

Weighted Ensemble Sampling in Spatial Stochastic Simulations

Rory Donovan
MMBios Workshop 28/4/15

Resampling

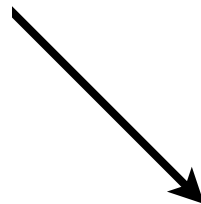
Original
Sample



Double Points,
Halve Weights



Halve Points,
Double Weights

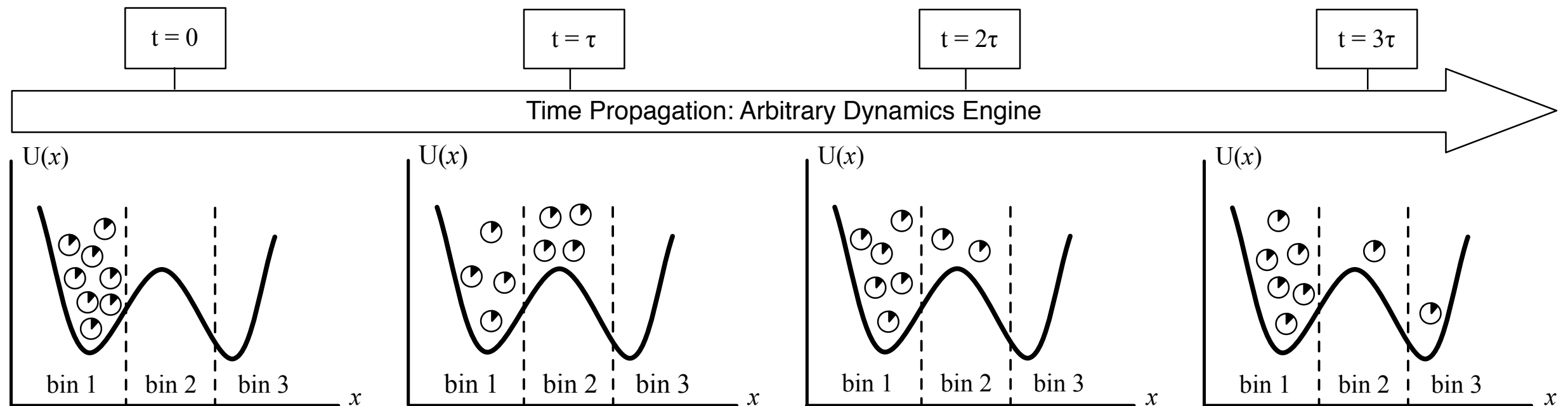


Do Both, In
Different Regions

A Weighted Ensemble is Just Repeated Resampling

- Of any stochastic process:
 - Molecular dynamics (protein motion)
 - Chemical kinetics (cell signaling)
 - Spatial stochastic models (MCell)
- Resampling is statistically exact

