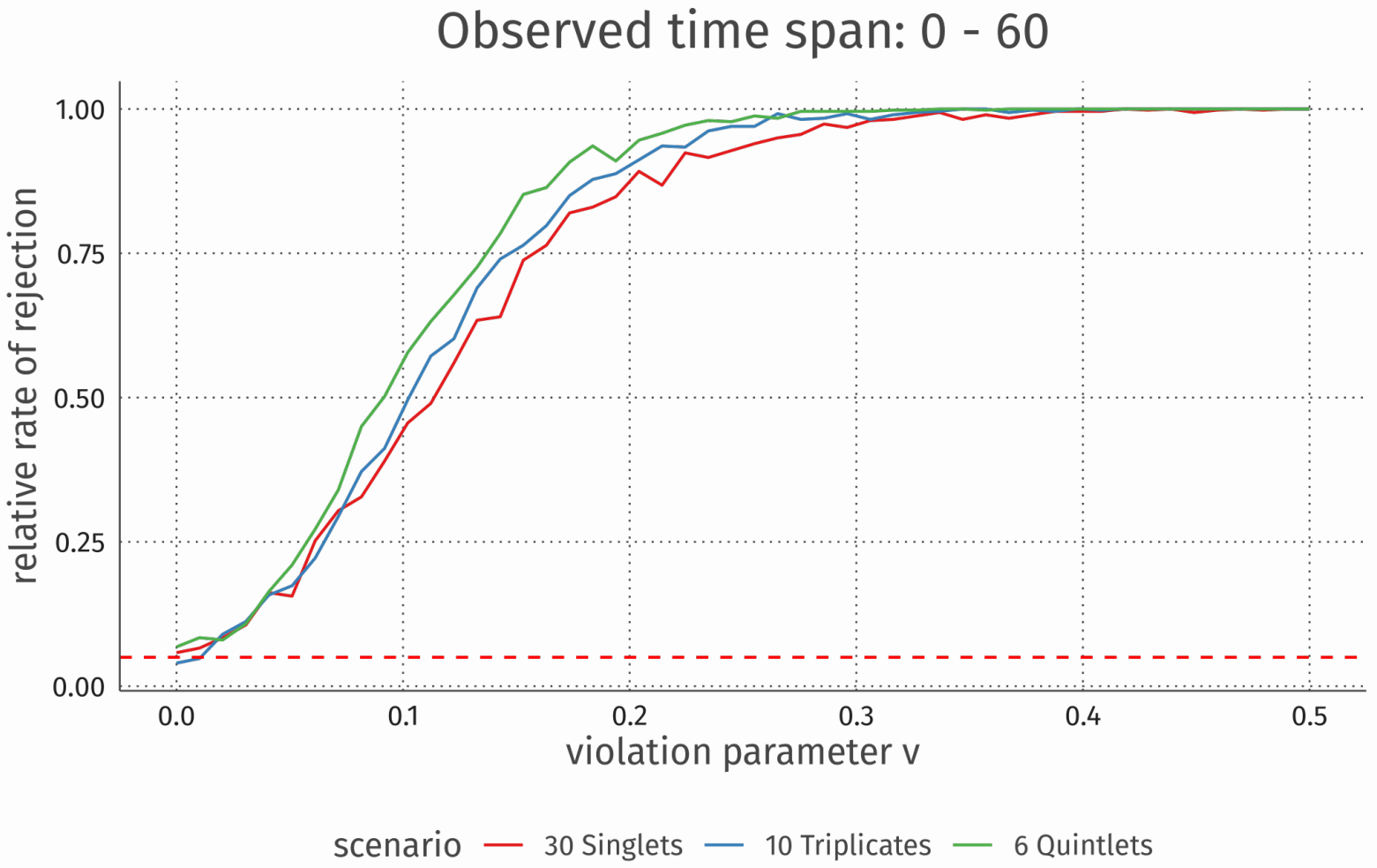
Example – slow dynamic



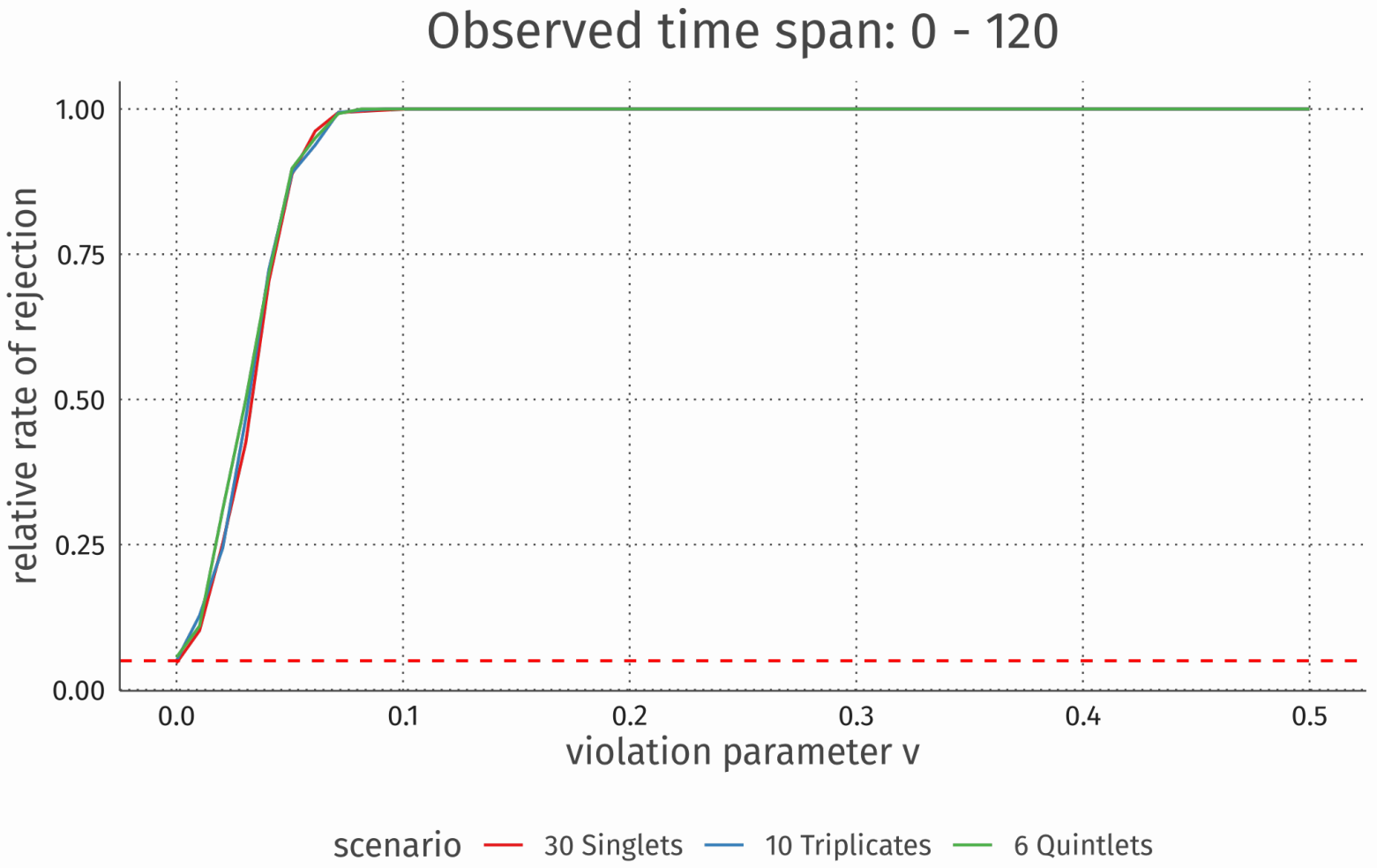
POTT plot

- Accept or reject reduced model based on p-value calculated from log-likelihood ratio between full and reduced model
- Plot the relative rate of rejection, i.e. proportion of realizations with $p < 0.05$ over violation v
- At $v = 0$: rejection rate should represent chosen p-value threshold. If not:
  - Less rejected: *conservative*
  - More rejected: *anti-conservative*

