How can Bayesian networks be used for uncertainty quantification in particle physics?

Christina Peters

Computer & Information Sciences
University of Delaware

Physics & Astronomy
Rice University

Aaron Higuera, Shixiao Liang, Venkat Roy
Waheed U. Bajwa, Hagit Shatkay, Christopher D. Tunnell
Data-Intensive Discovery Accelerated by Computational Techniques for Science (DIDACTS) Collaboration

3 November 2022
Rutgers University

The Need for Uncertainty Quantification

- Uncertainty is present in all experimental measurements
- Uncertainty is necessary to test hypotheses
- Efficacy of many ML algorithms for science is limited by inability to quantify uncertainty on parameters

Statistical Model Inference

- complex, high-dimensionality, high-fidelity models

  likelihood is intractable

- ML provides a variety of ways to estimate the likelihood

  with summary statistics or density estimation

The frontier of simulation-based inference Cranmer et al. 2020

Uncertainty Quantification by ML

- Provide a single locally optimal solution

  or

- Estimate of uncertainty, usually requires recalibration
- Uncertainty quantification techniques are not interpretable,

  not grounded in probability theory / statistics

Interpretable Uncertainty Quantification in AI for HEP Chen et al. 2022

Probabilistic Graphical Models

- relatively simple, but powerful
- physically interpretable
- based on a formalism for probabilistic reasoning
- graph-based representation to compactly represent the likelihood
- designed for inference / uncertainty quantification

Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference Pearl 1988

Example Use Case in Particle Physics

- reconstruction
- developed for use by a dark matter direct detection experiment
    - TPC
    - arrays of PMTs on top and bottom
    - ρ = 66.4 cm, z = 148.5 cm