Ensemble Sampling



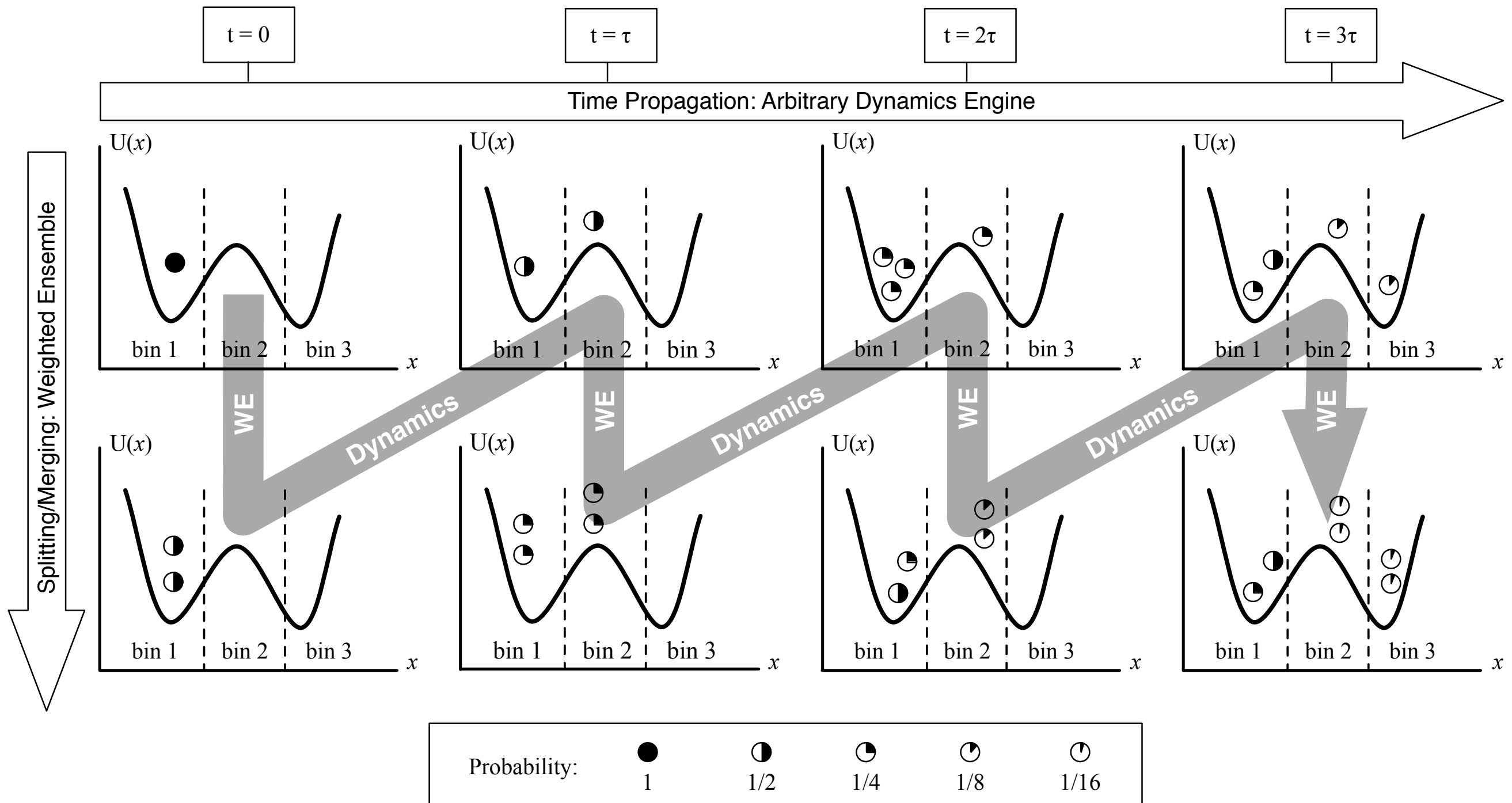
Probability: $\frac{1}{8}$

Repeated Runs

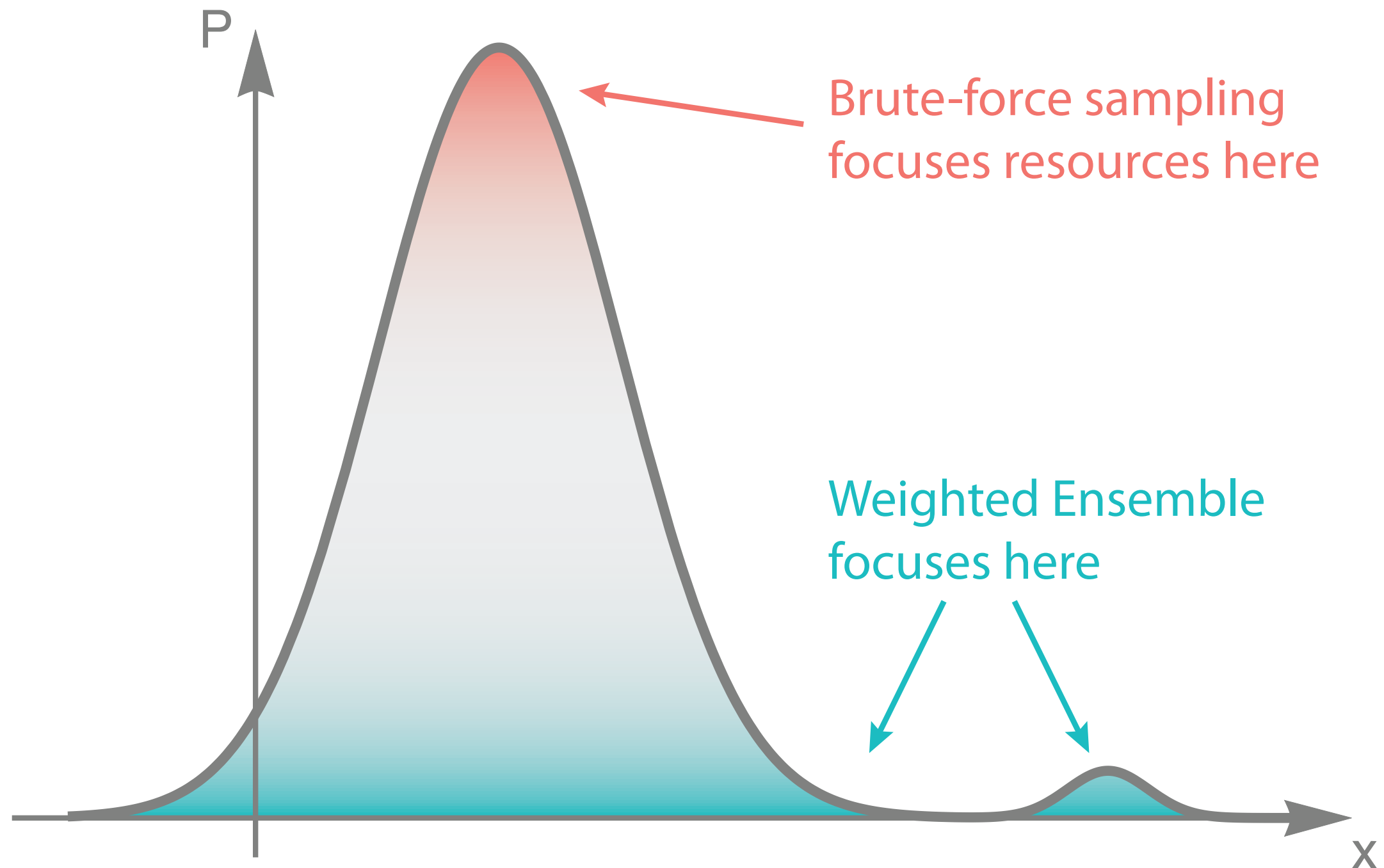