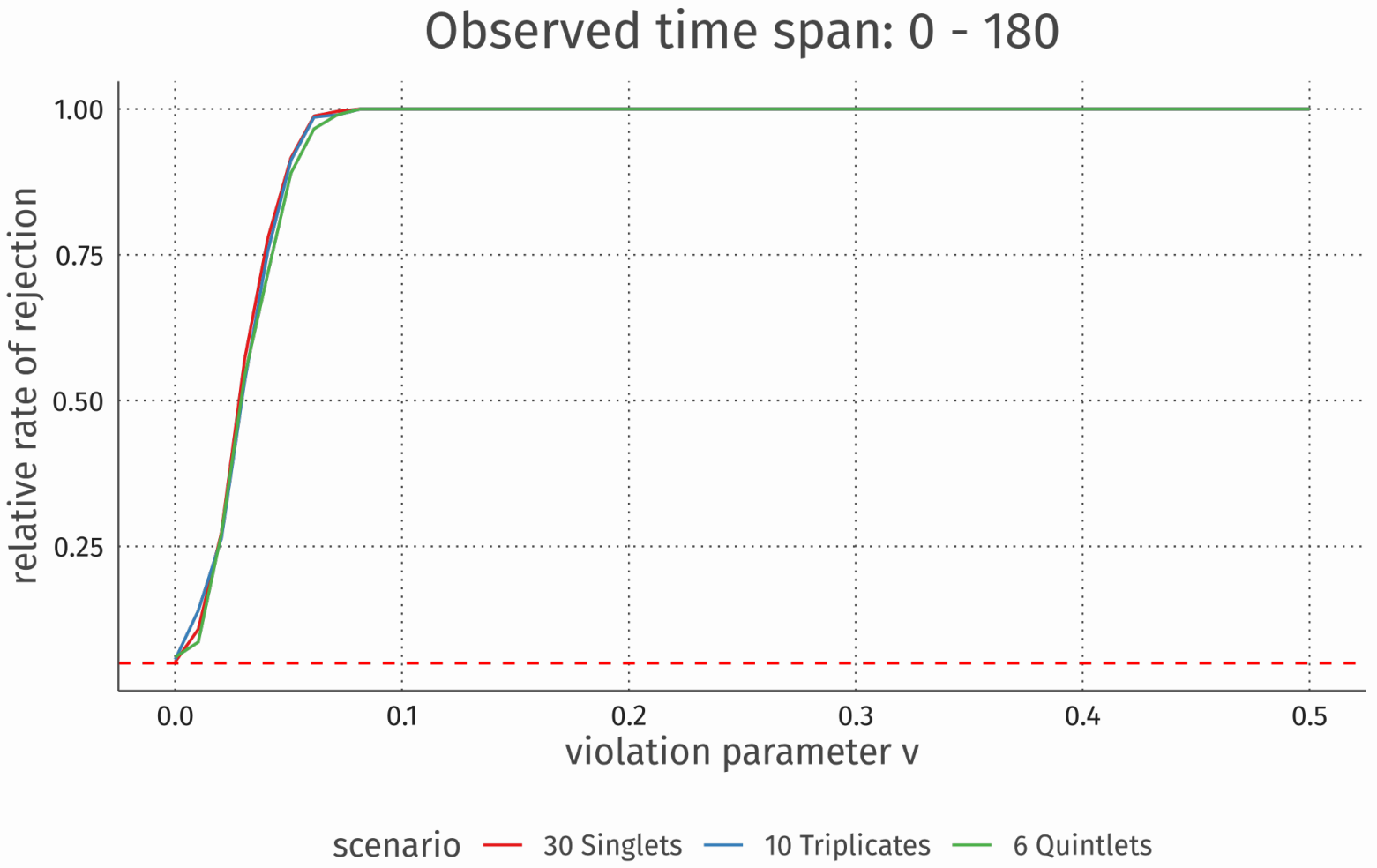
POTT-PLOT – SLOW DYNAMIC



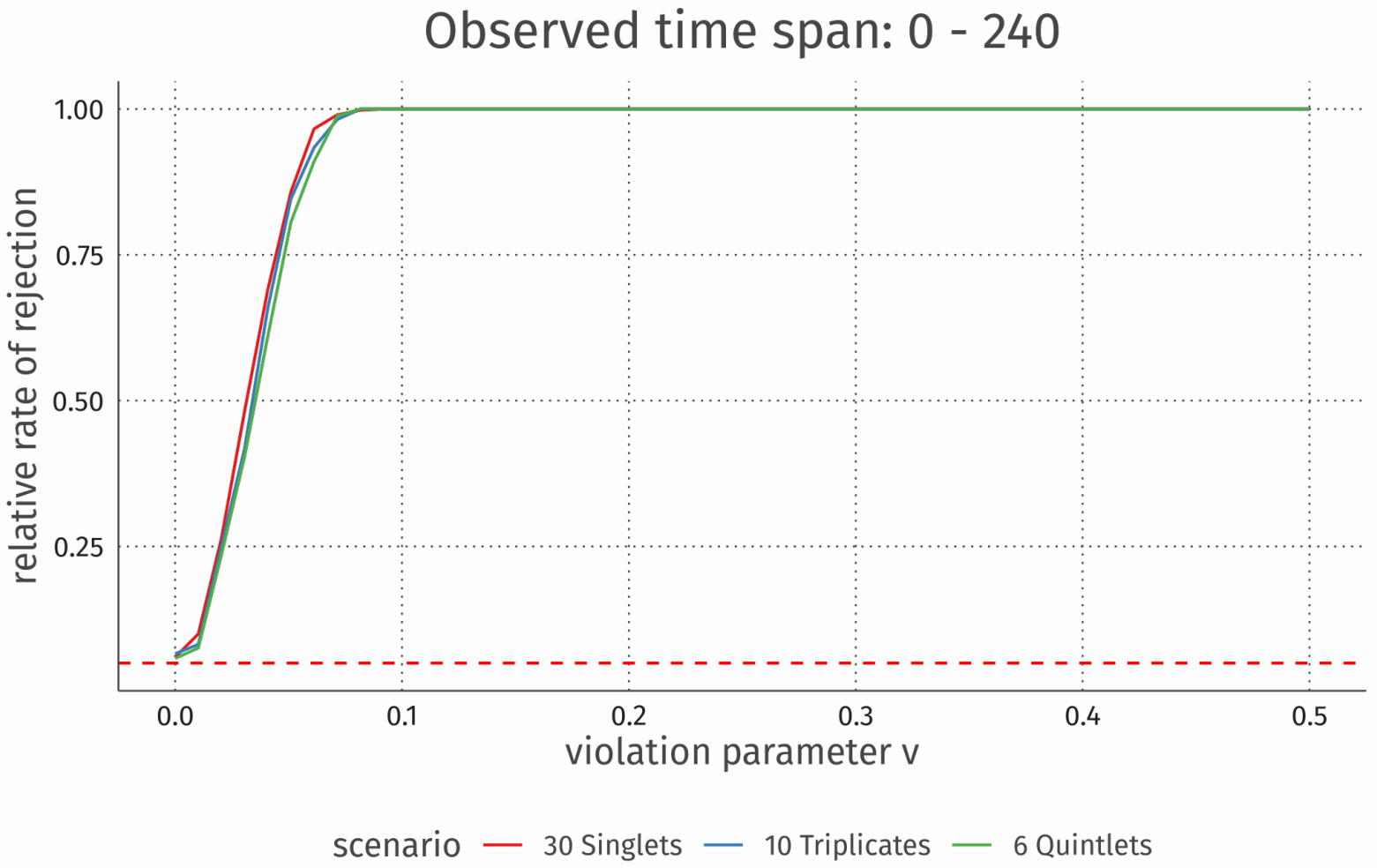
POTT-PLOT – SLOW DYNAMIC



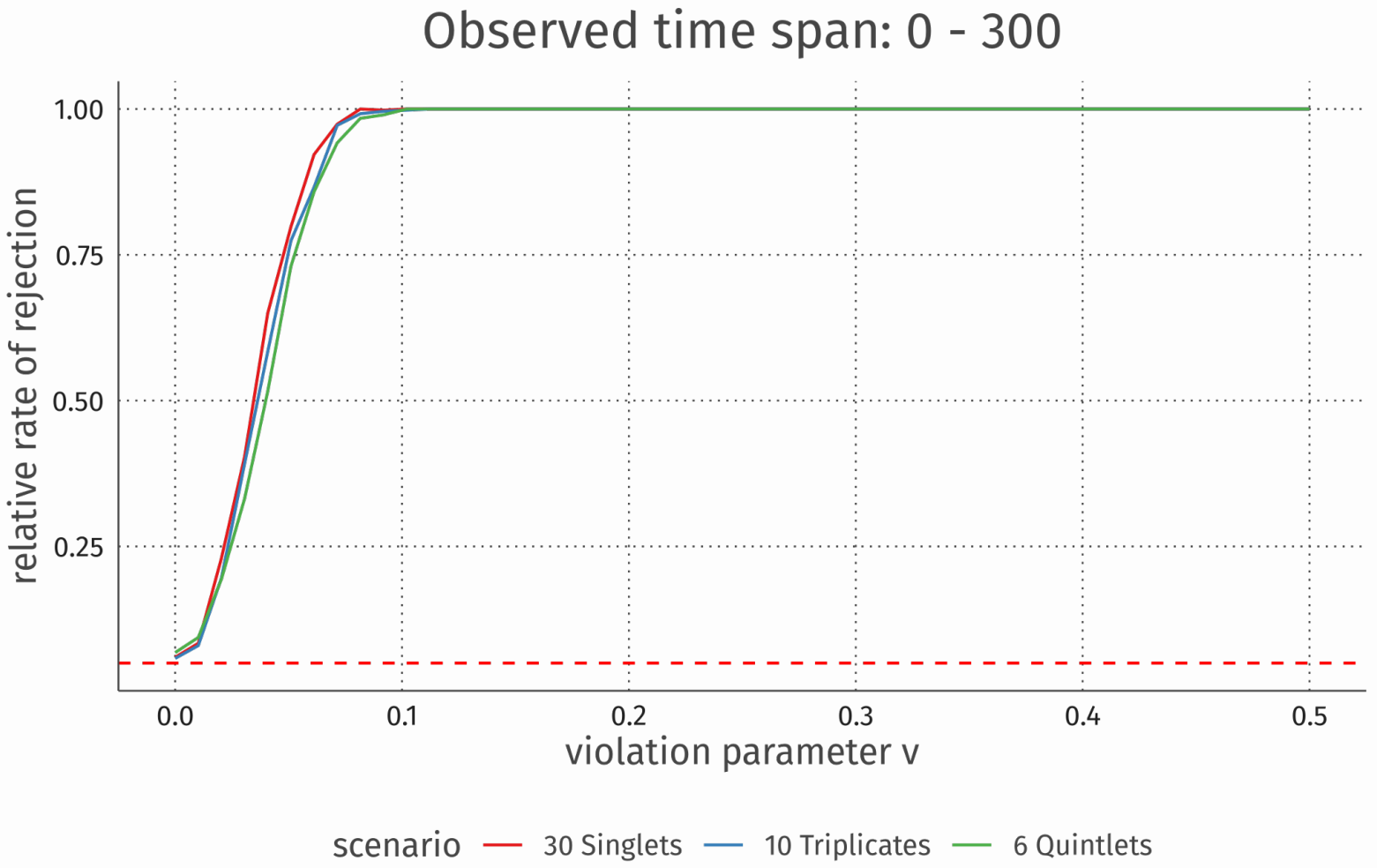
POTT-PLOT – SLOW DYNAMIC



POTT-PLOT – SLOW DYNAMIC

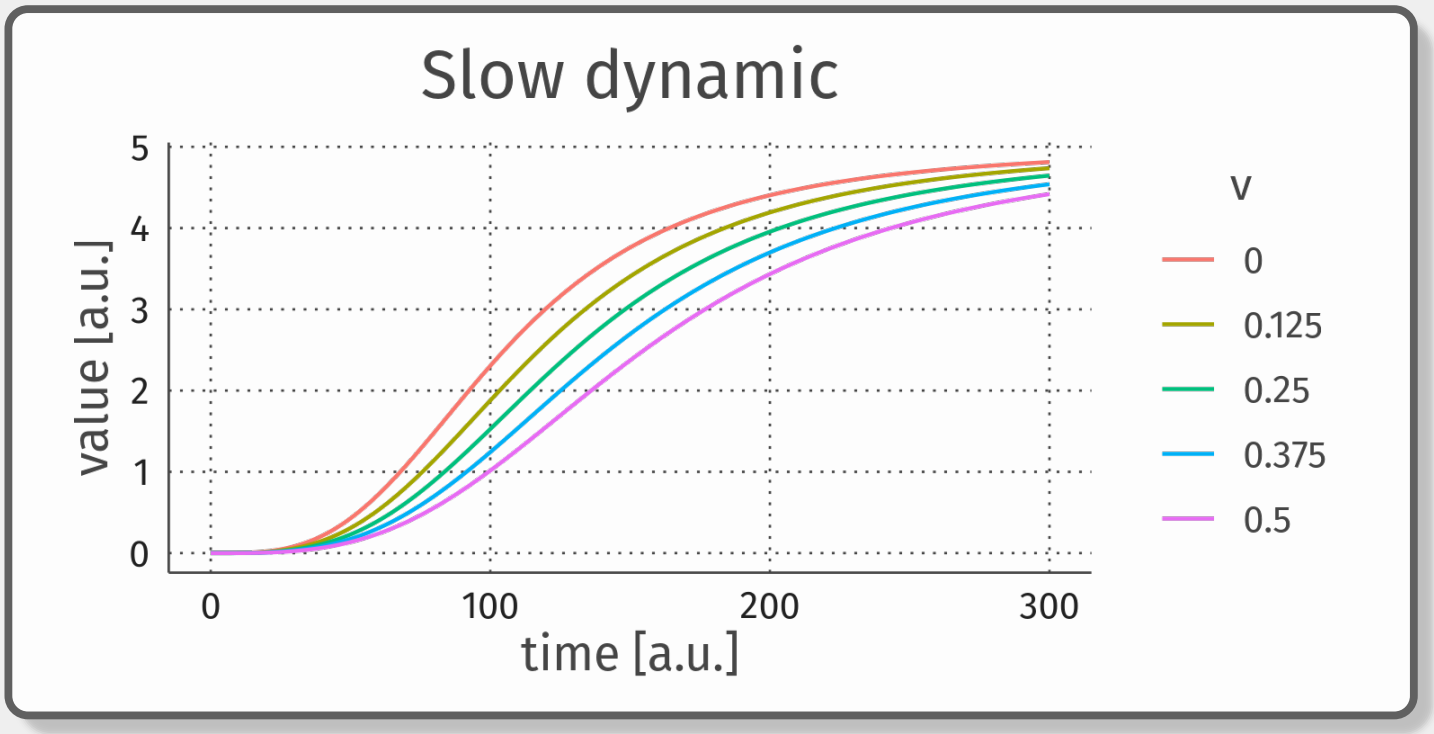