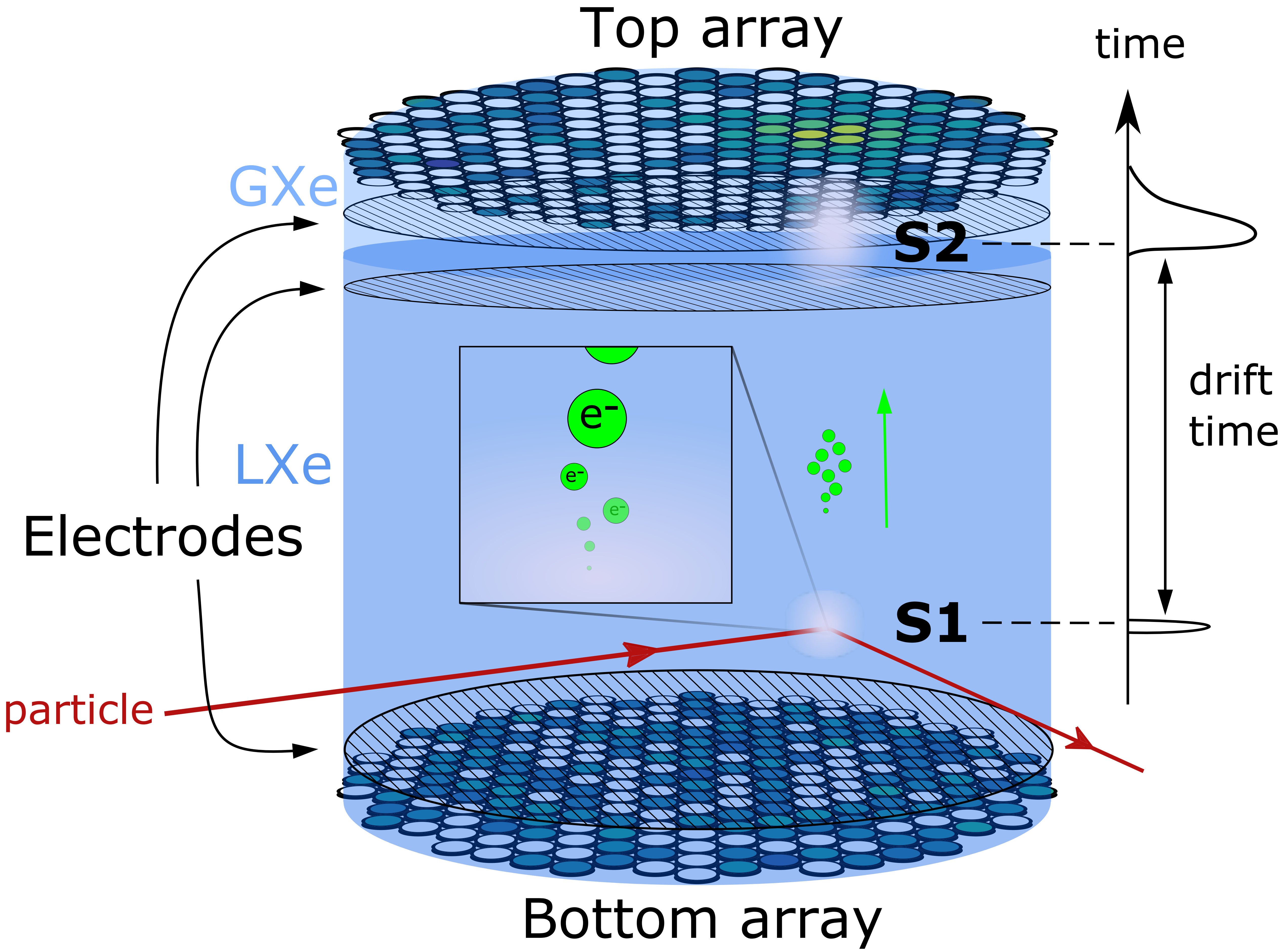


Figure credit: *Liang et al. 2022*

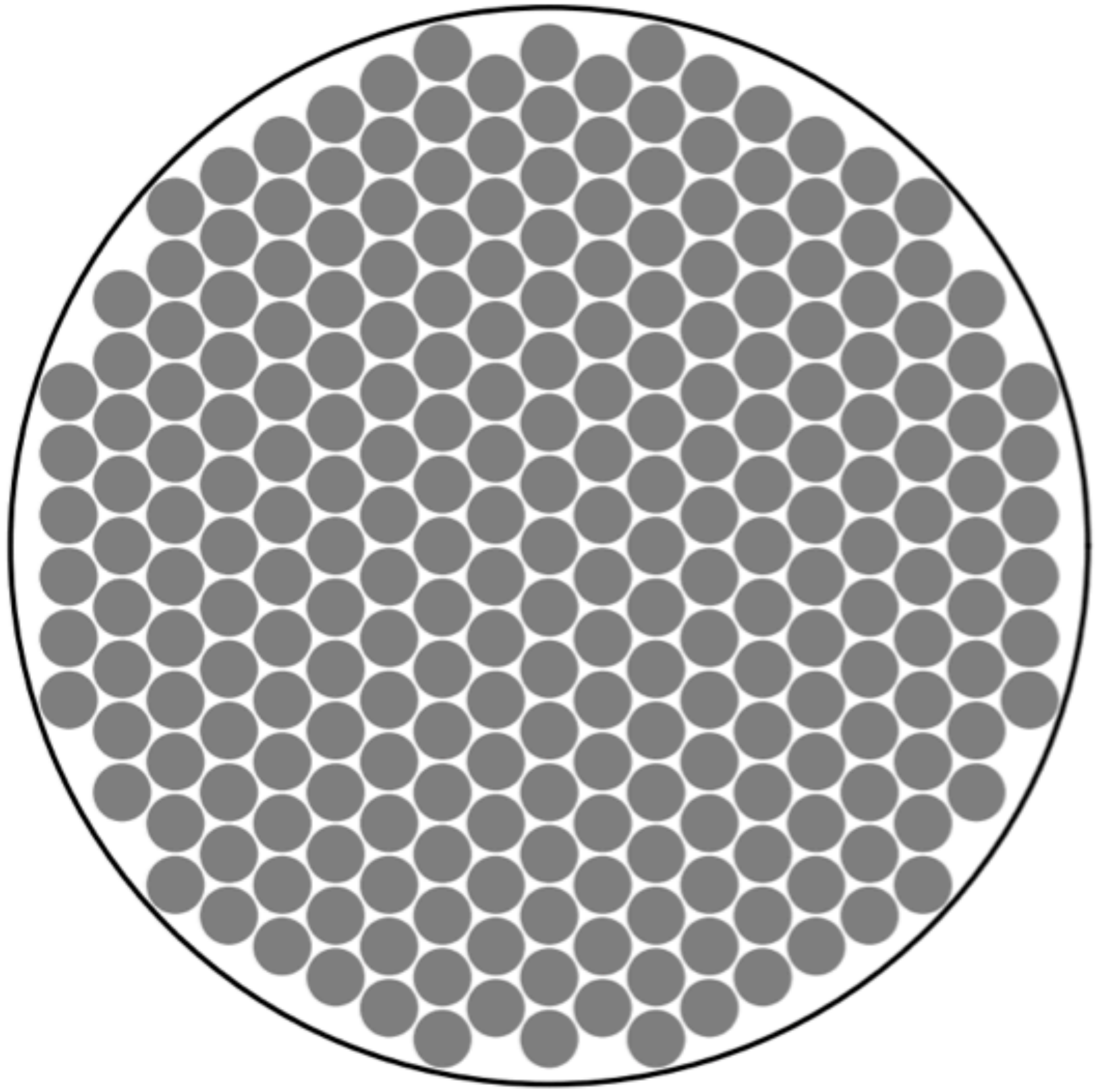
Bayes' Rule - for reconstruction

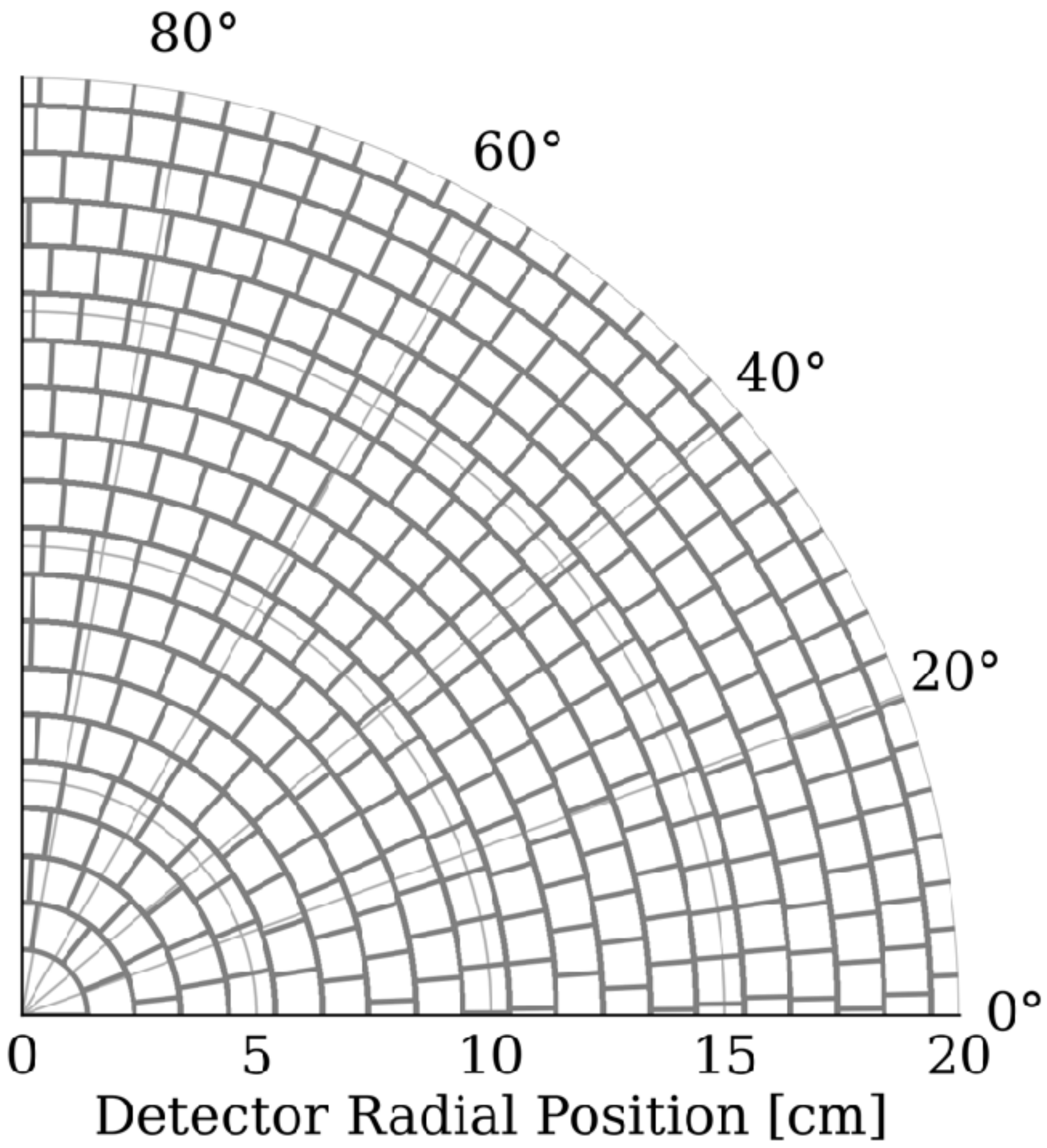
- Physical processes occurring within a detector, such as the interaction position, are inferred from sensor observables

- We know (from simulations) $P(\text{detected by sensors} \,|\, \text{event location})$
- We want to know $P(\text{event location} \,|\, \text{detected by sensors})$

$$P(\text{event location} \,|\, \text{detected by sensors}) = \frac{P(\text{detected by sensors} \,|\, \text{event location})\; P(\text{event location})}{P(\text{detected by sensors})}$$

Nodes - representation of the variables





#0 #1 #2 ... #250 #251 #252

Range = { measured intensities }

L

Range = { 0, … , Number of Pixels}

Graphical Model - syntax

Node – a random variable (discrete or continuous)
Edge – indicates dependencies between nodes

Graphical Model - syntax

Node – a random variable (discrete or continuous)
Edge – indicates dependencies between nodes

Directed edges, indicated with arrows, can be used to indicate believed causality.

Definition

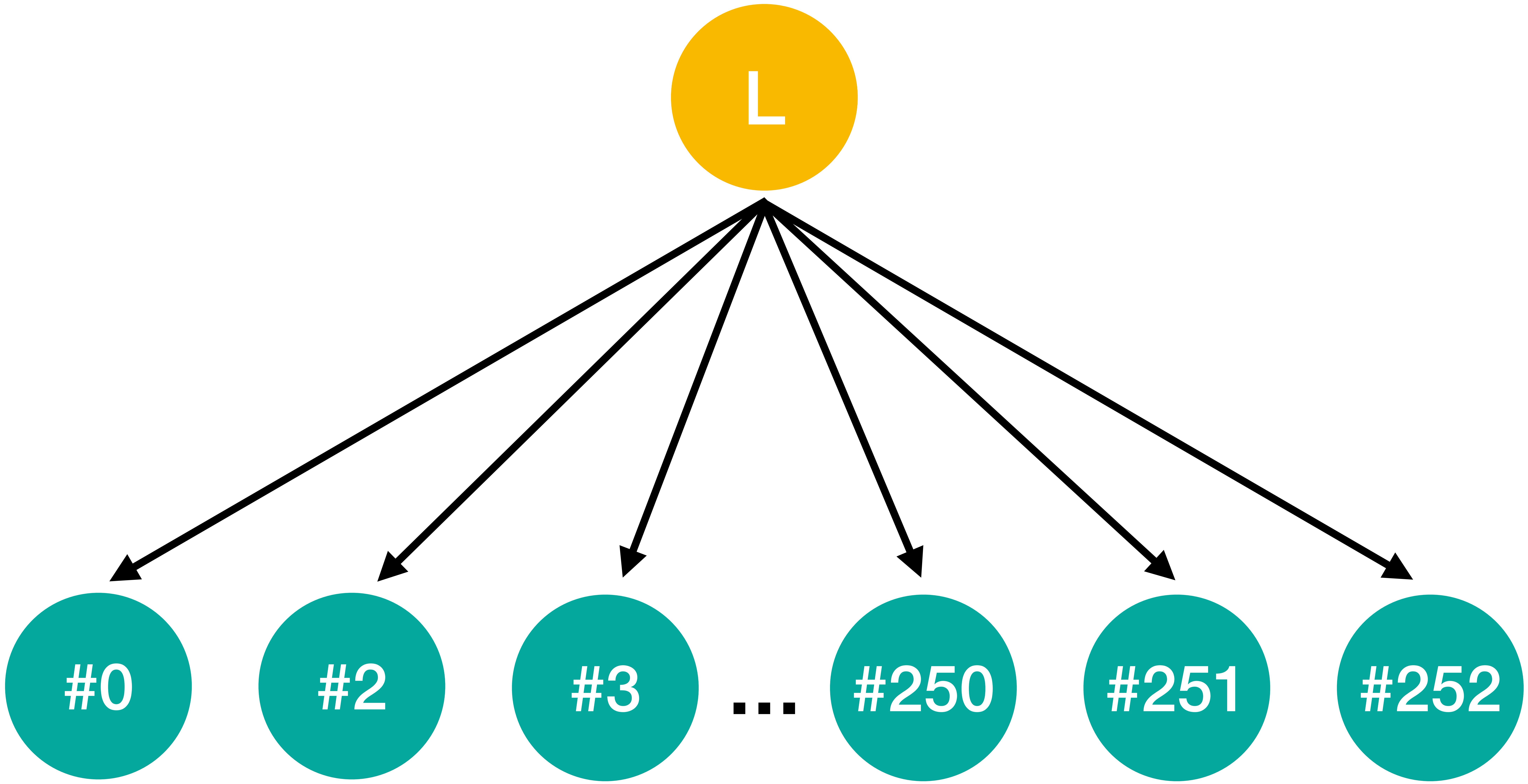
A Bayesian Network is a probabilistic graphical model that has:

- only directed edges
- no cycles
  i.e. arrows can not be followed in a closed loop

A directed acyclic graph (DAG).