Normalized Histogram

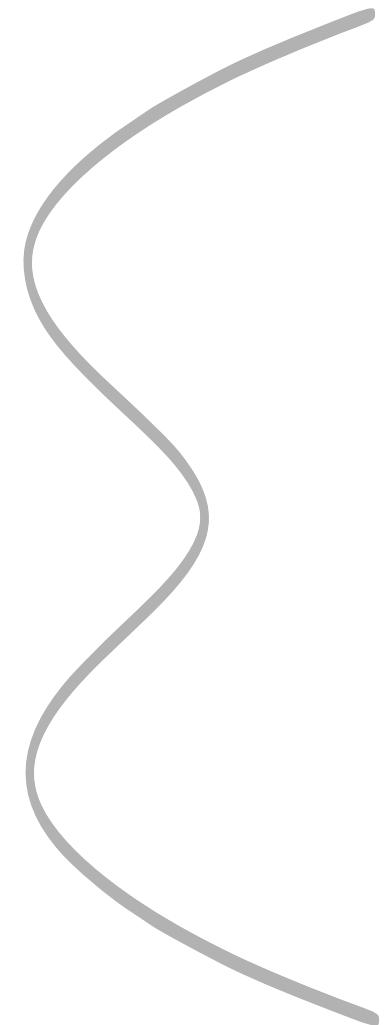
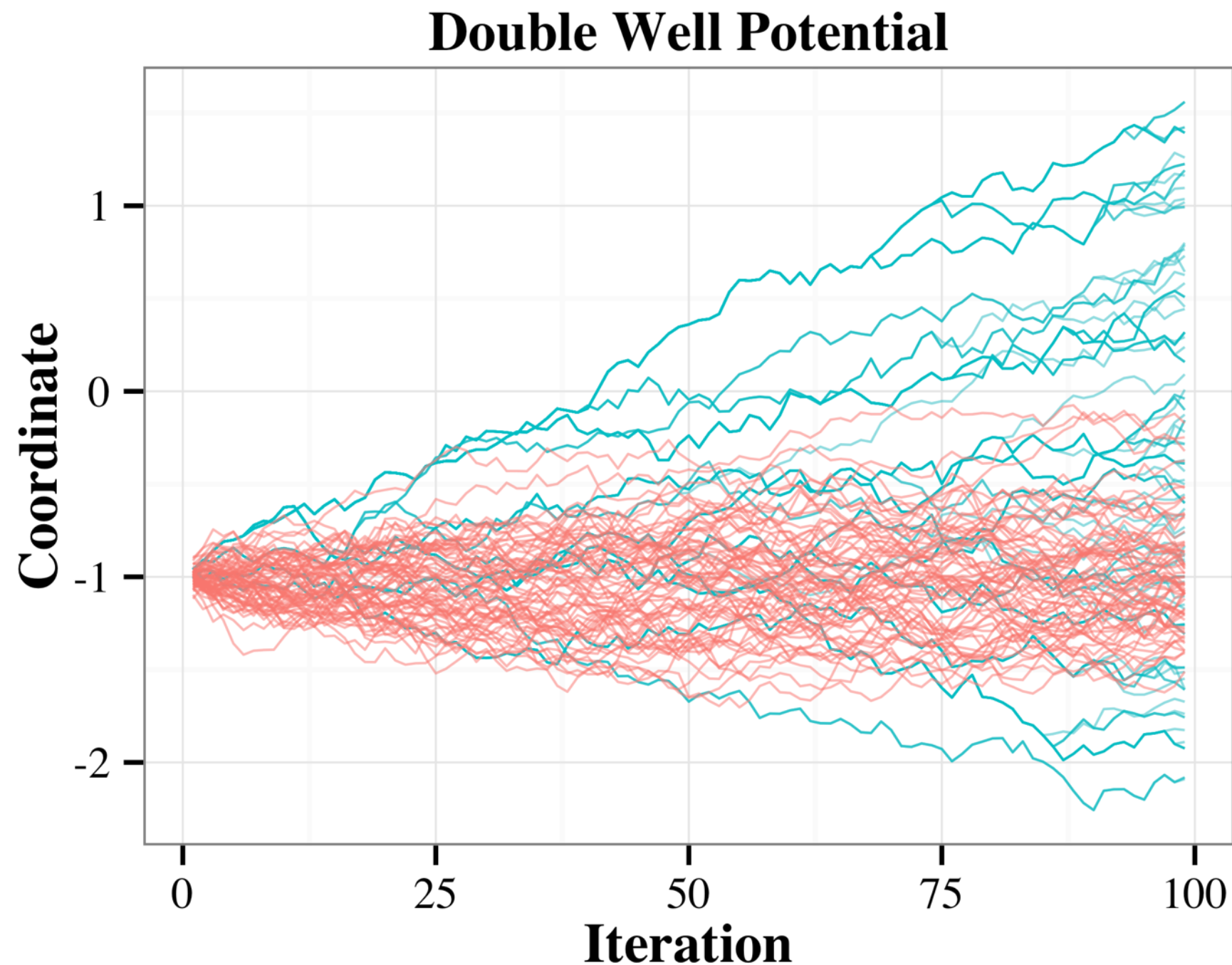
Weighted Ensemble Sampling