POTT-PLOT – SLOW DYNAMIC



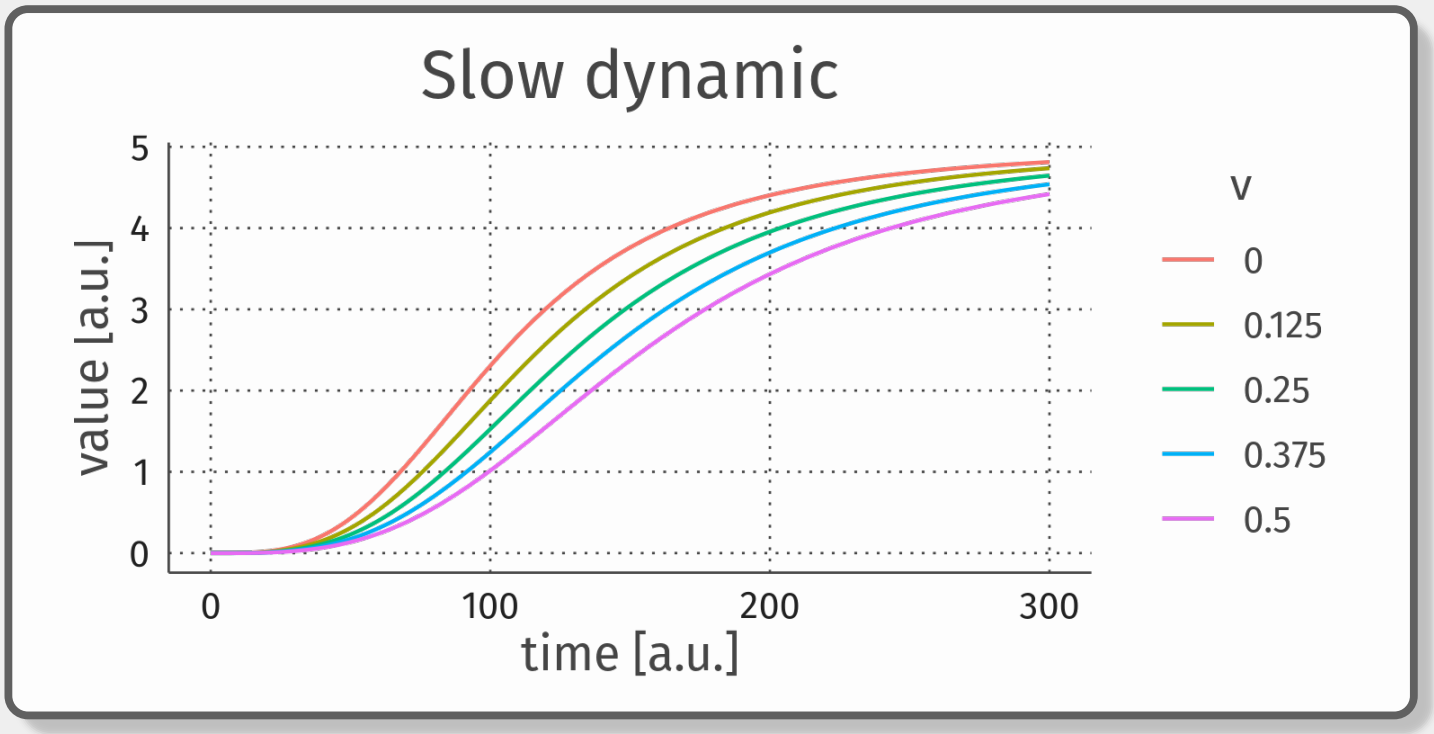
Insights/Recap

- Slow dynamic: only little change up until 60



Insights/Recap

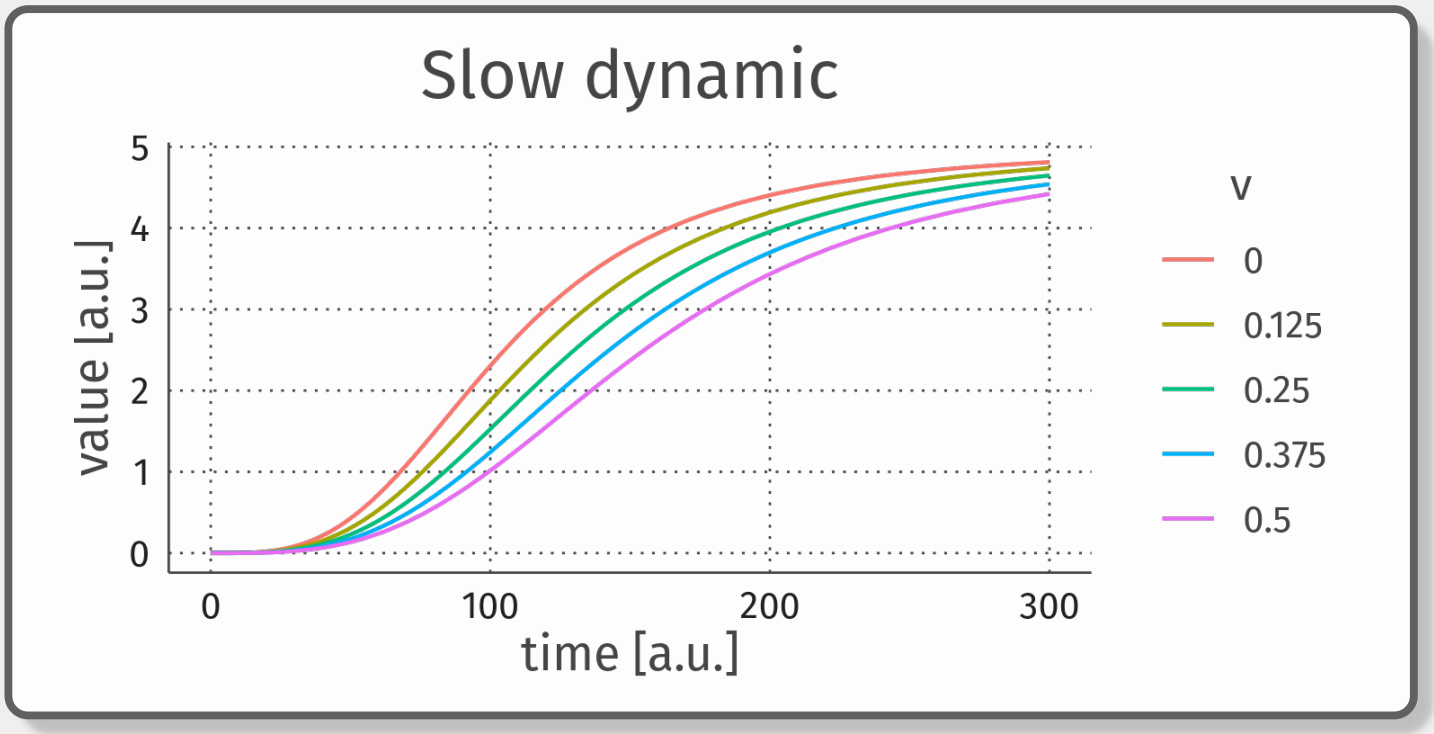
- Slow dynamic: only little change up until 60