Simplest Graph Structure

known as a Naive Bayes Classifier



Simplest Graph Structure

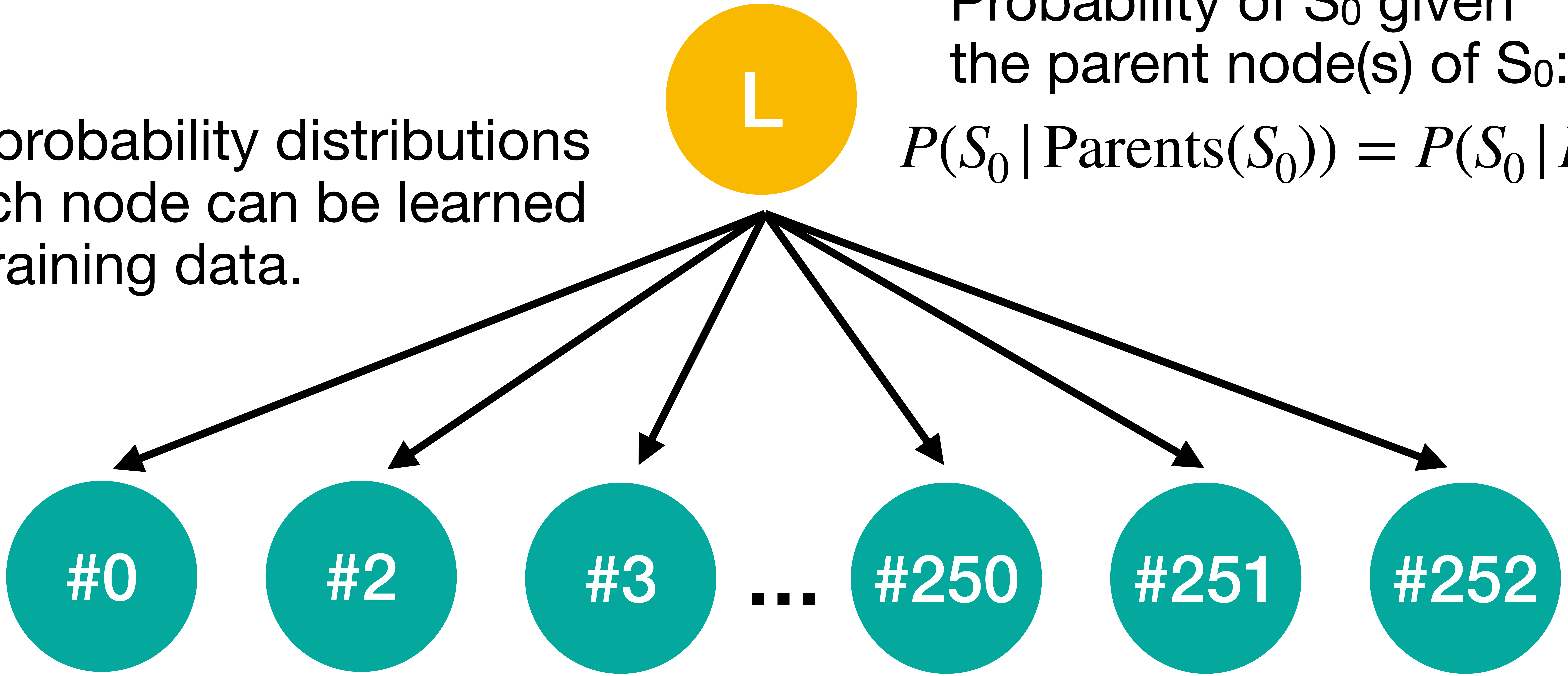
known as a Naive Bayes Classifier

Conditional Probability

Probability of S_0 given the parent node(s) of S_0:

$$P(S_0 \,|\, \text{Parents}(S_0)) = P(S_0 \,|\, L)$$

Local probability distributions for each node can be learned from training data.