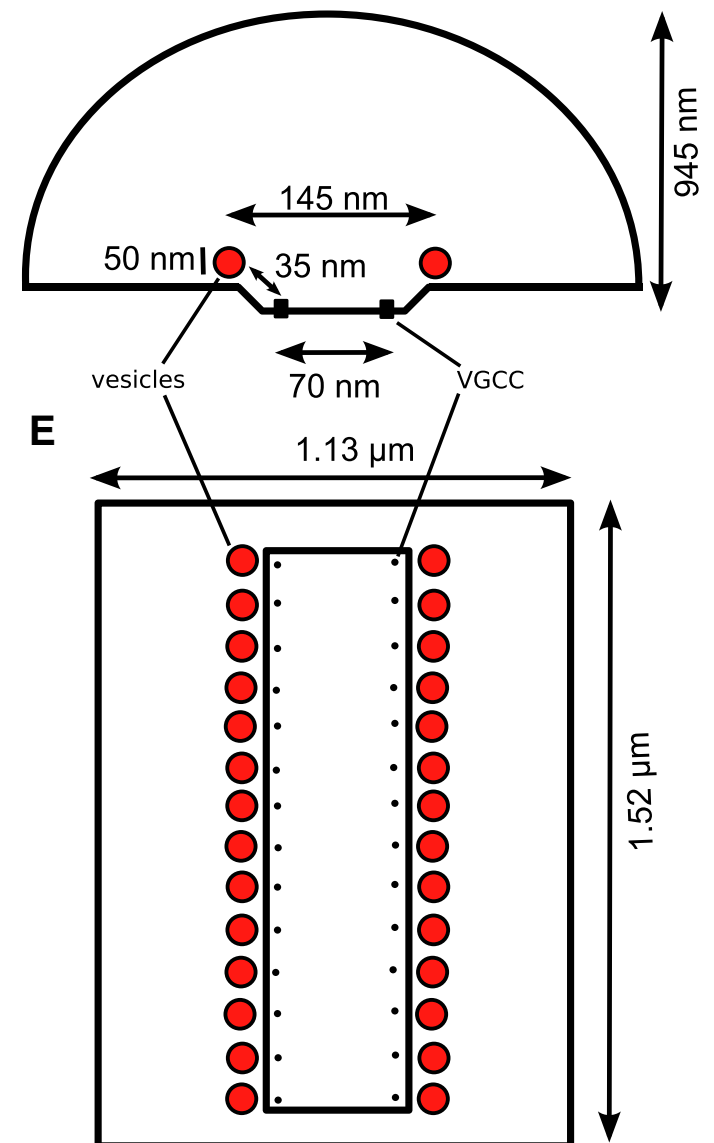
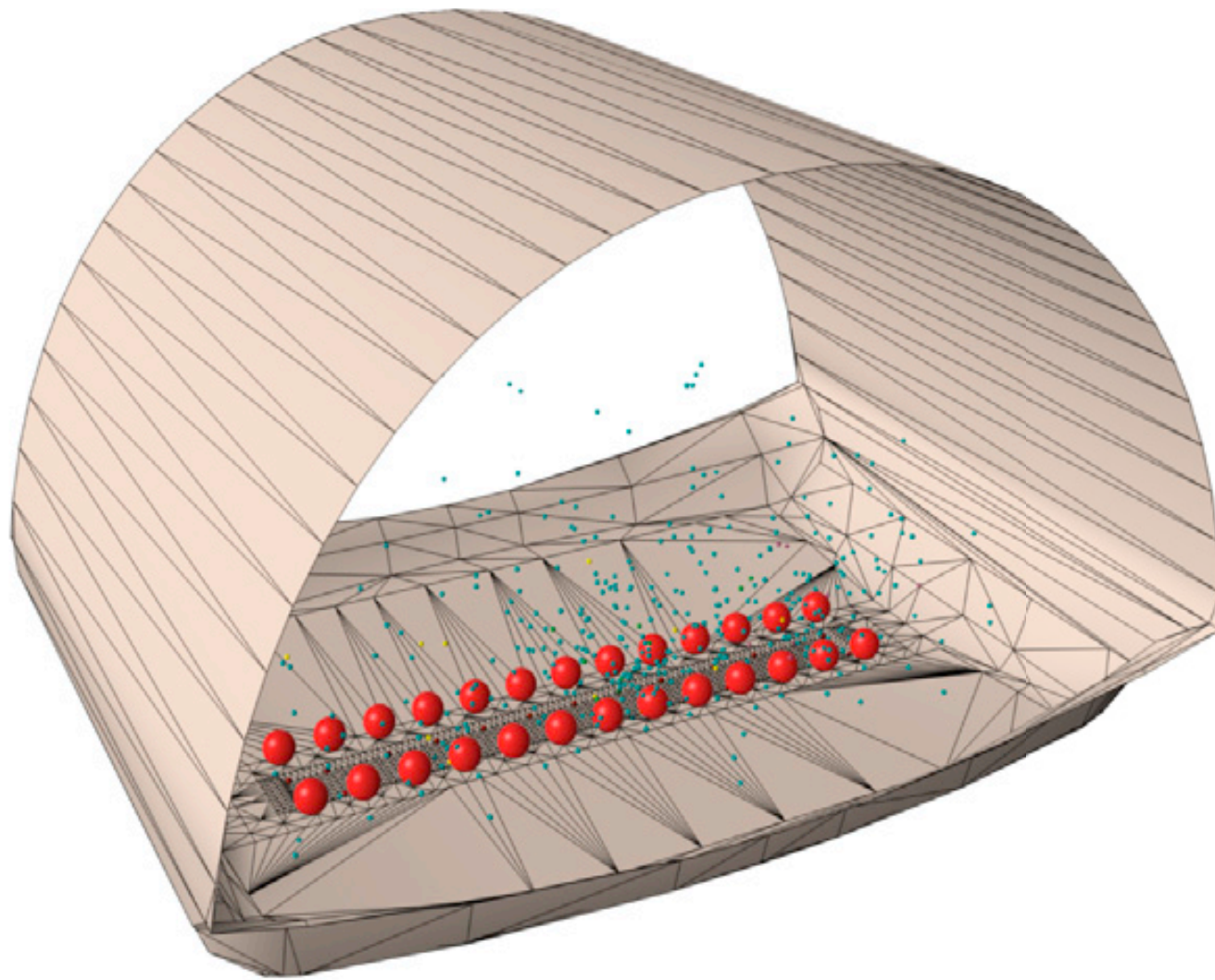
Weighted Ensemble Sampling