- Acquisition schemes make basically no difference

Insights/Recap

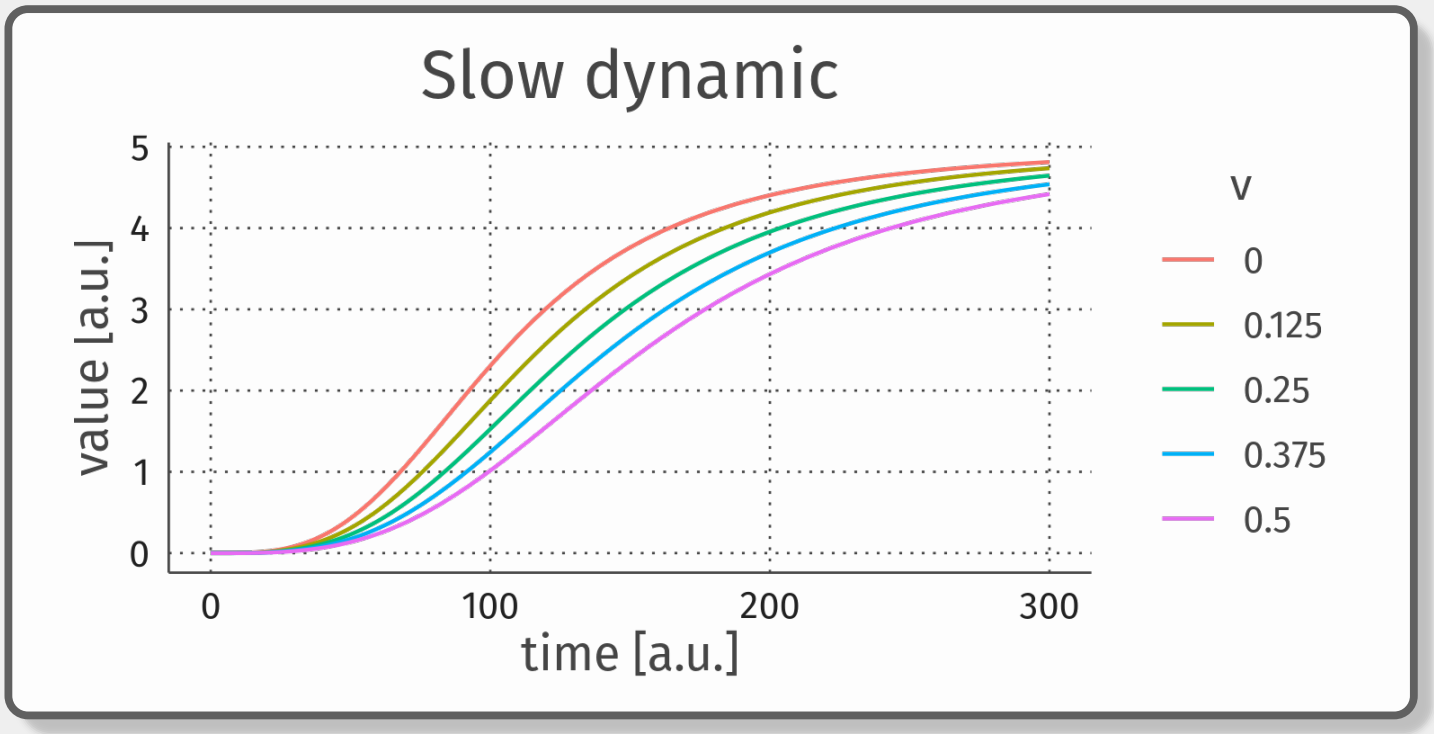
- Slow dynamic: only little change up until 60



- Acquisition schemes make basically no difference
- Violation manifests as offset

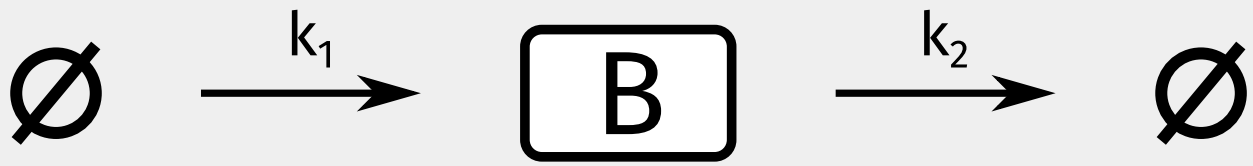
Insights/Recap

- Slow dynamic: only little change up until 60

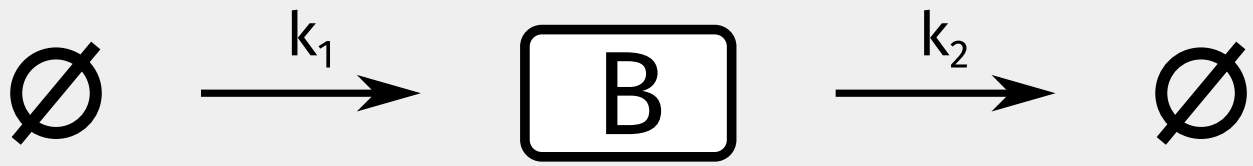


- Acquisition schemes make basically no difference
- Violation manifests as offset
- ...so look at the fast dynamic

Fast dynamic

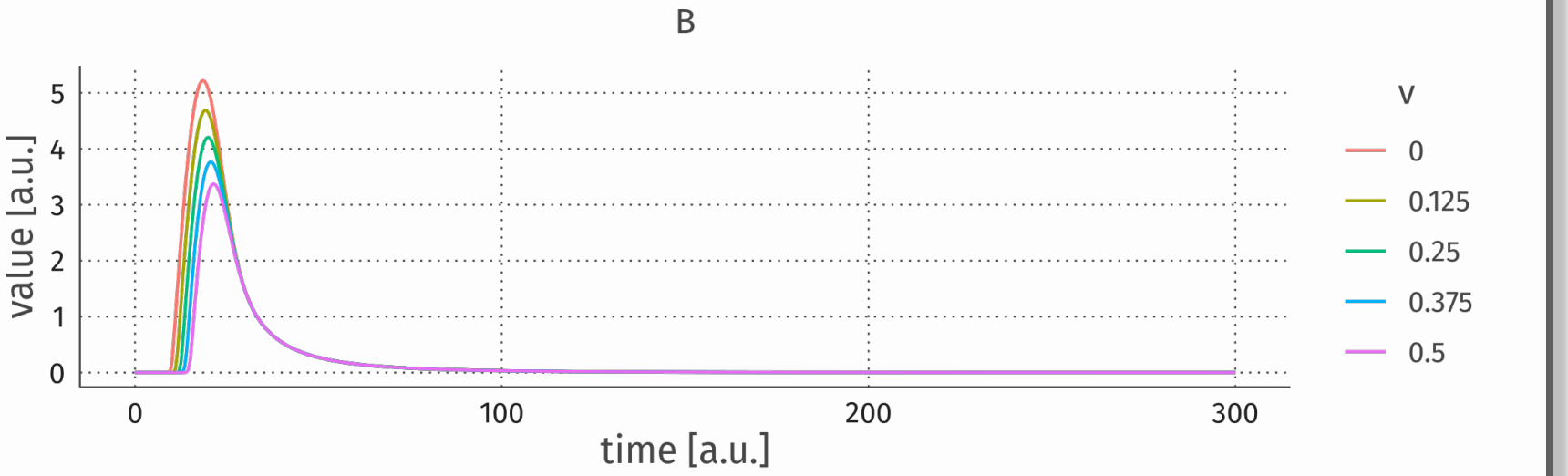


Fast dynamic

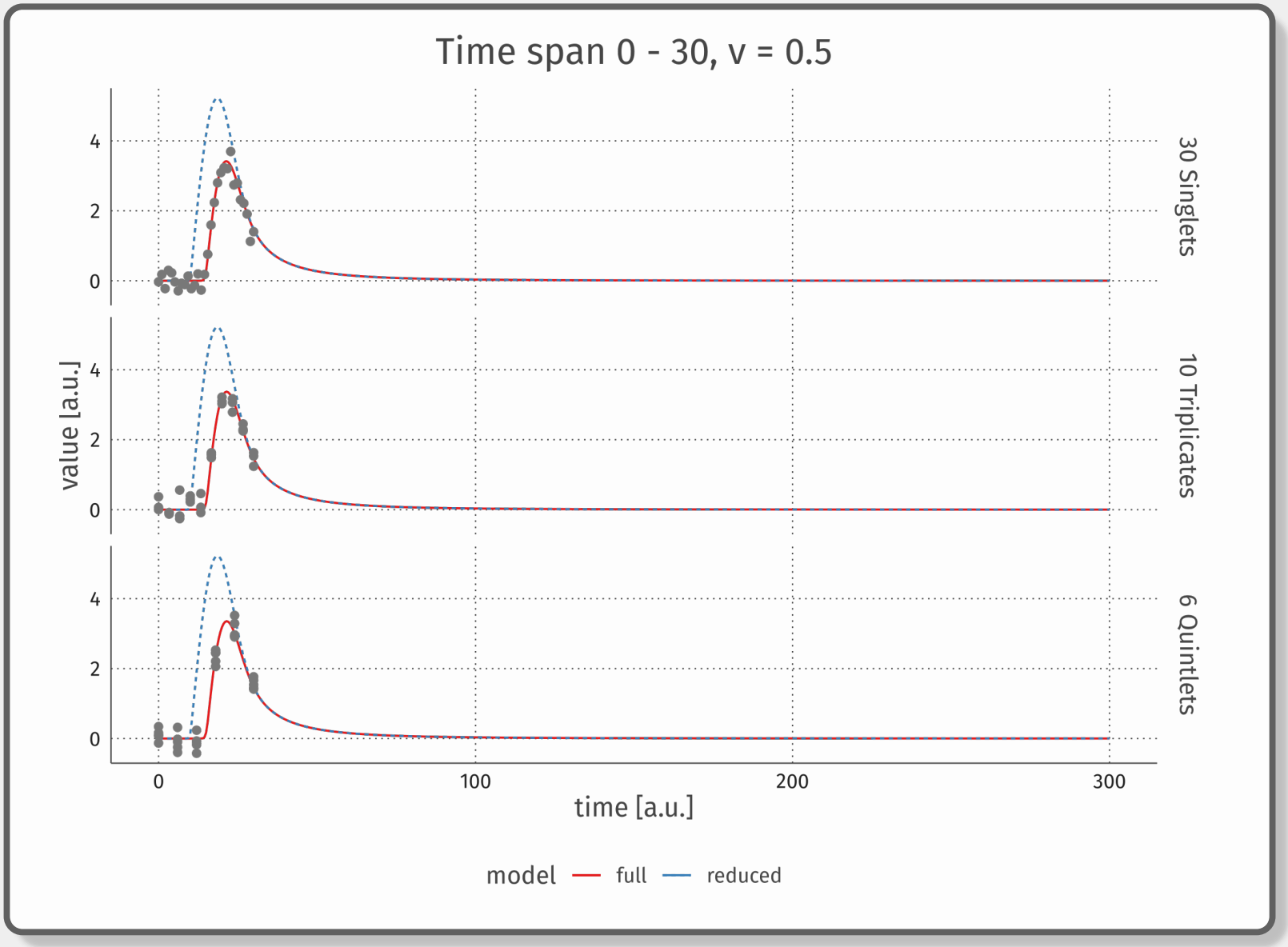


- Generation of *B* with Hill type kinetic, a linear *driver A* and cubic degradation:

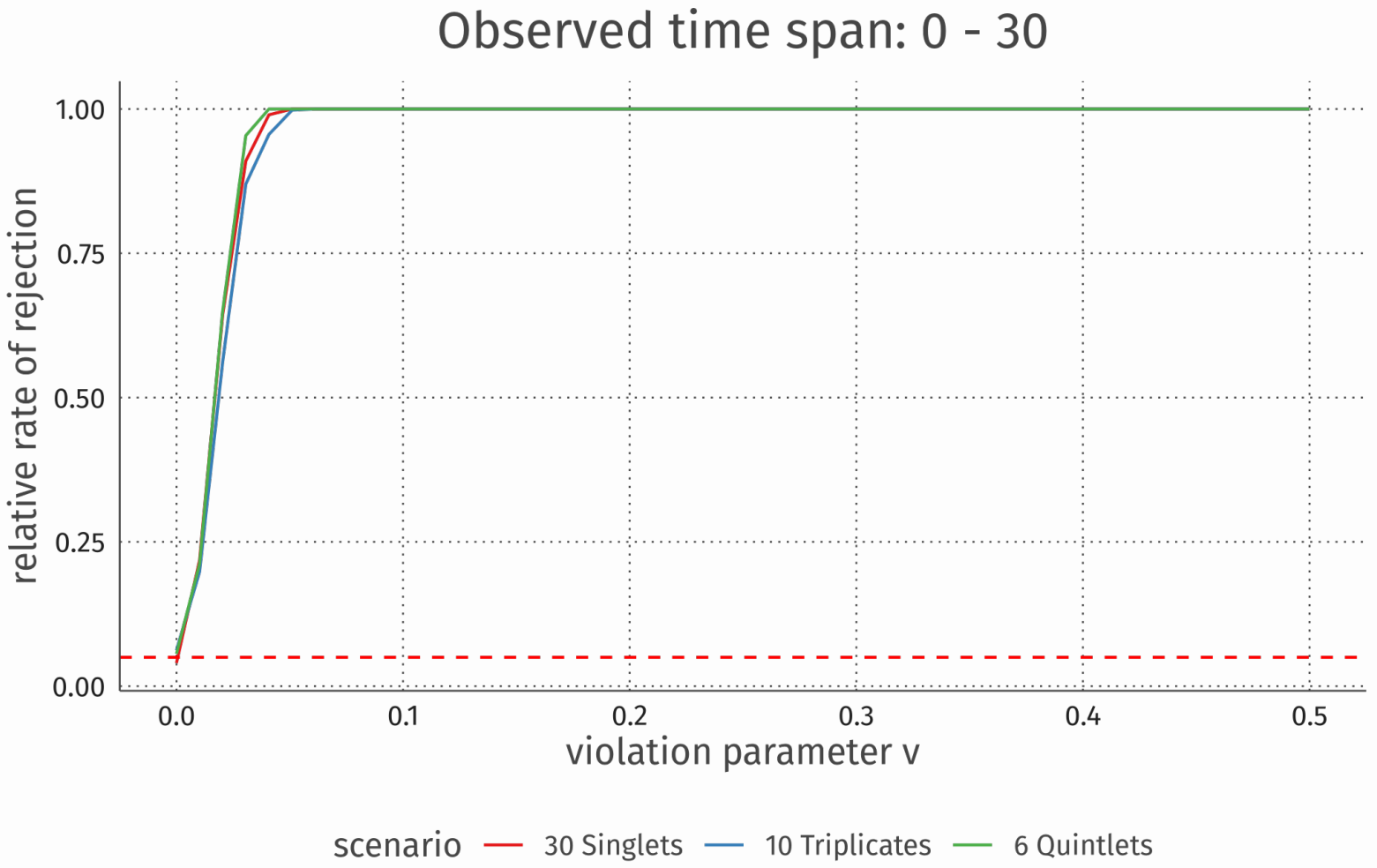
$$k_1 = \frac{A^n}{(k_M \cdot (1 + v))^n + A^n}, \quad k_2 = B \cdot k_3 \cdot A^3$$