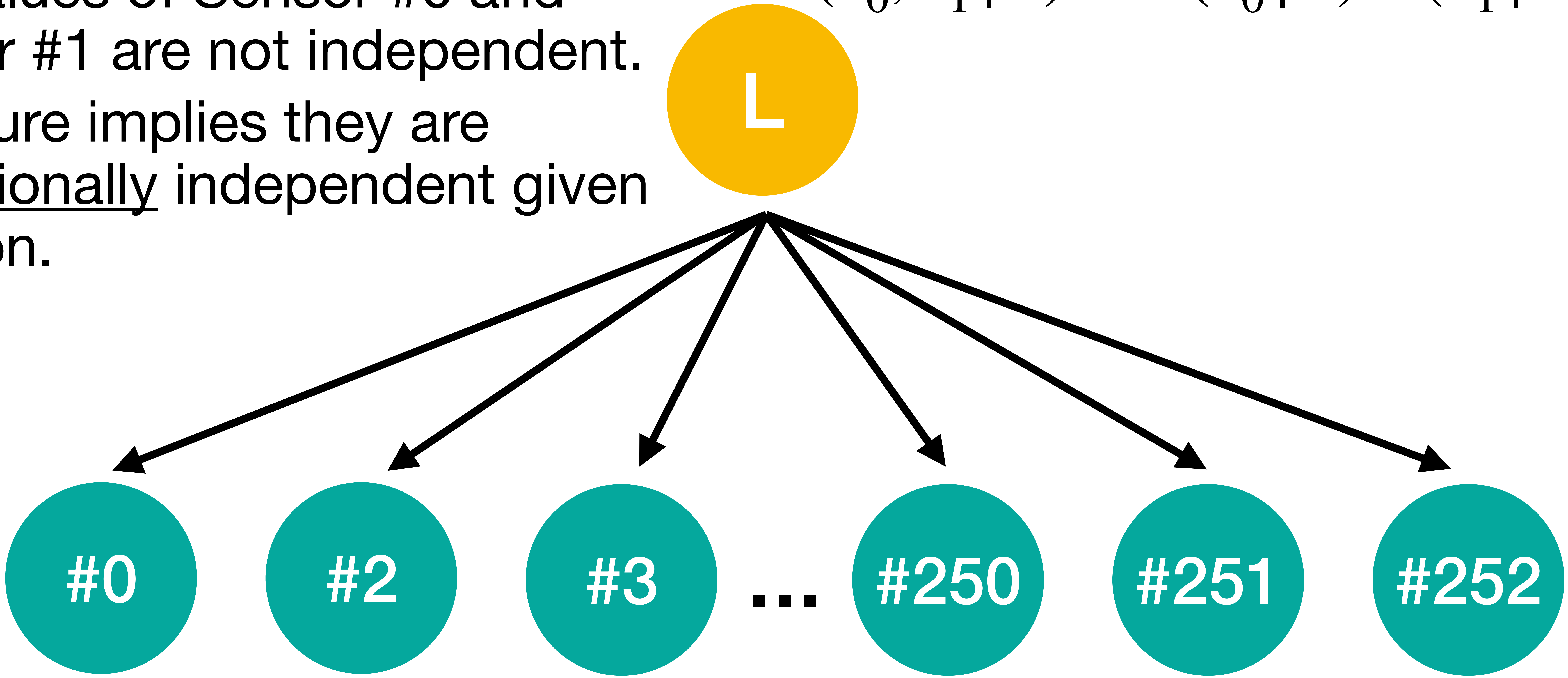
Simplest Graph Structure

known as a Naive Bayes Classifier

Conditional Independence

$$P(S_0, S_1 | L) = P(S_0 | L)\ P(S_1 | L)$$

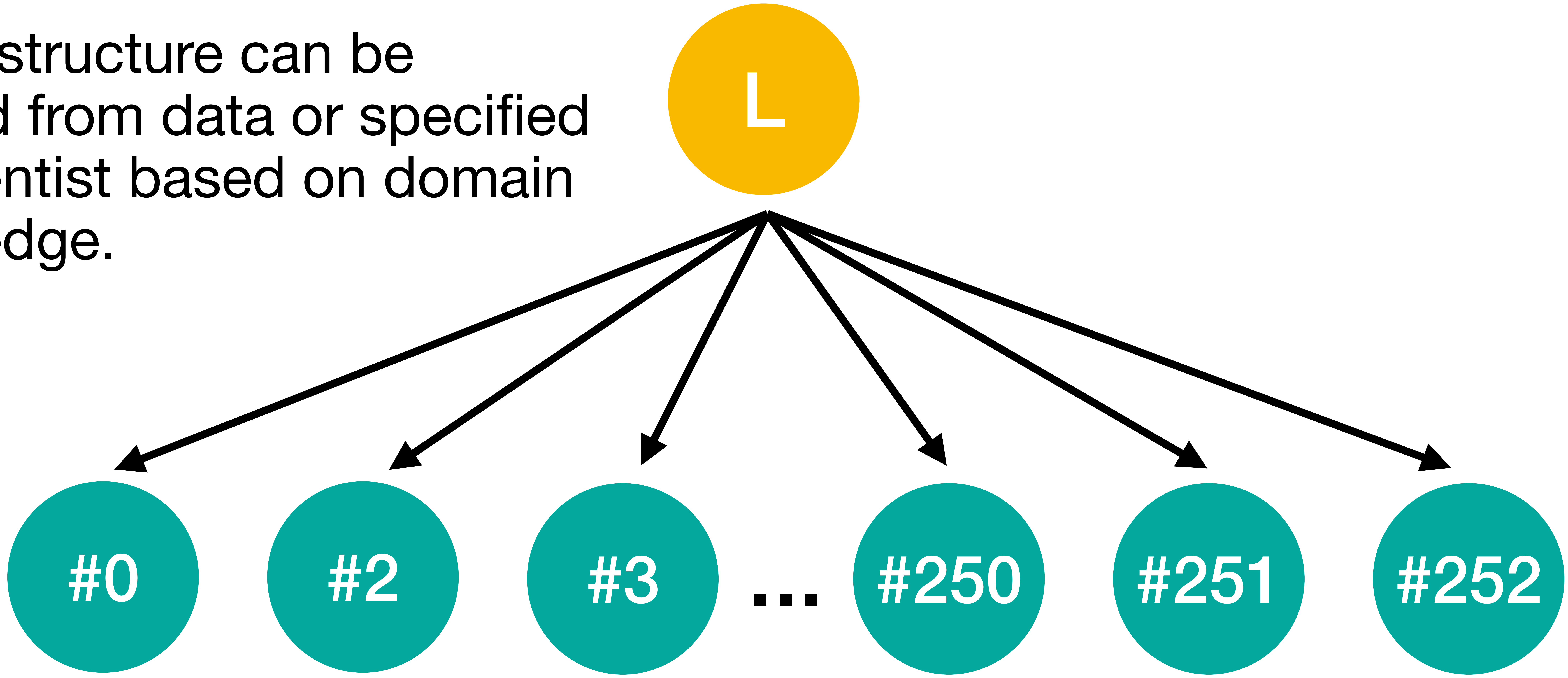
- The values of Sensor #0 and Sensor #1 are not independent.
- Structure implies they are <u>conditionally</u> independent given location.



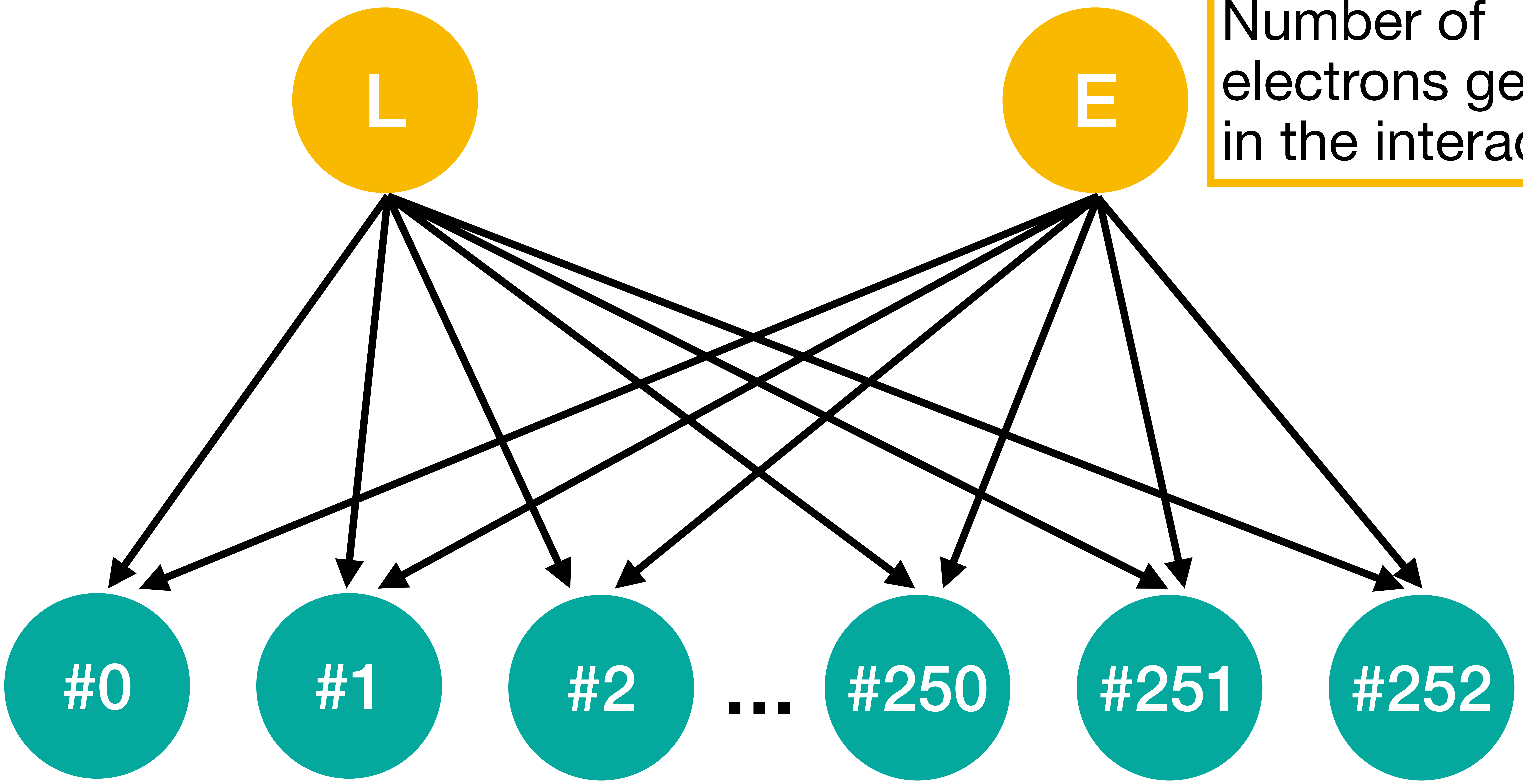
Simplest Graph Structure

known as a Naive Bayes Classifier

Graph structure can be learned from data or specified by scientist based on domain knowledge.



Less Simple Graph Structure



Number of electrons generated in the interaction.

Inference

Evidence Variables: #0 = #0 , #1 = #1 , #2 = #2 ...

Query Variables: L E

Non-Evidence & Non-Query Variables: **none**

$$P(L, E \mid \#0 = \#0, \#1 = \#1, \ldots)$$

$$= P(L, E)\, P(\#0 = \#0, L, E)\, P(\#1 = \#1, L, E) \ldots$$

Inference

Evidence Variables: #0 = #0 , #1 = #1 , #3 = #3 ...