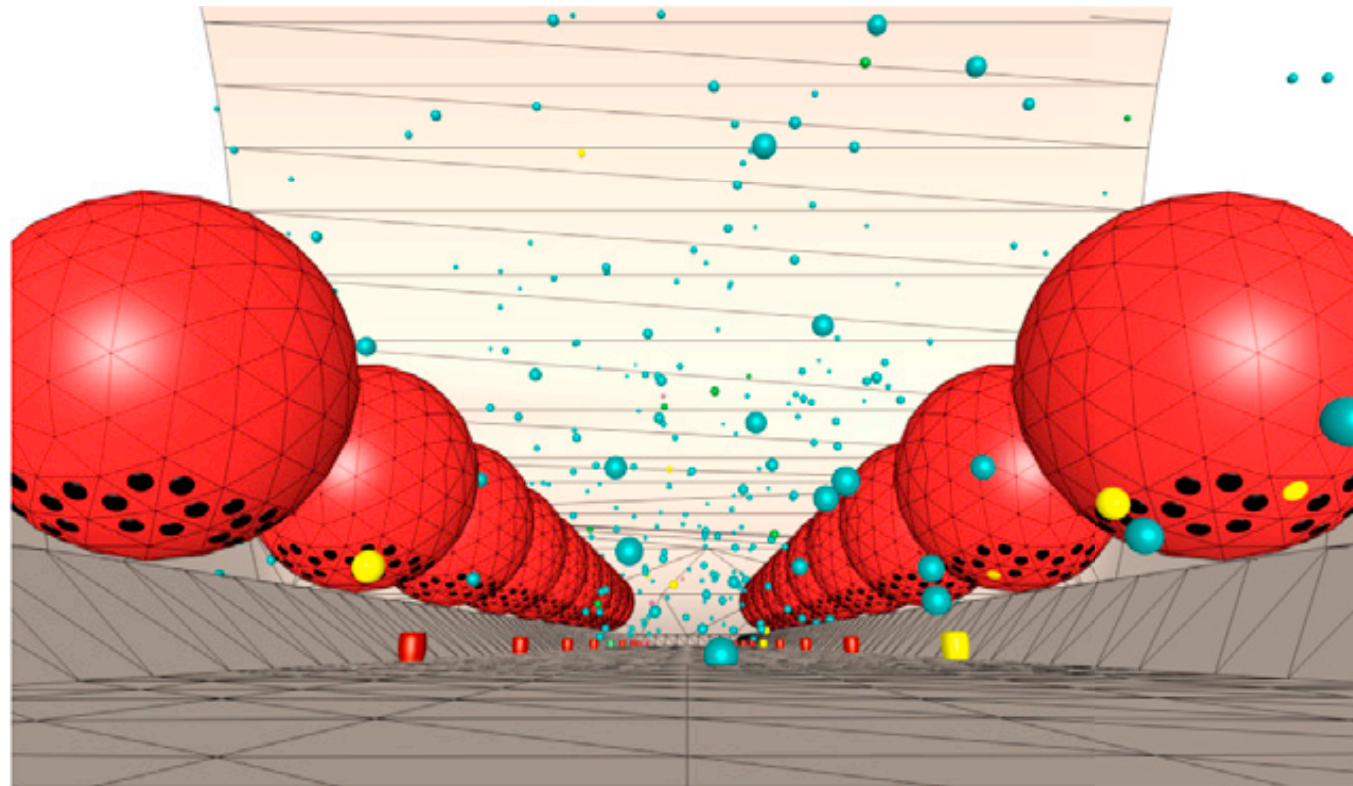
Ensembles: Weighted vs Unweighted



Neuromuscular Junction Model

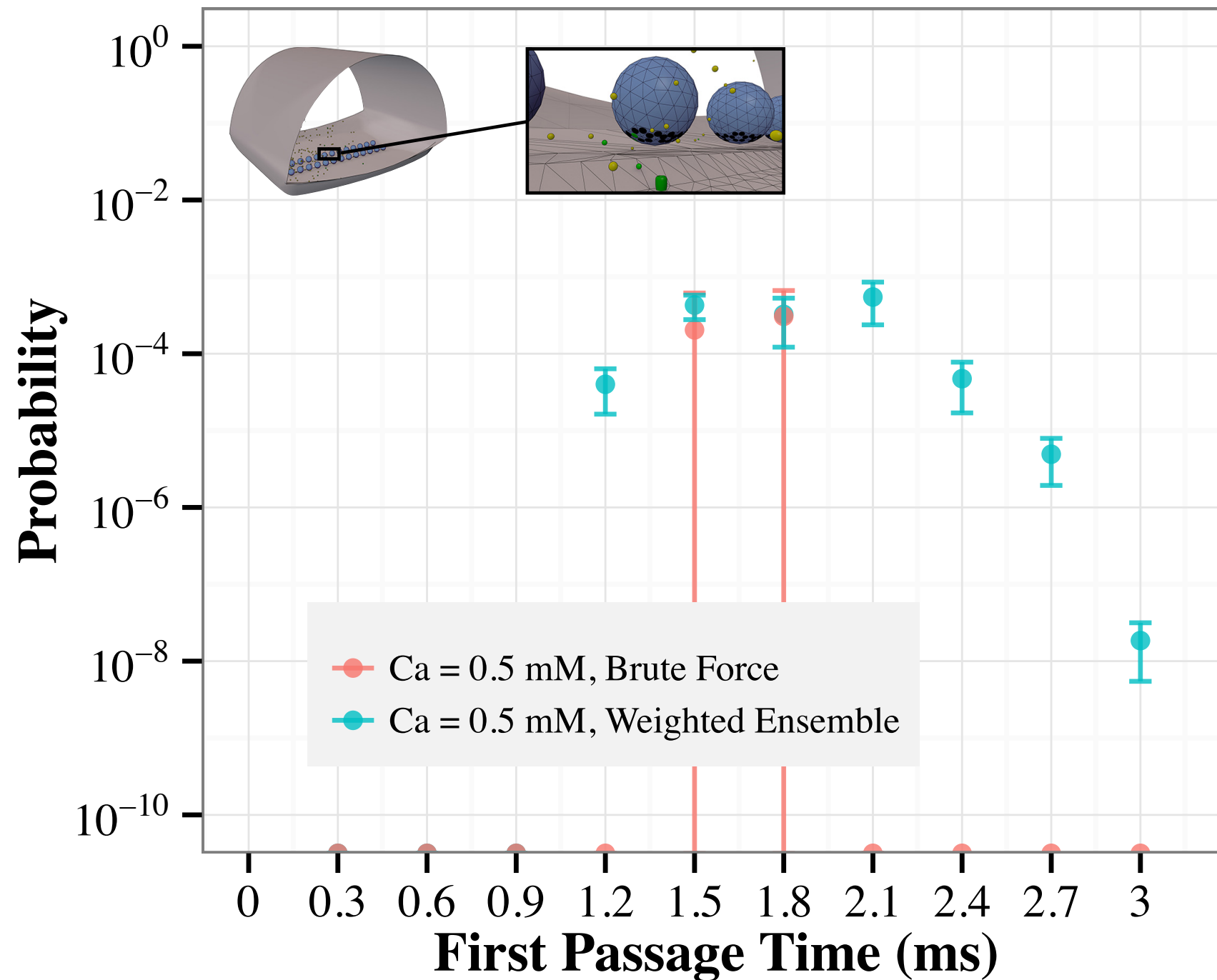


NMJ Zoomed-In

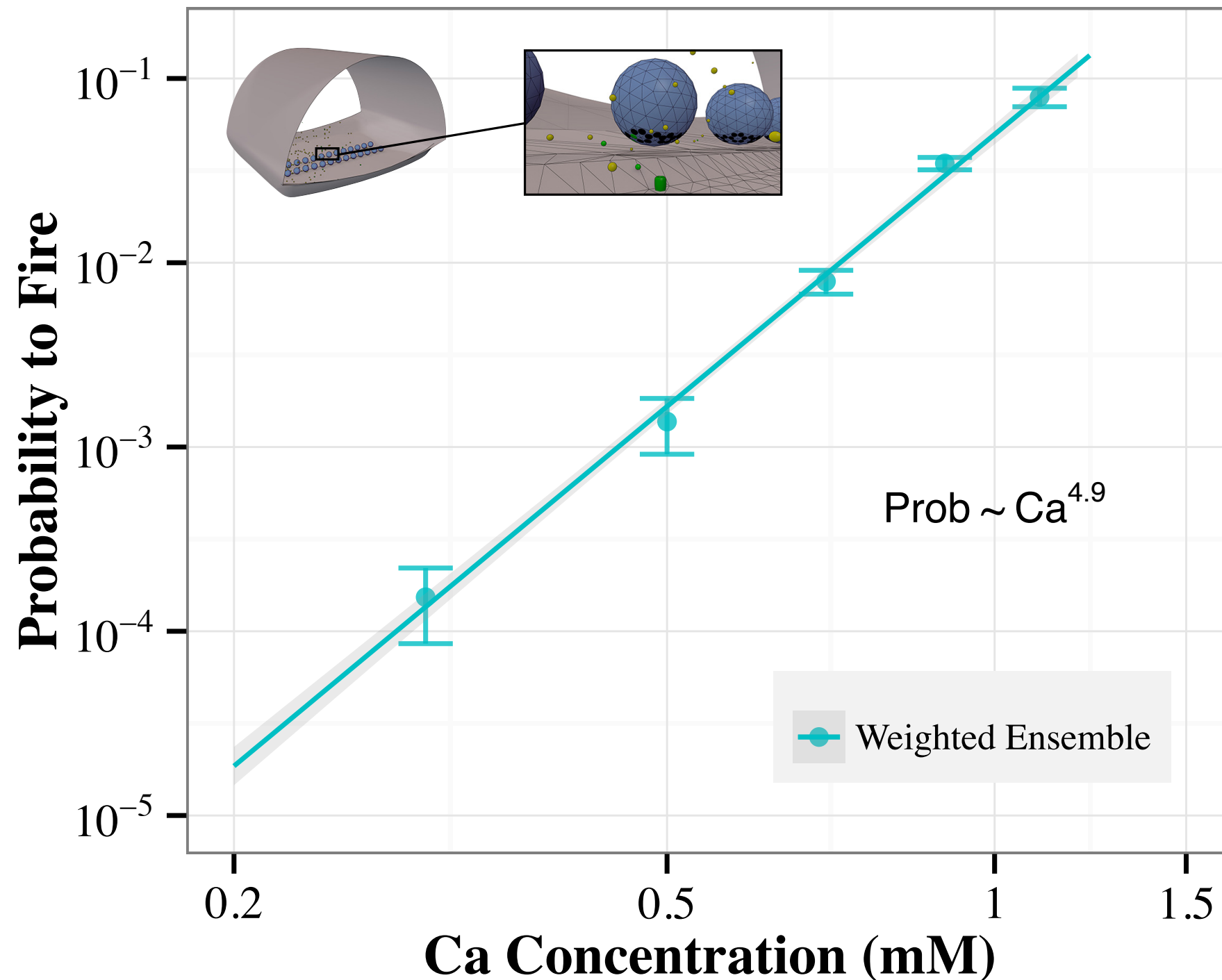


- Calcium is released from bottom, diffuses, and can bind to synaptotagmin vesicles
- Model: if enough calcium bind to one vesicle, in the right pattern, a release event occurs