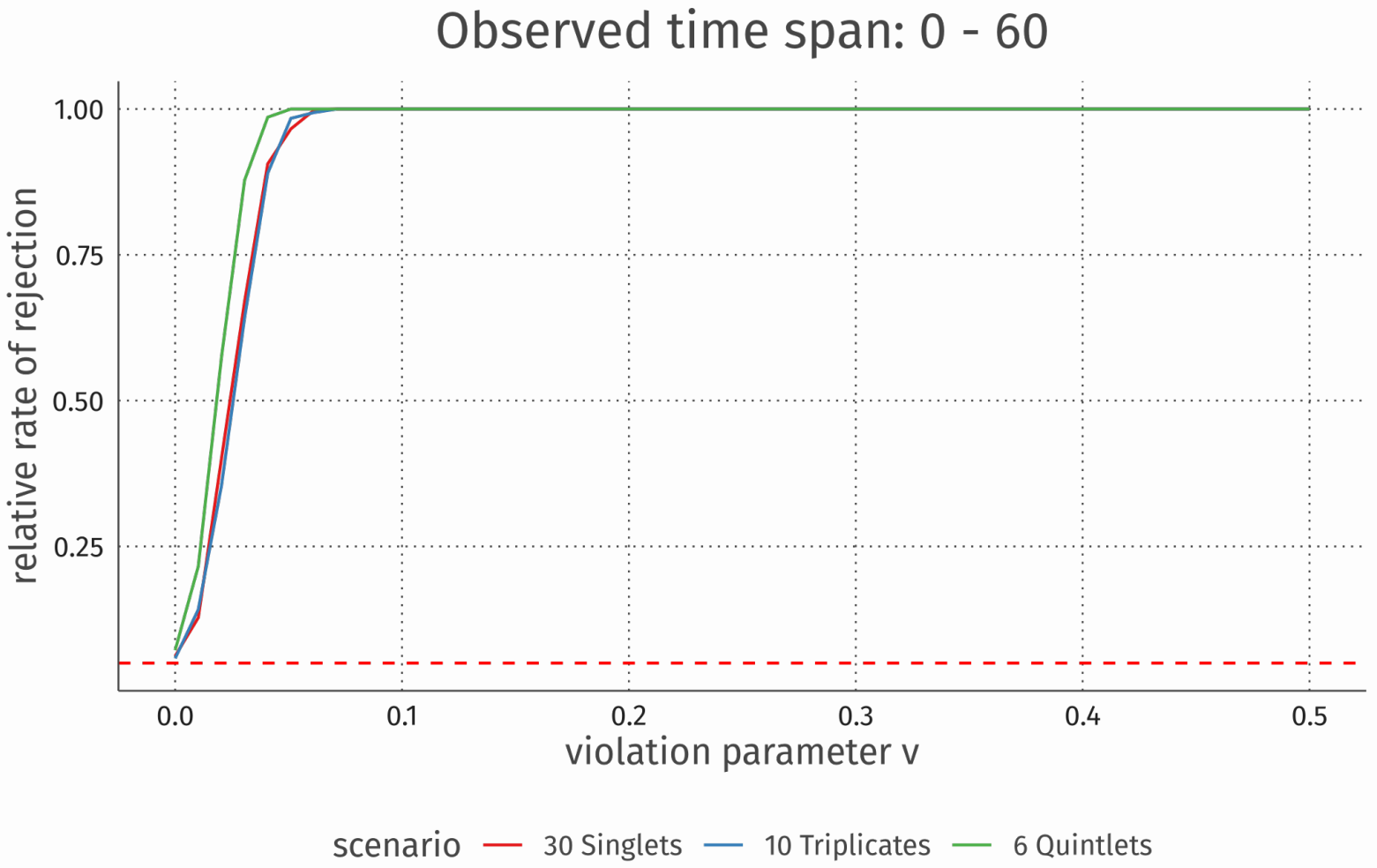
Example – fast dynamic



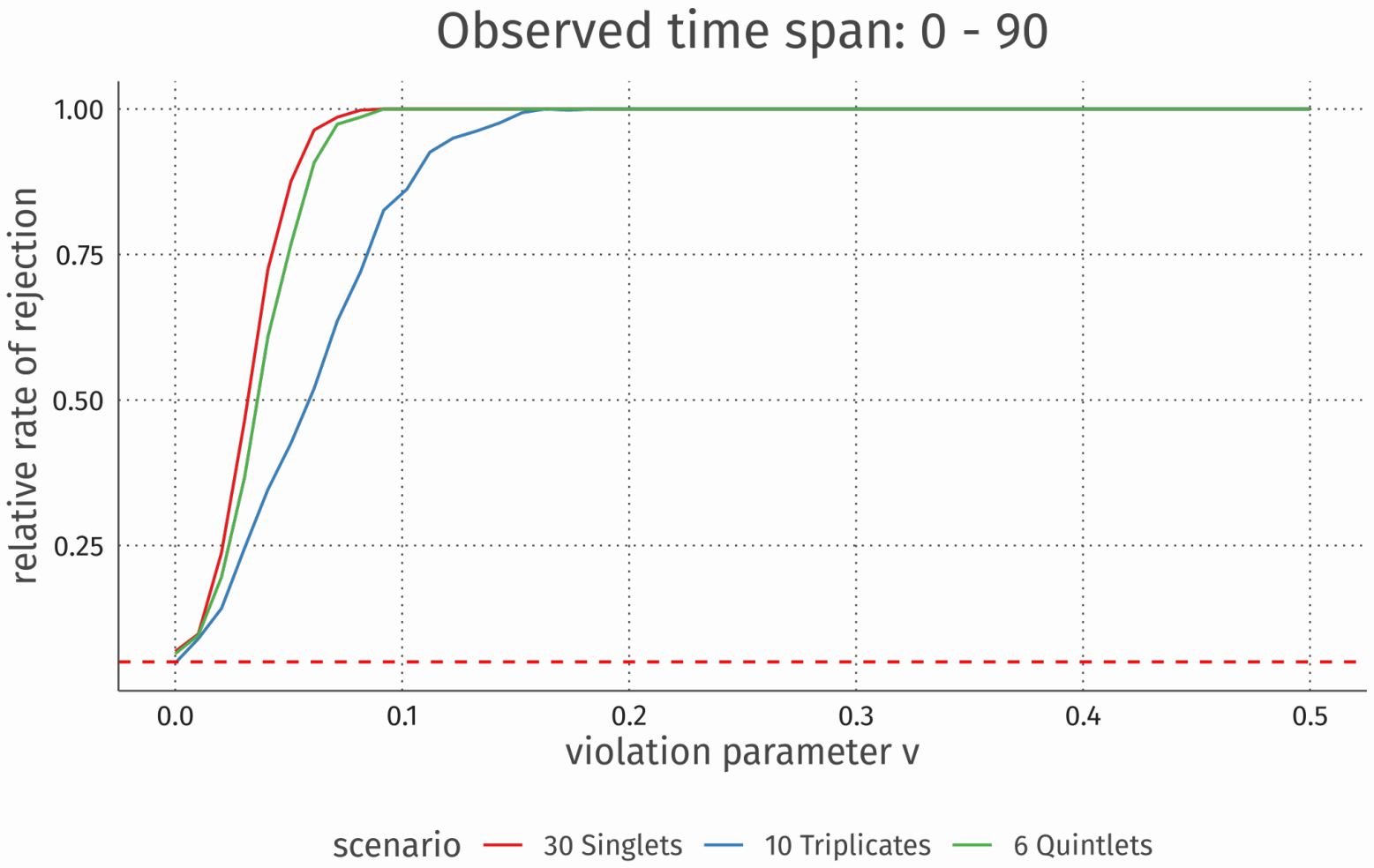
POTT-PLOT – FAST DYNAMIC



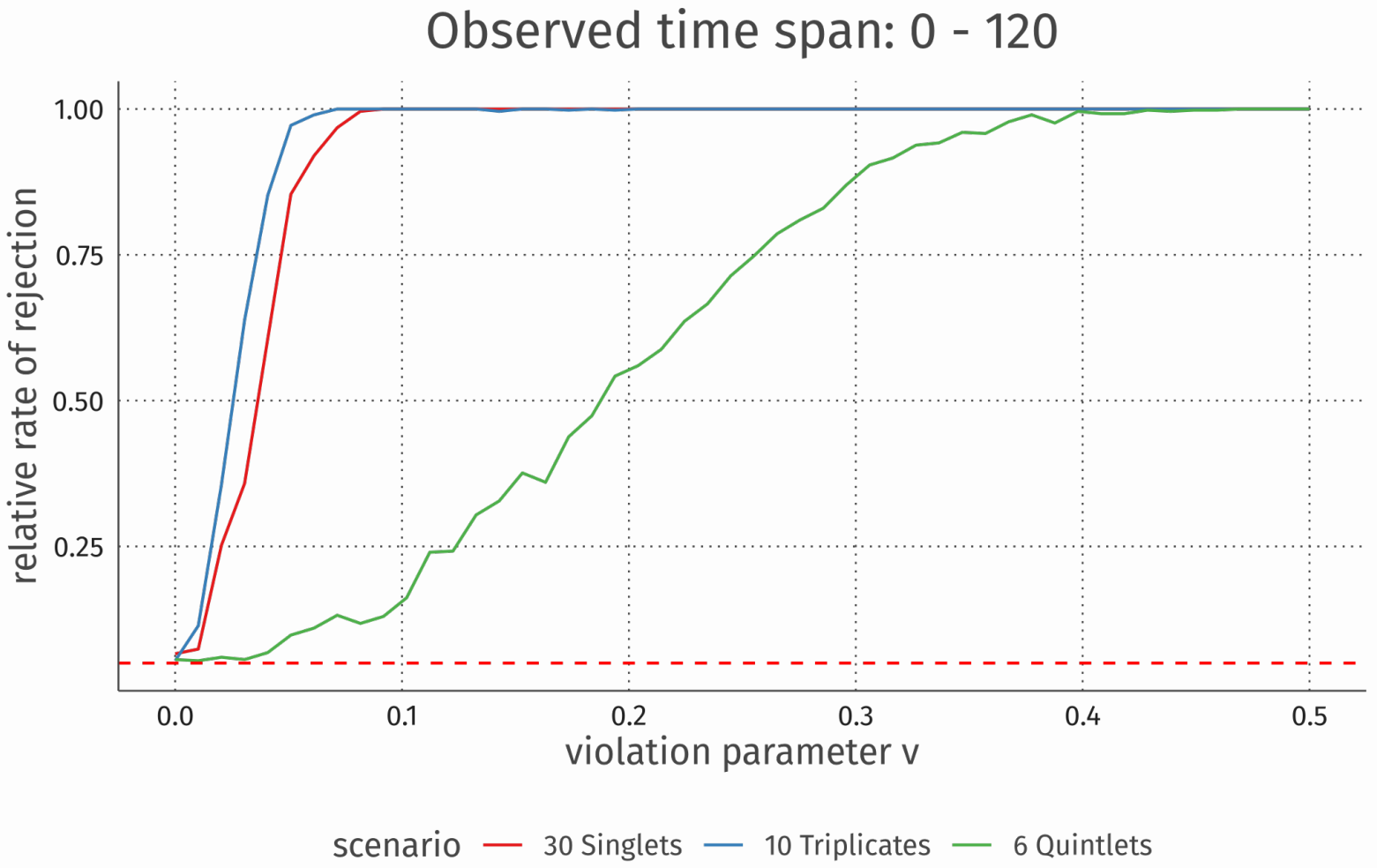
POTT-PLOT – FAST DYNAMIC



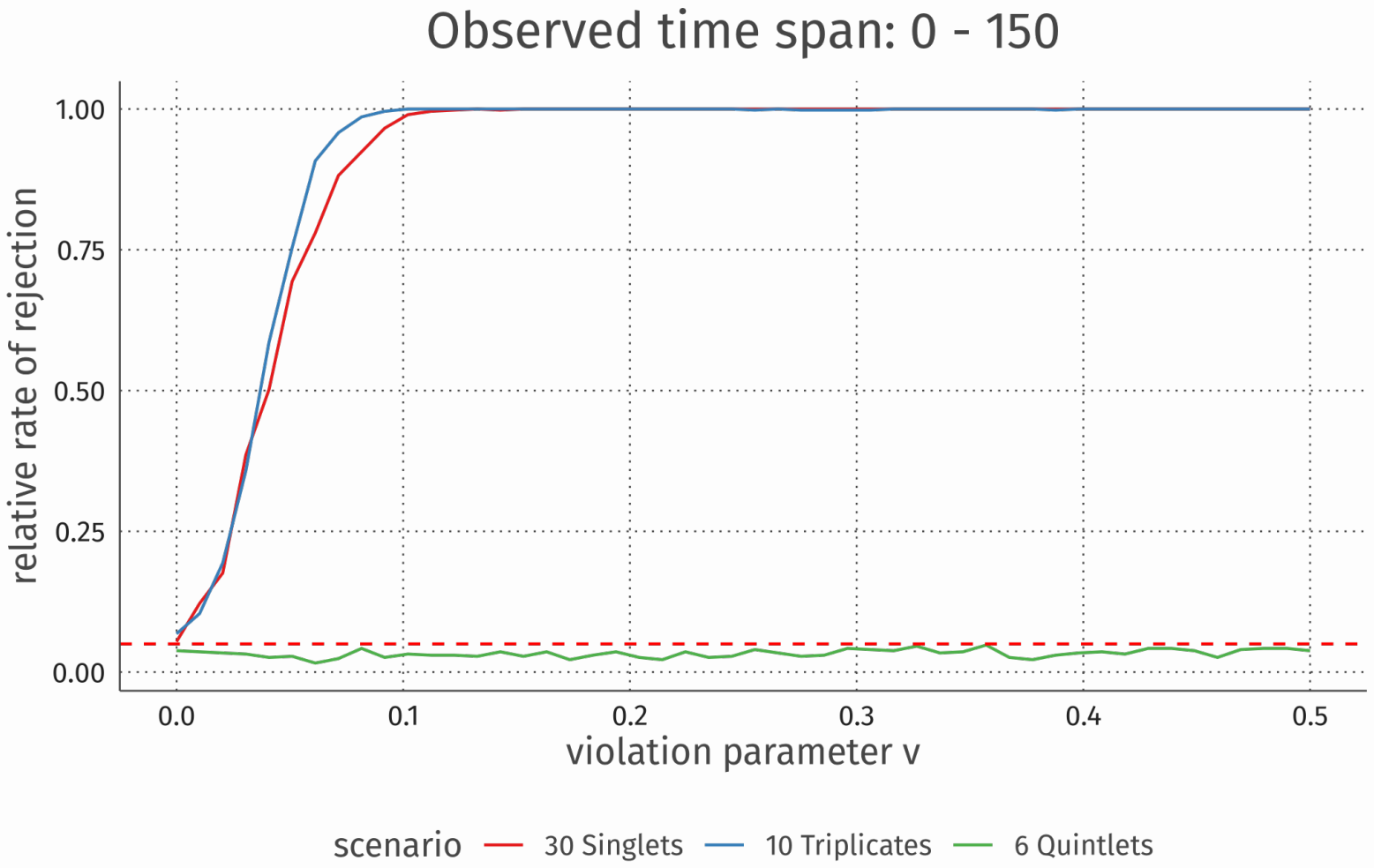
POTT-PLOT – FAST DYNAMIC



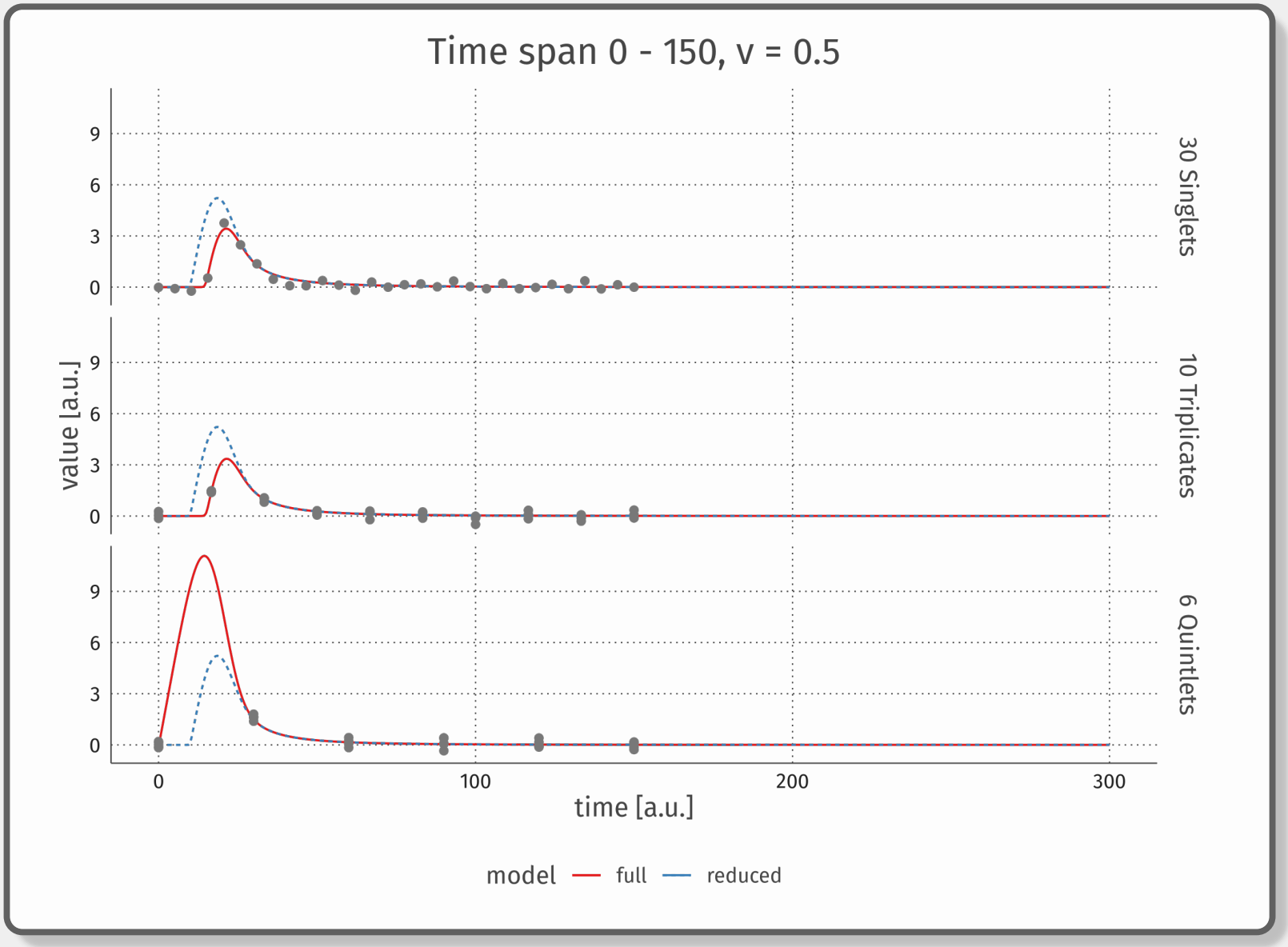
POTT-PLOT – FAST DYNAMIC



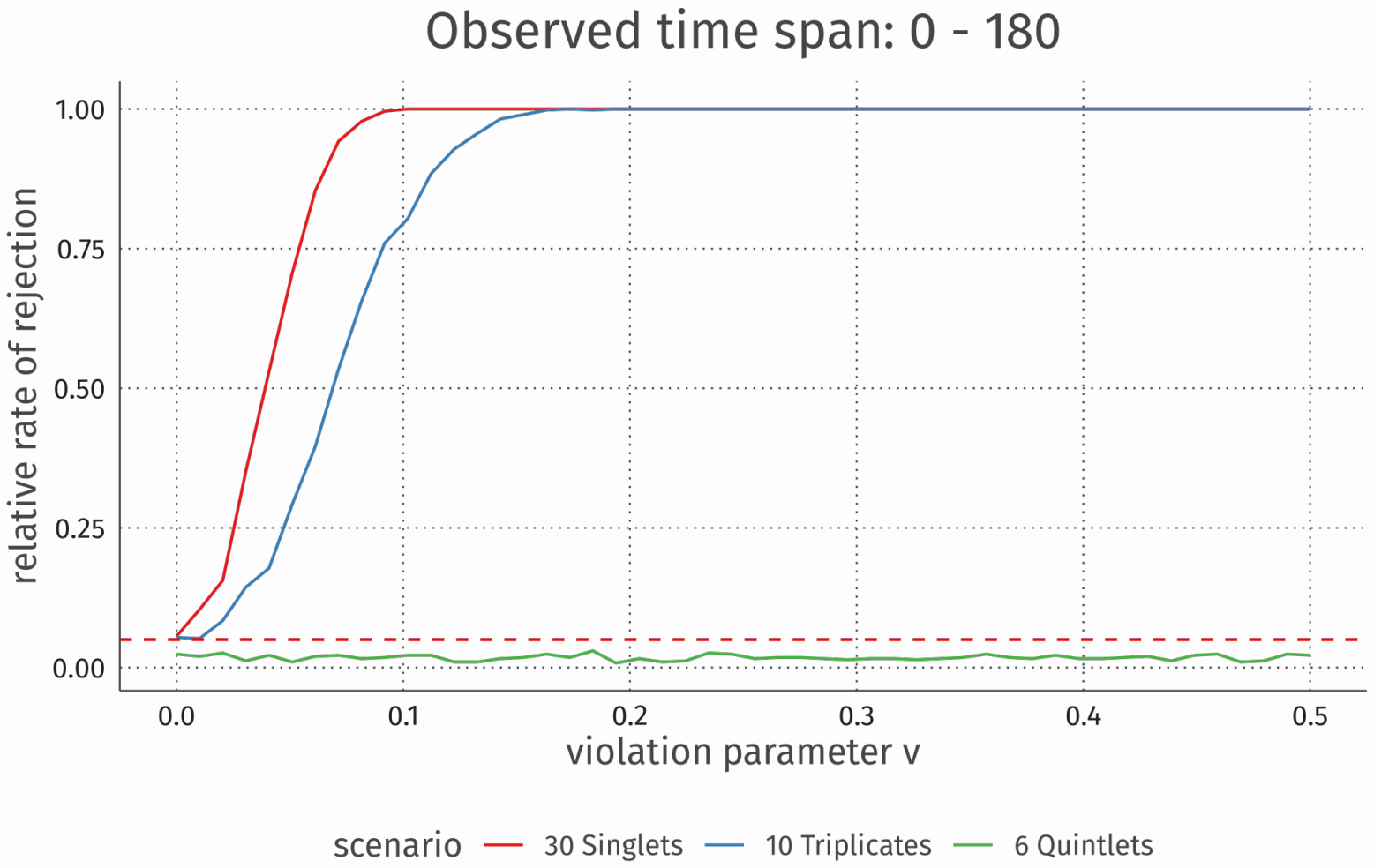
POTT-PLOT – FAST DYNAMIC



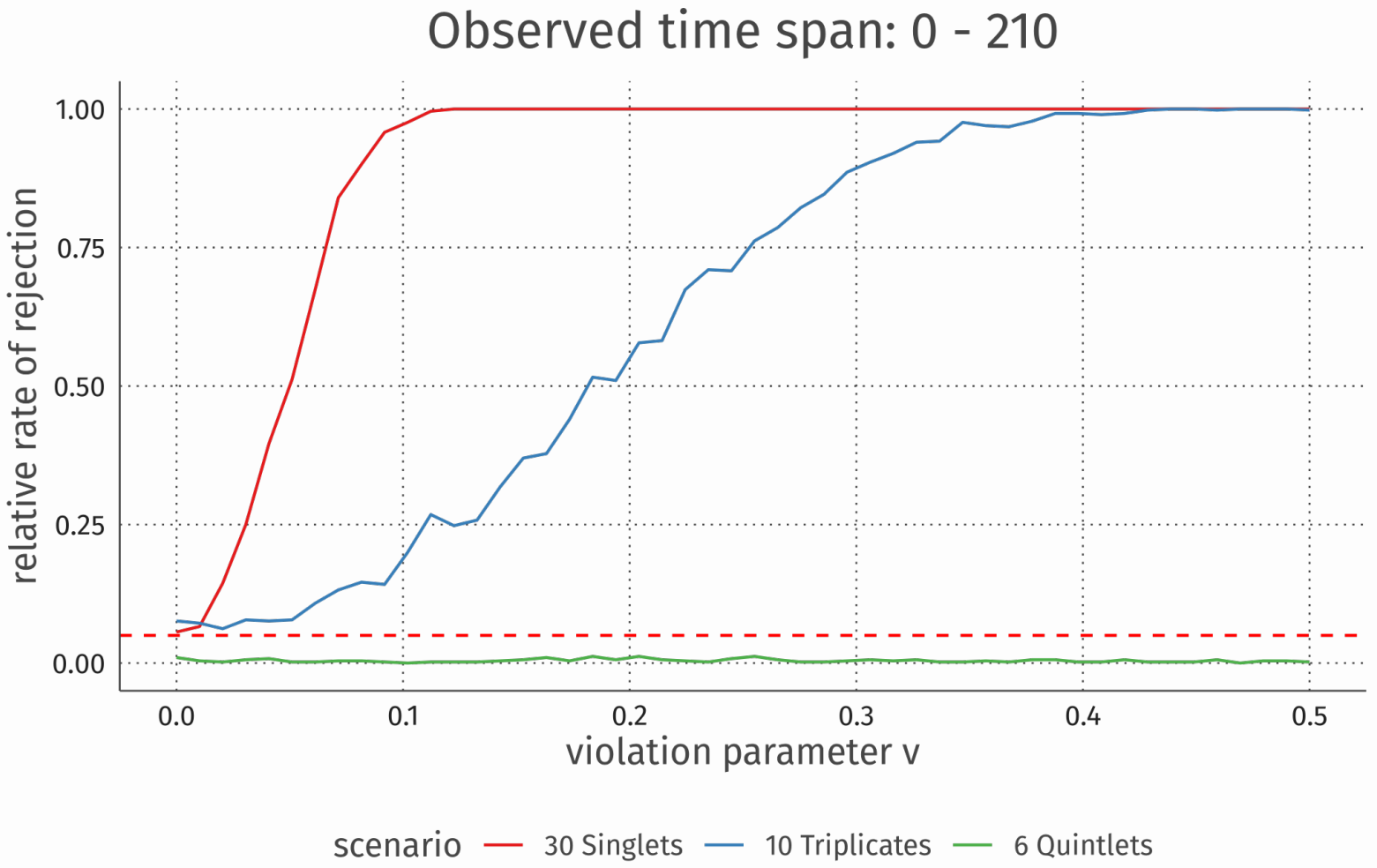
POTT-PLOT – FAST DYNAMIC



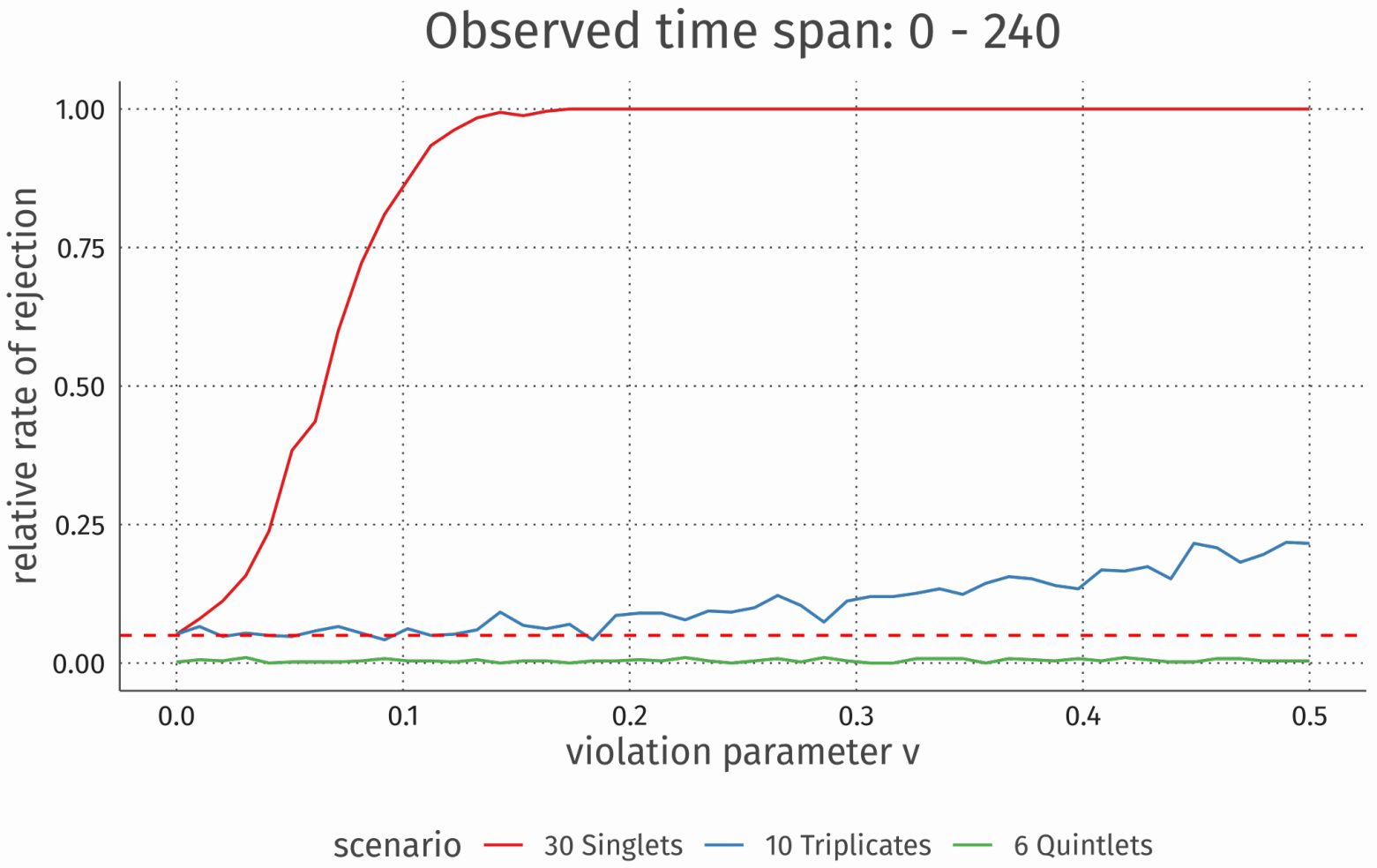