Query Variables: L E

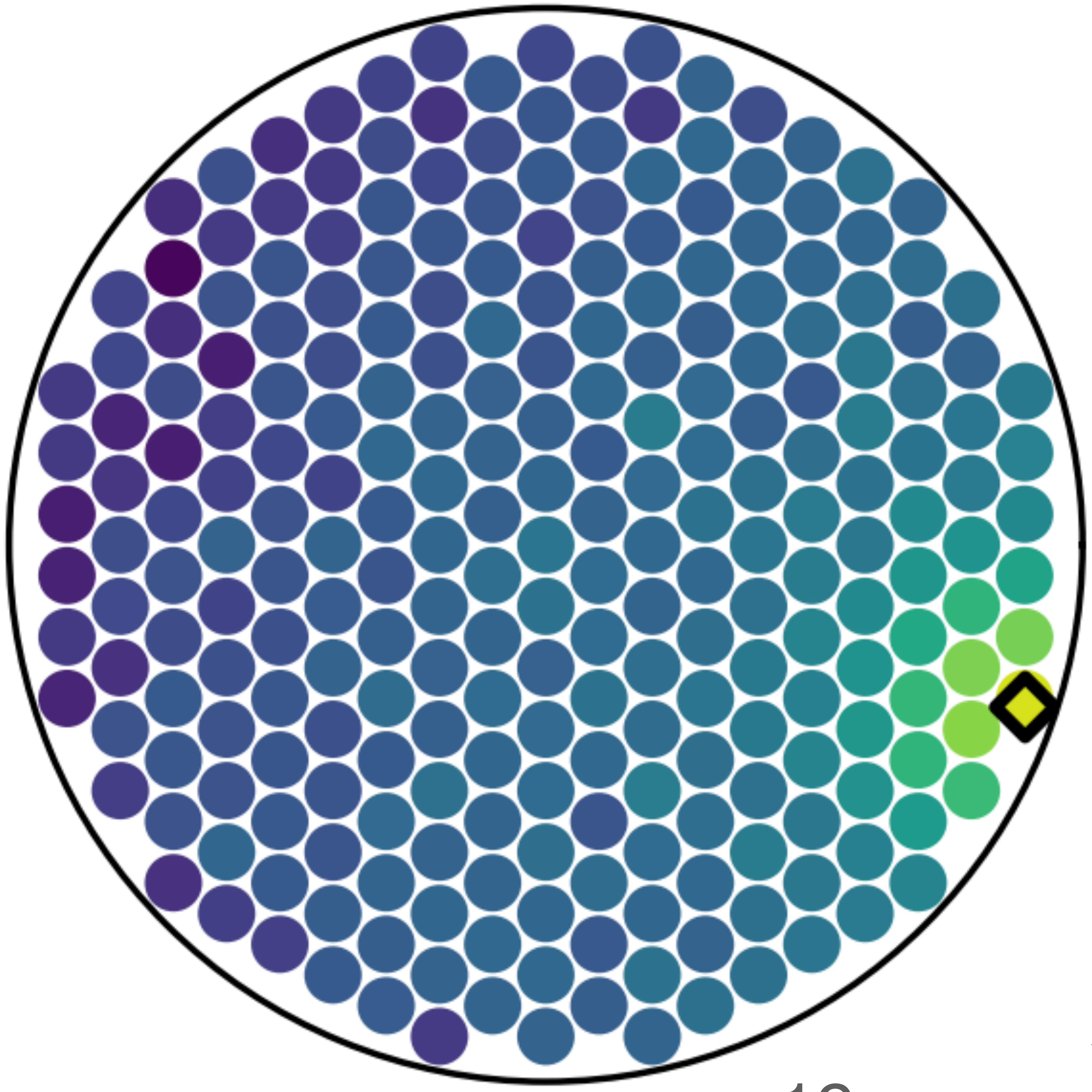
Non-Evidence & Non-Query Variables: #2

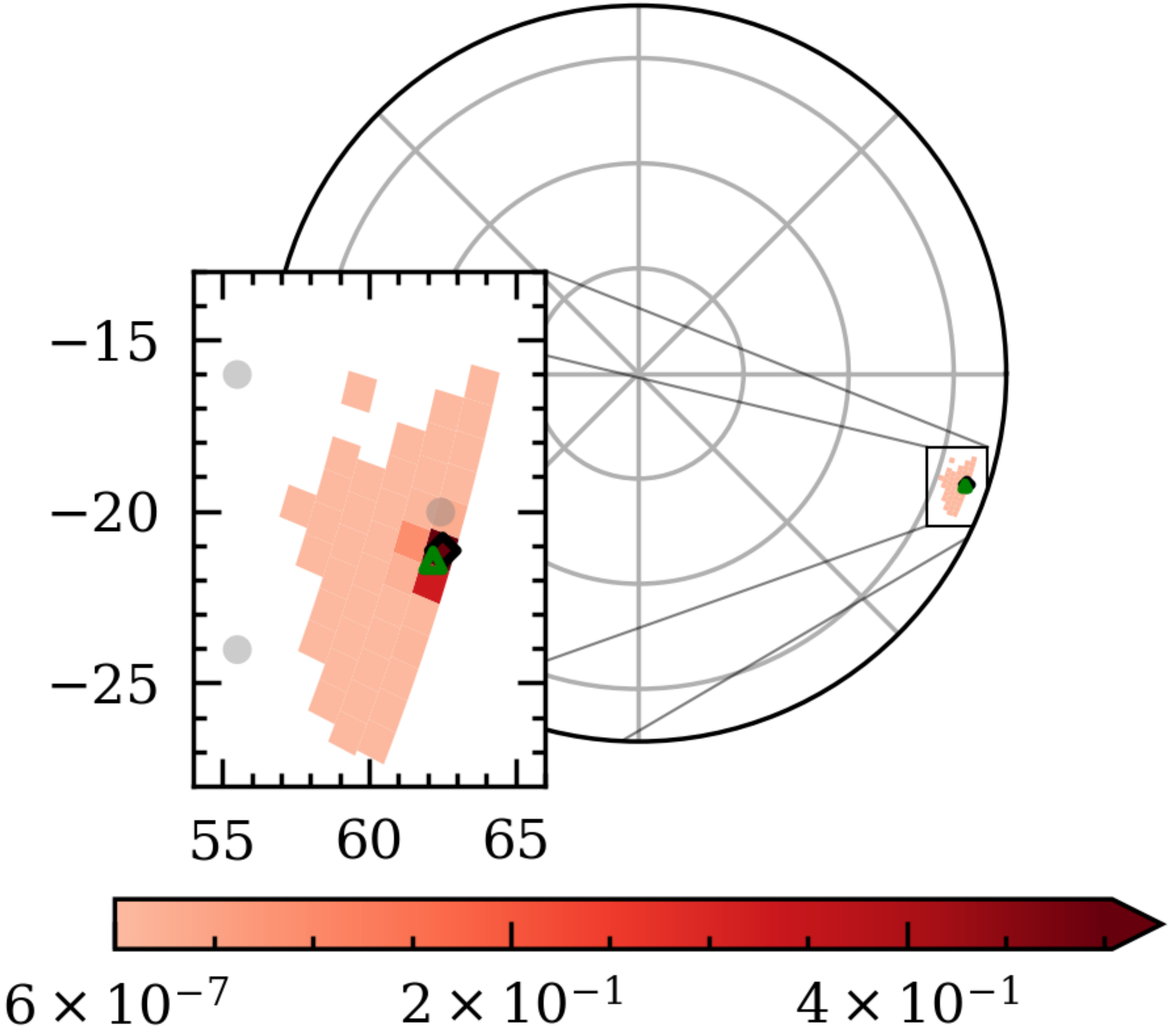
e.g. sensor #2 is broken

$$P(L, E \mid \#0 = \#0, \#1 = \#1, \#3 = \#3, \ldots) = P(L, E) \sum_{\text{Possible Values of \#2}} P(\#0 = \#0, \#2, L, E)\, P(\#1 = \#1, \#2, L, E) \ldots$$