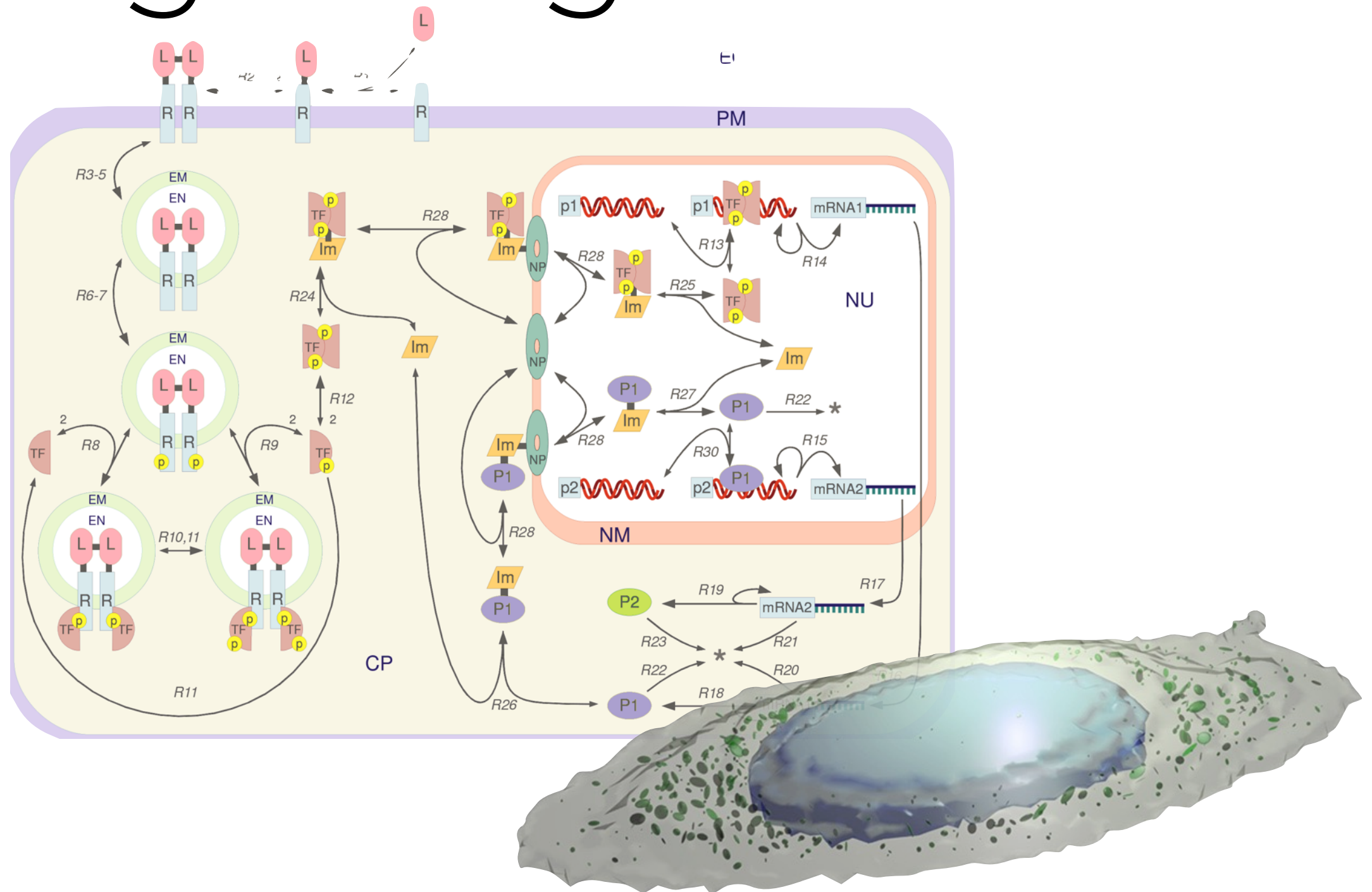
First Passage Times



Scaling Relationship



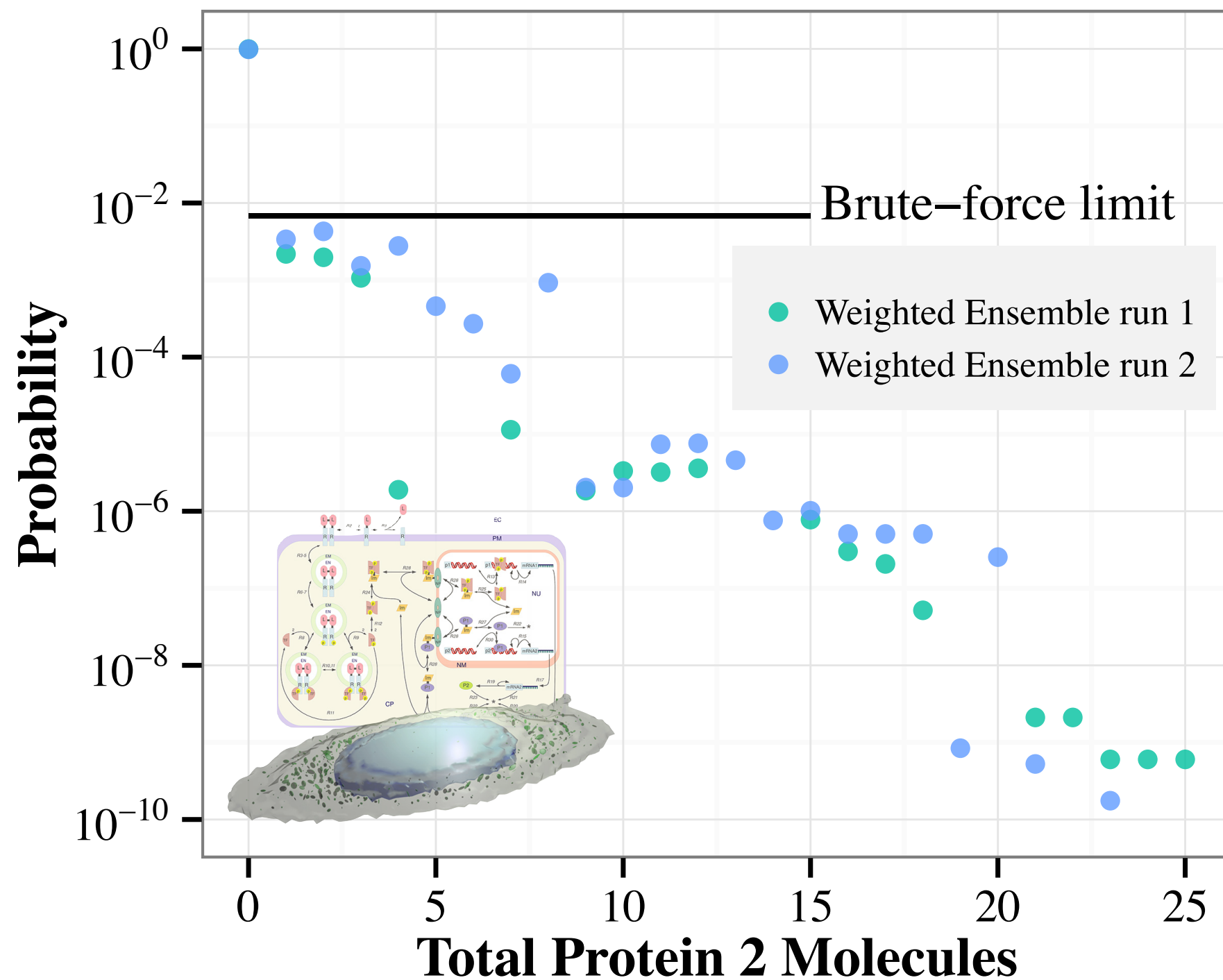