POTT-PLOT – FAST DYNAMIC



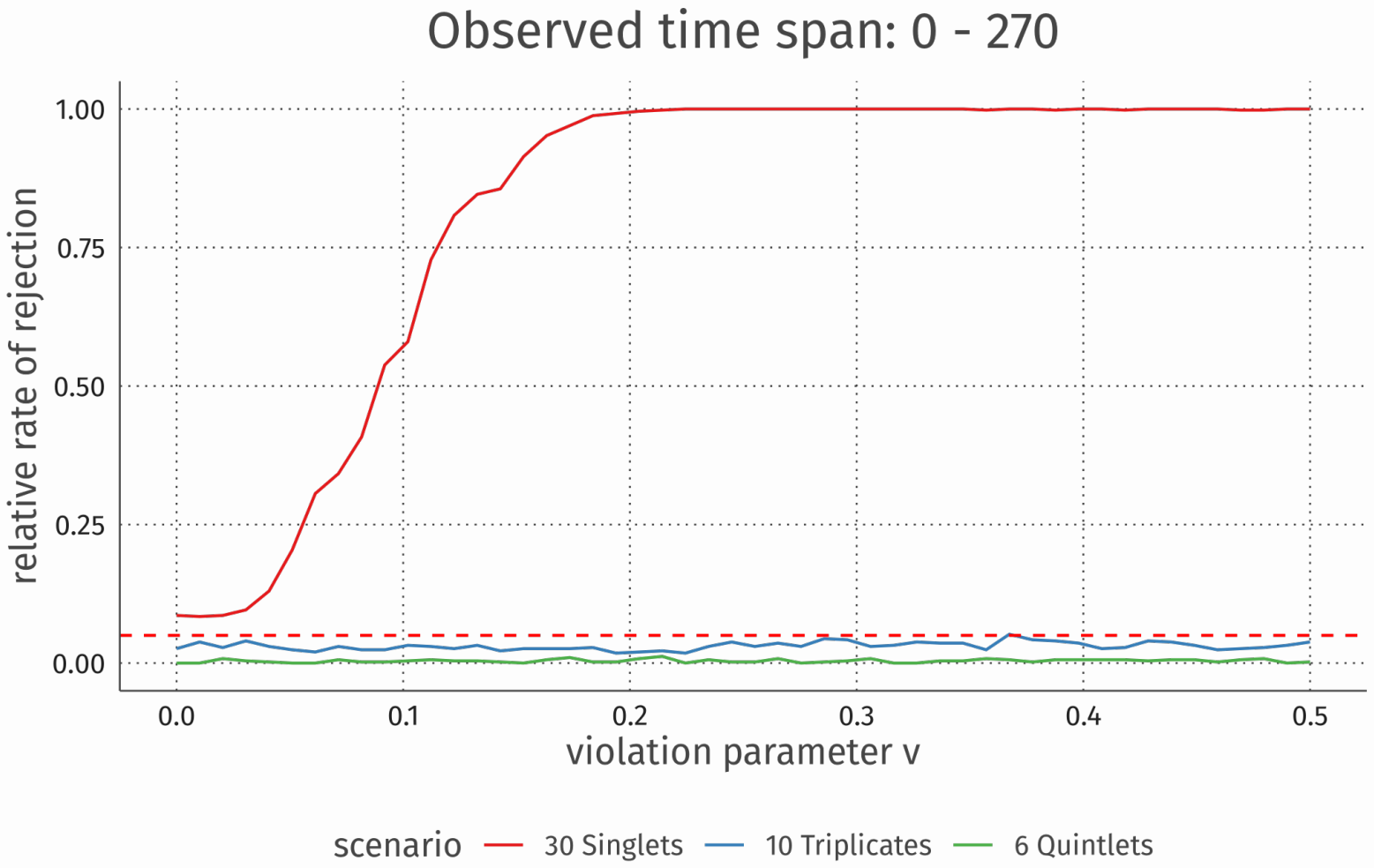
POTT-PLOT – FAST DYNAMIC



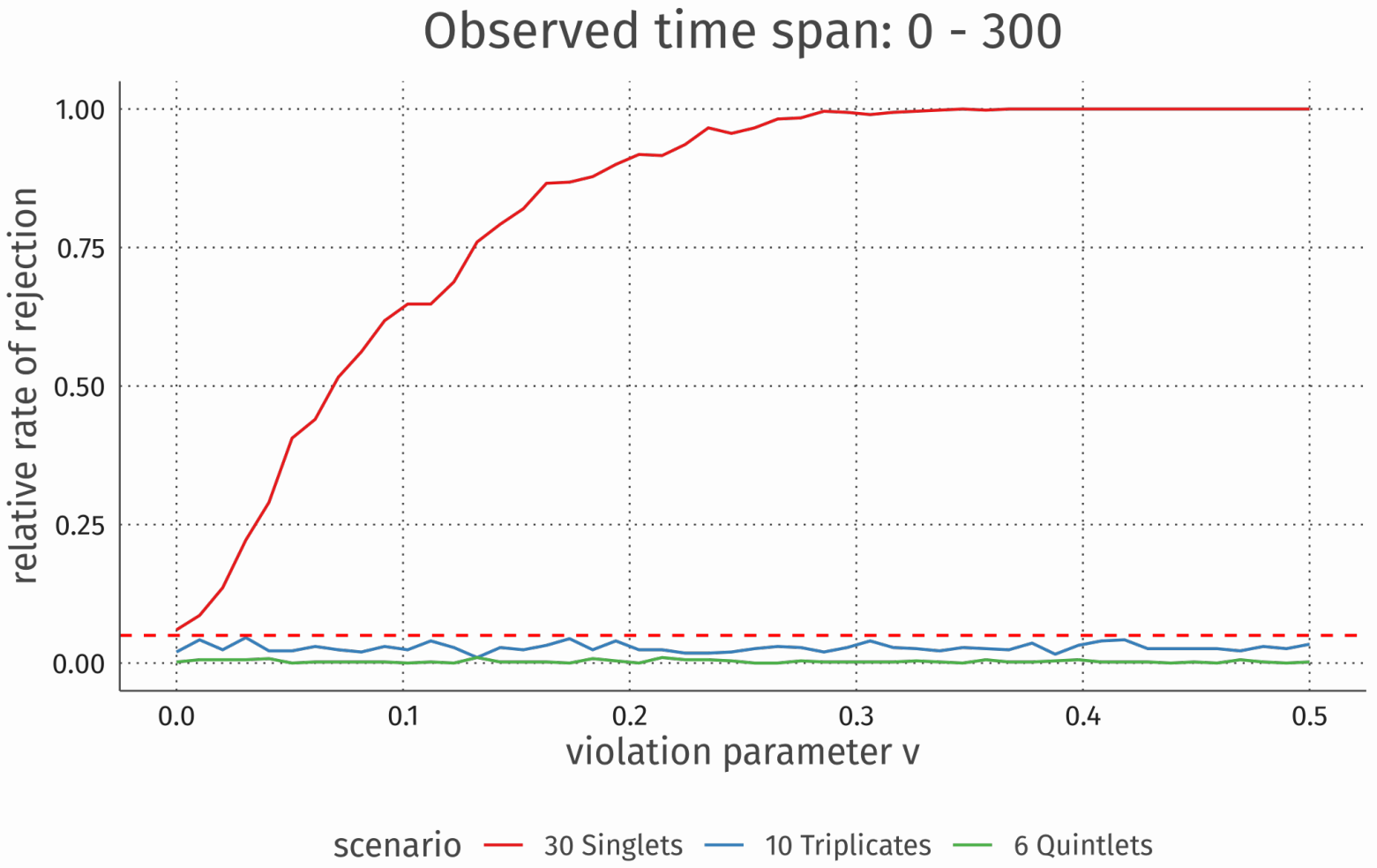
POTT-PLOT – FAST DYNAMIC



POTT-PLOT – FAST DYNAMIC



POTT-PLOT – FAST DYNAMIC



In short

- Insights:
  - No silver bullet
  - Replicates make only sense on points of high dynamic
  - Most efficient would be: pinpoint dynamic by singlets, additional replicates only on relevant points
  - But, that is only one aspect, there are more reasons for replicates (outliers, validation etc.)
- Outlook:
  - Test on typical dynamics
  - Evaluate averages vs. error model
  - Investigate alternative possibilities to detect outliers
  - ...

Thank you...

- ...to Jens Timmer & the whole JetiLab
- ...to Marcel Schilling
- ...to the bwForCluster Helix
- ...to you!

If this talk was one slide...

- Take home message:

  - Measuring more time points in fewer replicates is in general more informative
  - Replicates only useful when at time points of maximum dynamic

- If you want to know more:

Website email github