Inferred Location

Sensor Measurements

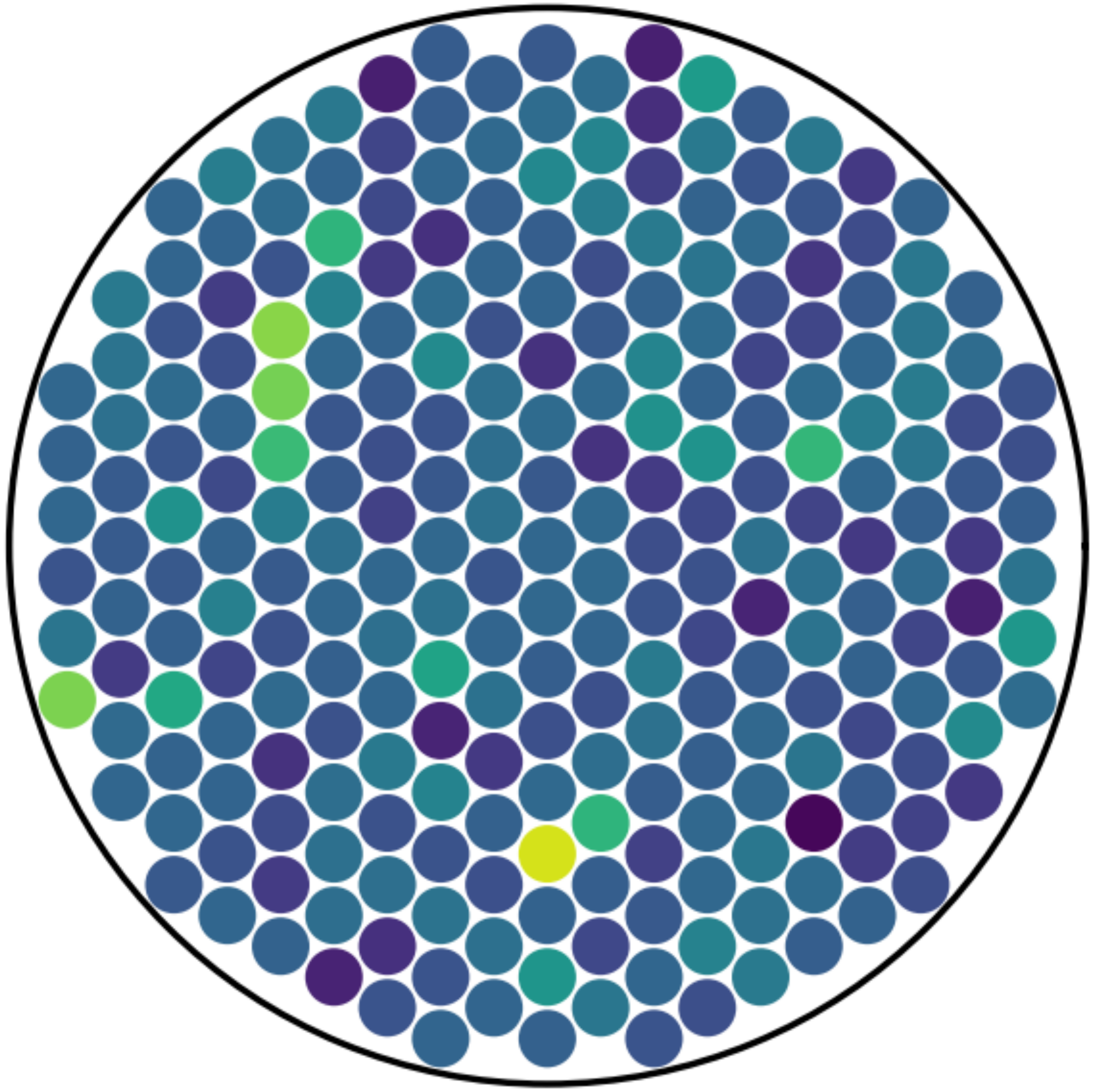


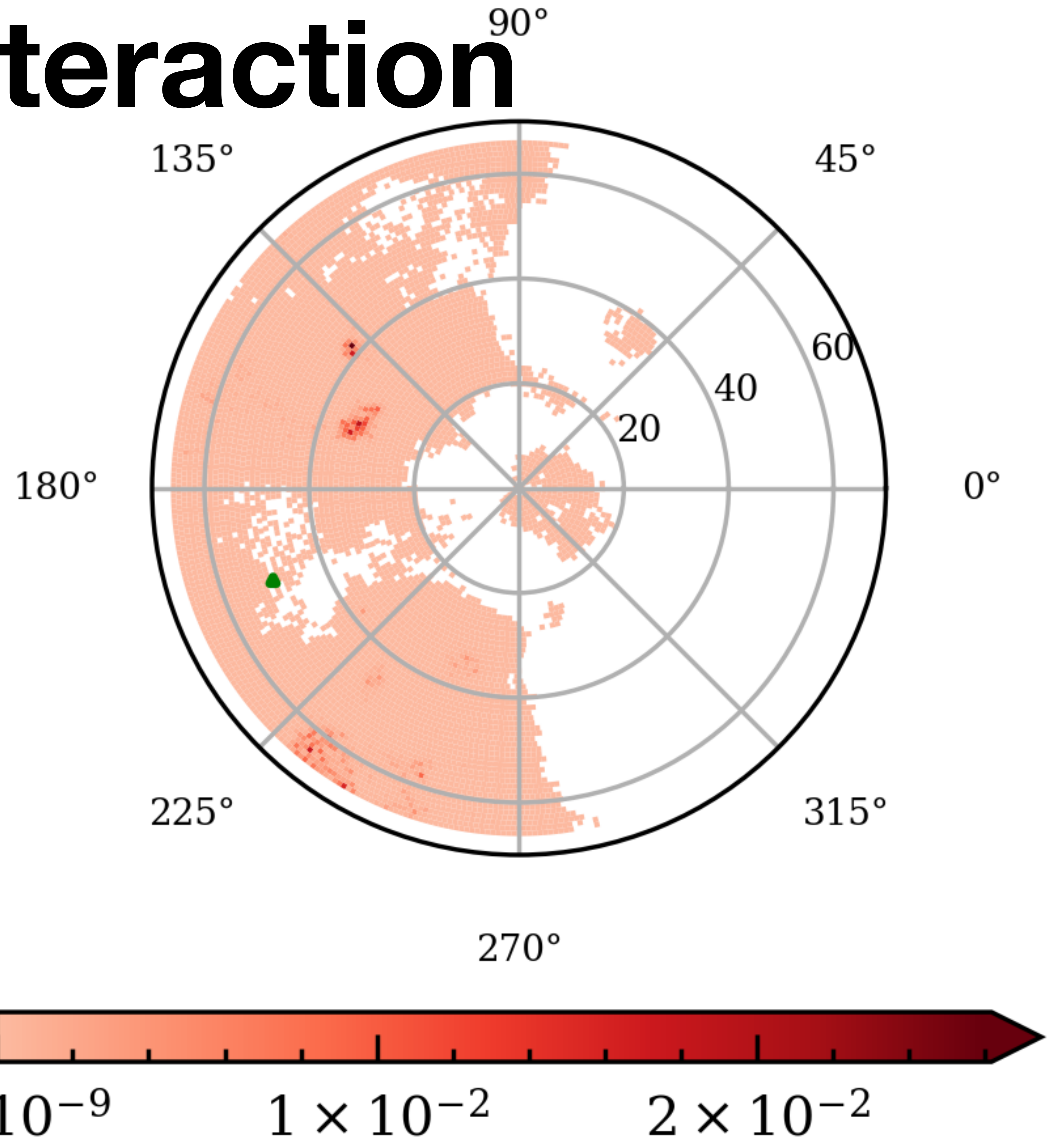


$$\sum_{\text{Possible Values of E}} P(L, E \mid \#0, \#1, \#2, \ldots, \#250, \#251, \#252)$$