Signaling Network



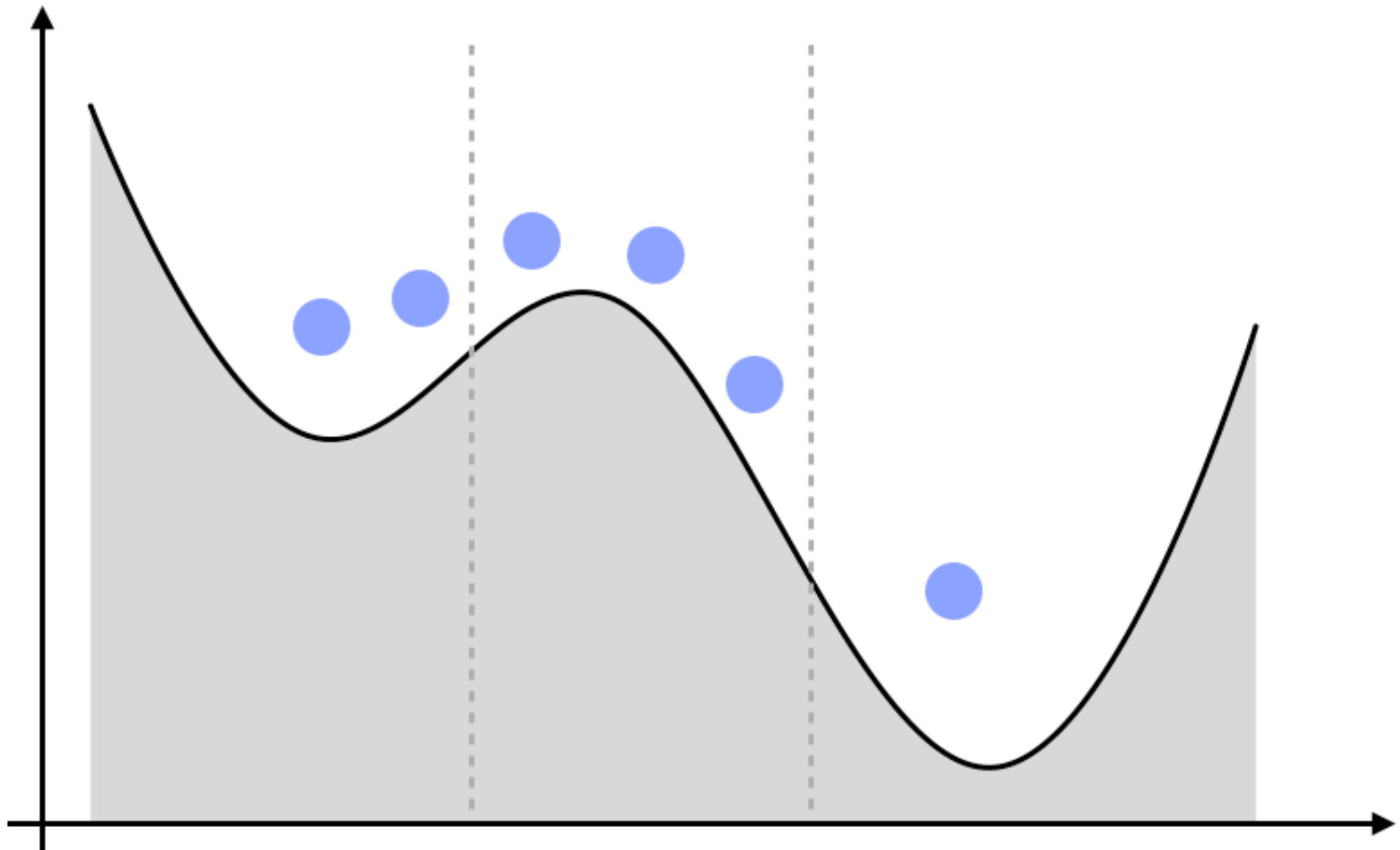
Pipeline: CellOrganizer → BioNetGen → MCell

Protein Histograms

$t = 400$ seconds



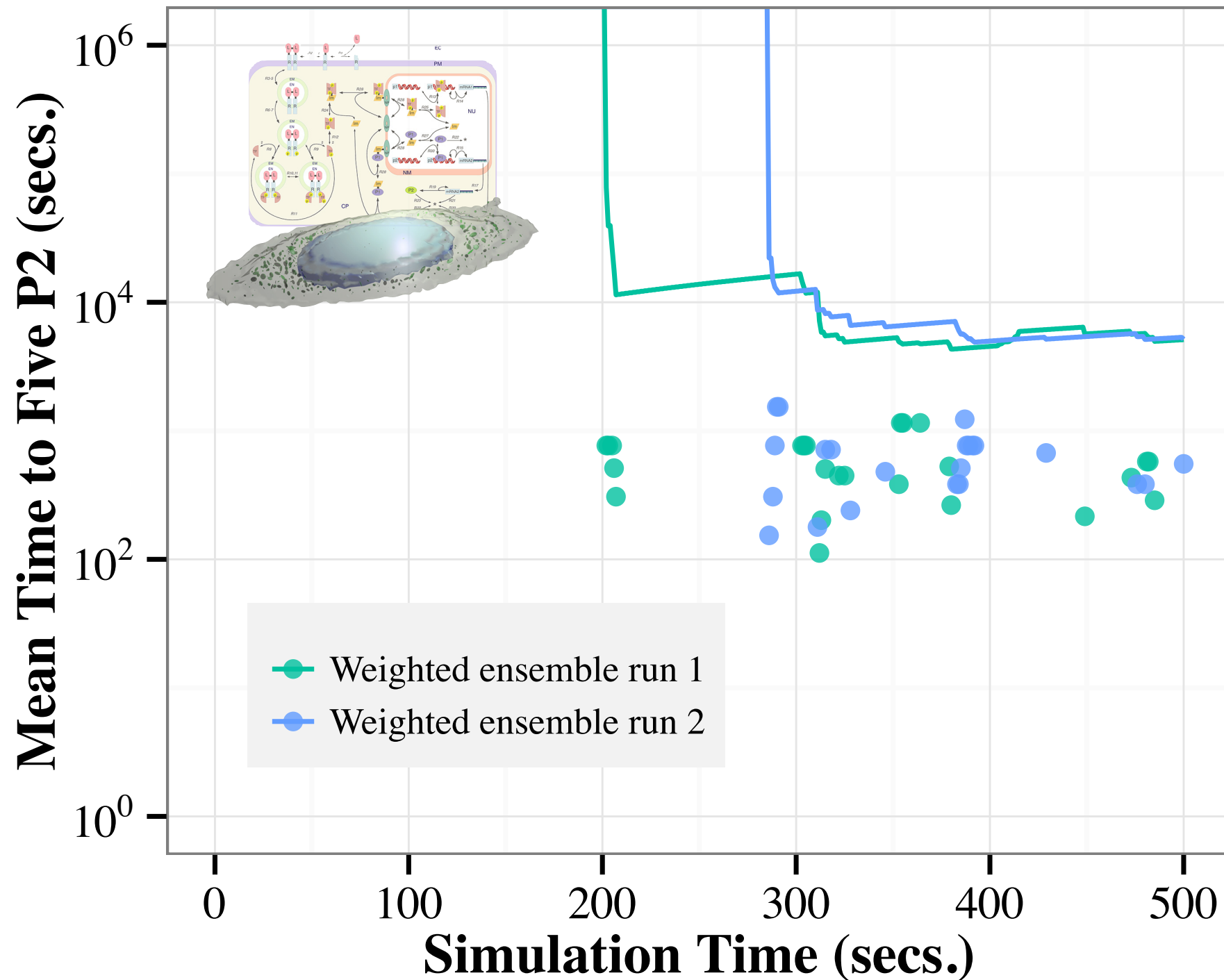
Steady-State



In steady state,

$$\text{MFPT}(A \rightarrow B) = 1/\text{Flux}(A \rightarrow B)$$

First Passage Times



Conclusions

- Able to sample the rare events and full distributions for stochastic systems biology models over a wide range of complexity
- Speed-up over brute-force is dramatic enough encourage the design of more complex, more realistic models
- Long time-scale behavior can be extrapolated from short simulations: can bridge dynamics over multiple timescales

More

- Code & help using it:
chong.chem.pitt.edu/WESTPA
github.com/westpa
- Workshop:
July 13-14, 2015