Inferred Location - no interaction

Sensor Measurements

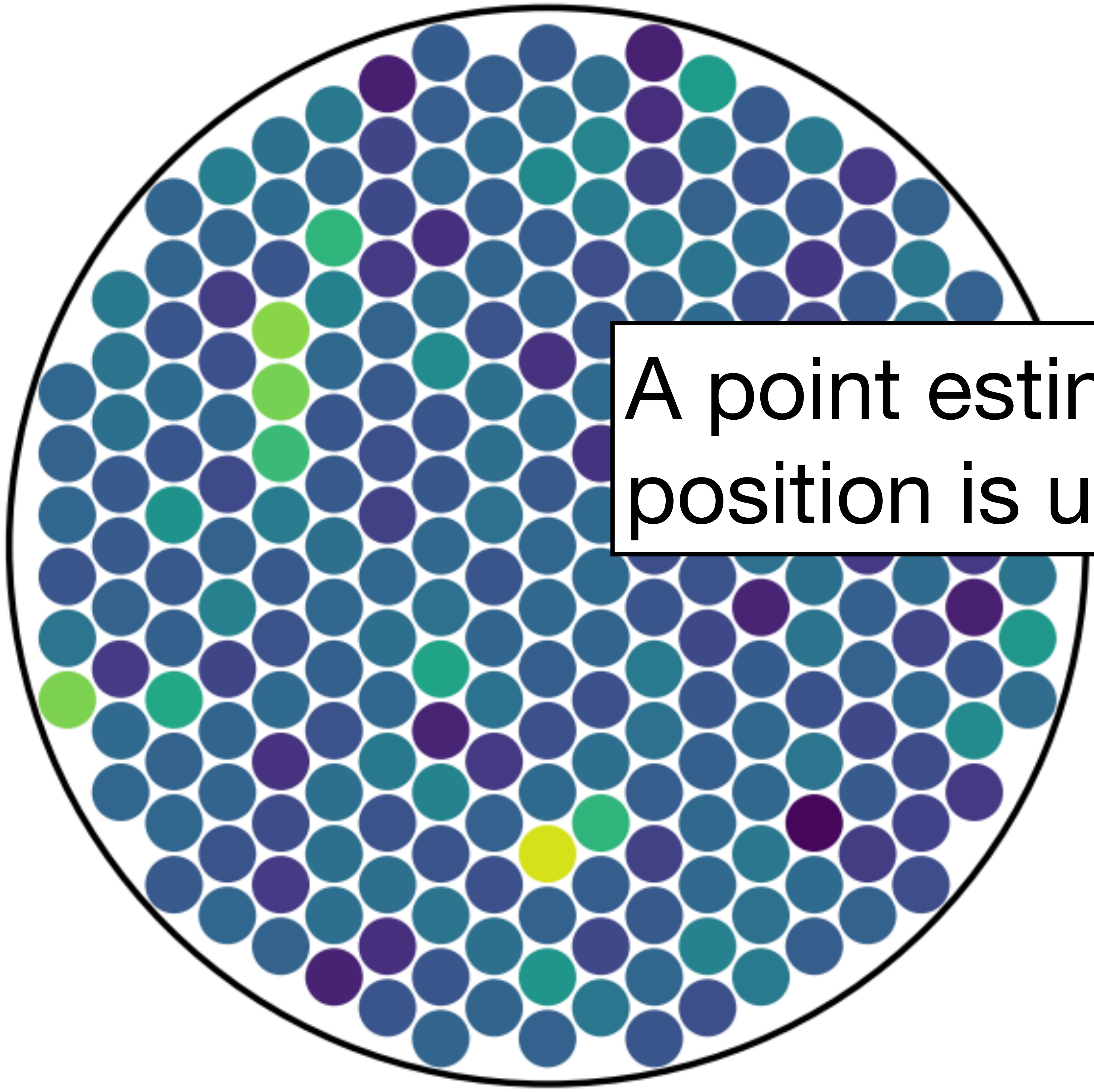




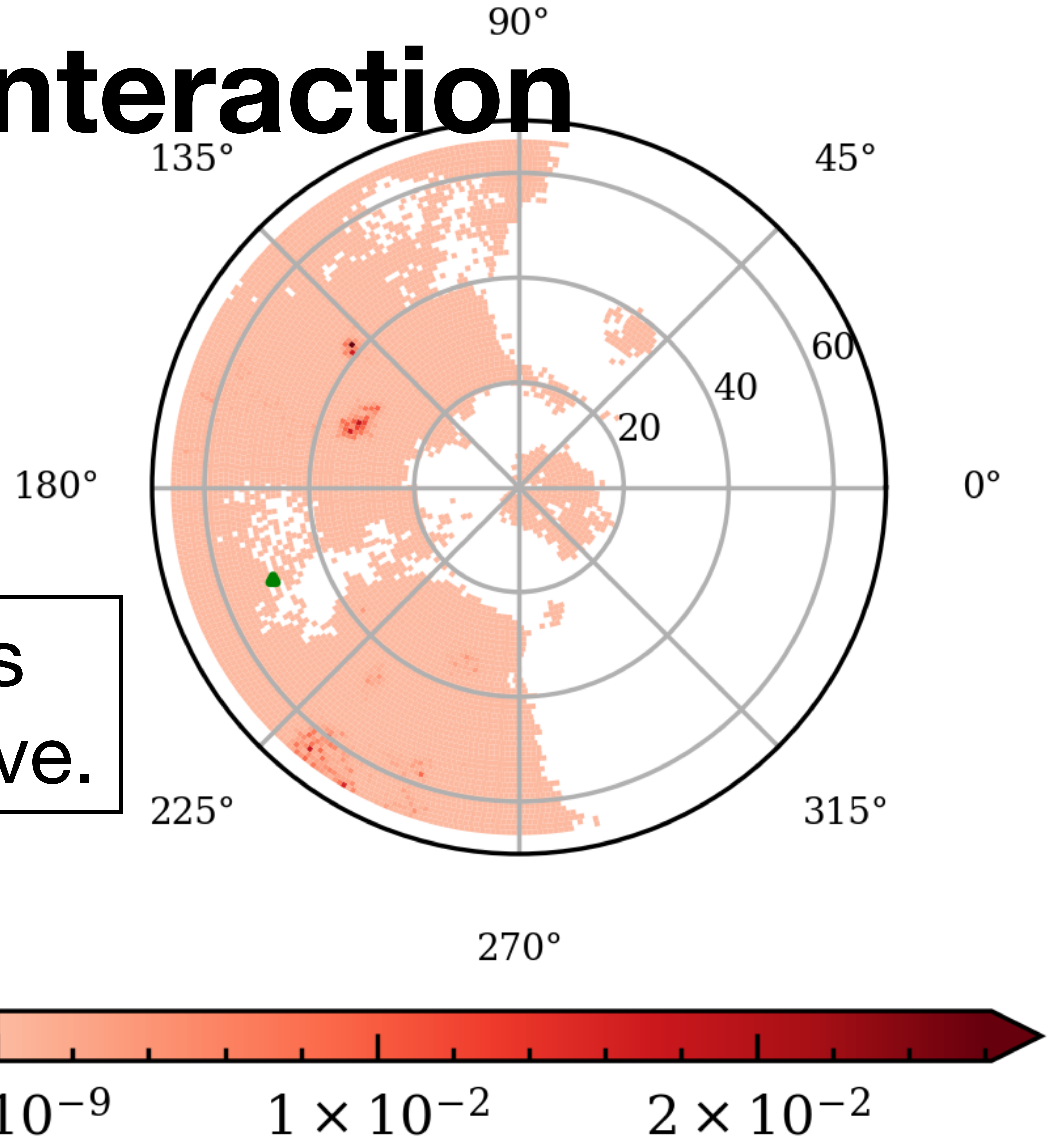
$$\sum_{\text{Possible Values of E}} P(L, E \mid \#0, \#1, \#2, \#250, \#251, \#252)$$

Inferred Positions - no interaction

Sensor Measurements

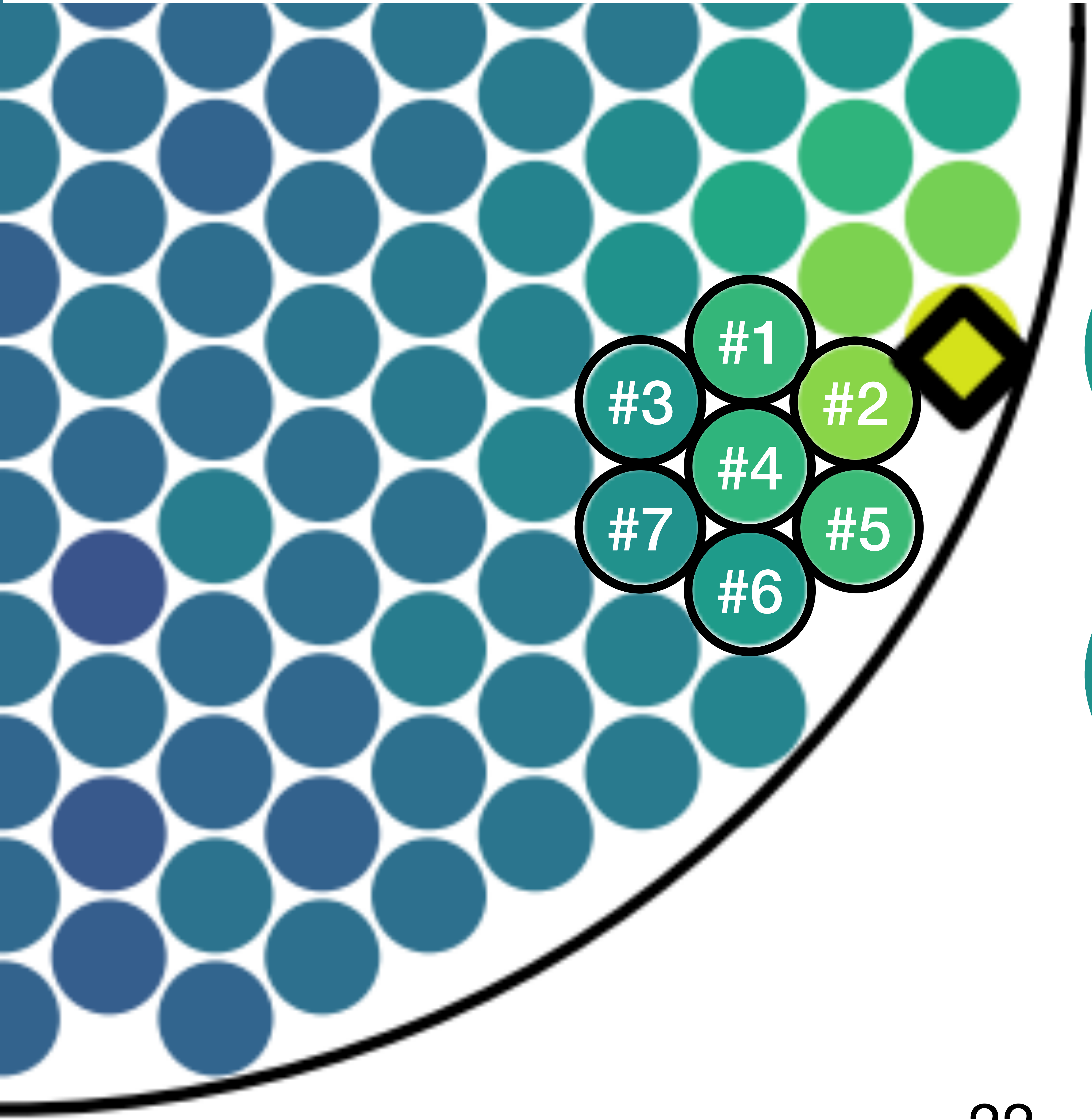


A point estimate of this position is uninformative.



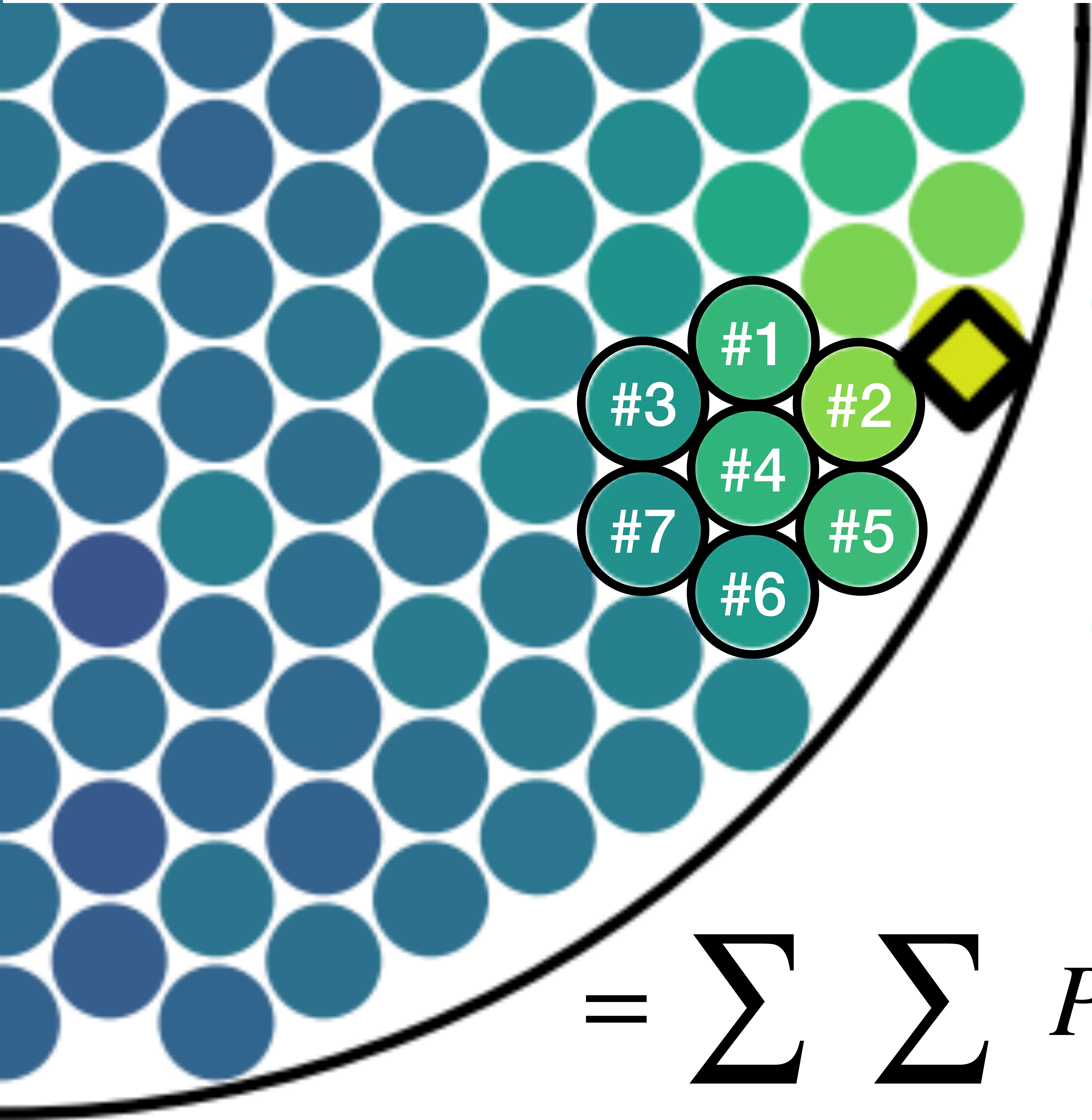
$$\sum_{\text{Possible Values of E}} P(\text{L}, \text{E} \mid \#0, \#1, \#2, \ldots, \#250, \#251, \#252)$$

Towards a More Complex Graph Structure



Zoë Bilodeau
IRIS-HEP Fellow

Towards a More Complex Graph Structure



Zoë Bilodeau
IRIS-HEP Fellow

$$= \sum_{\text{Possible Values of \#4}} \sum_{\text{Possible Values of \#6}} P(C, E, \#1, \#2, \#3, \#4, \#5, \#6, \#7)$$

When is this framework most applicable?

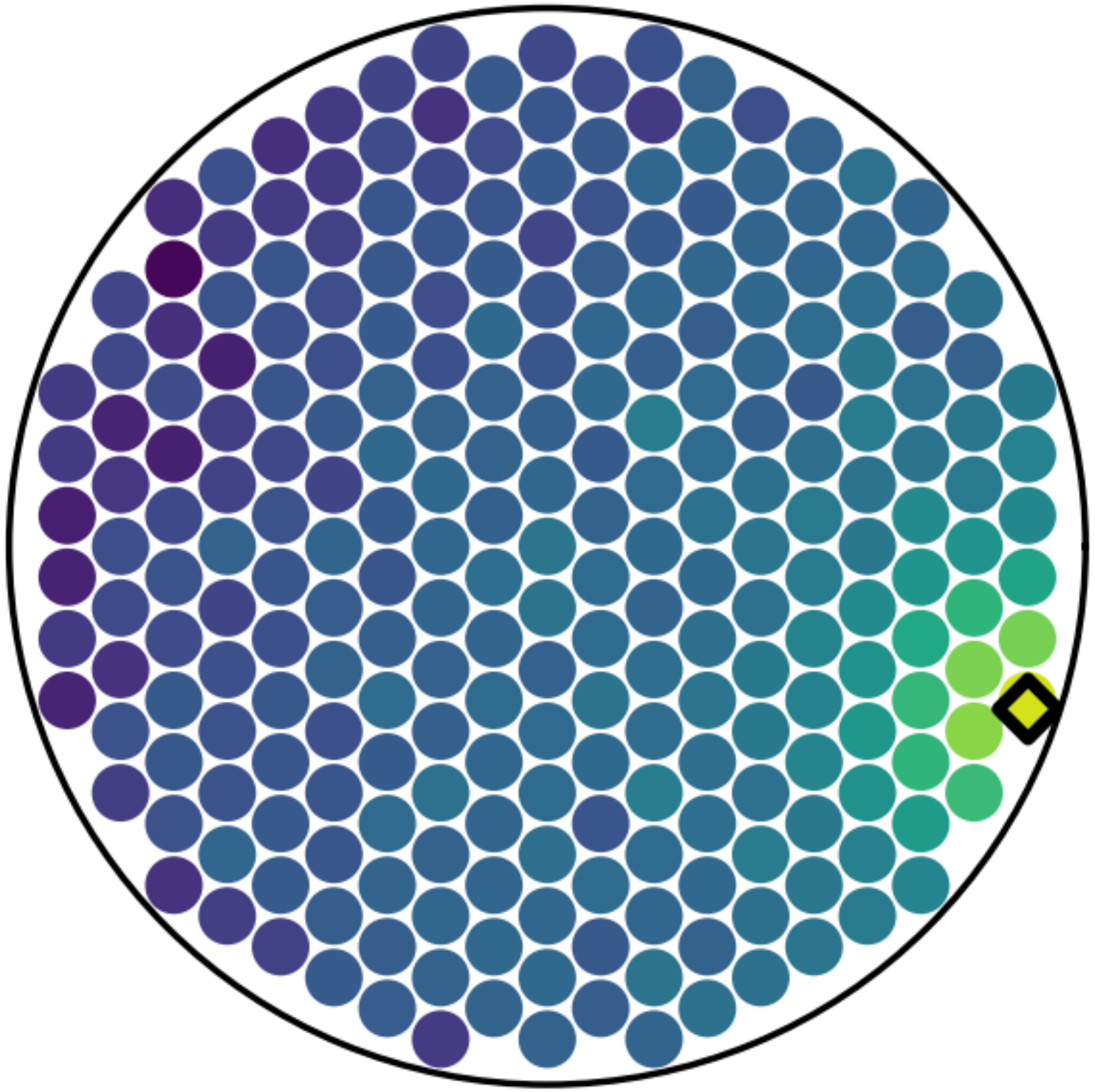
- uncertainty quantification and interpretability is paramount
- wish to constrain the model based on knowledge of the system
- set of variables is (relatively) small

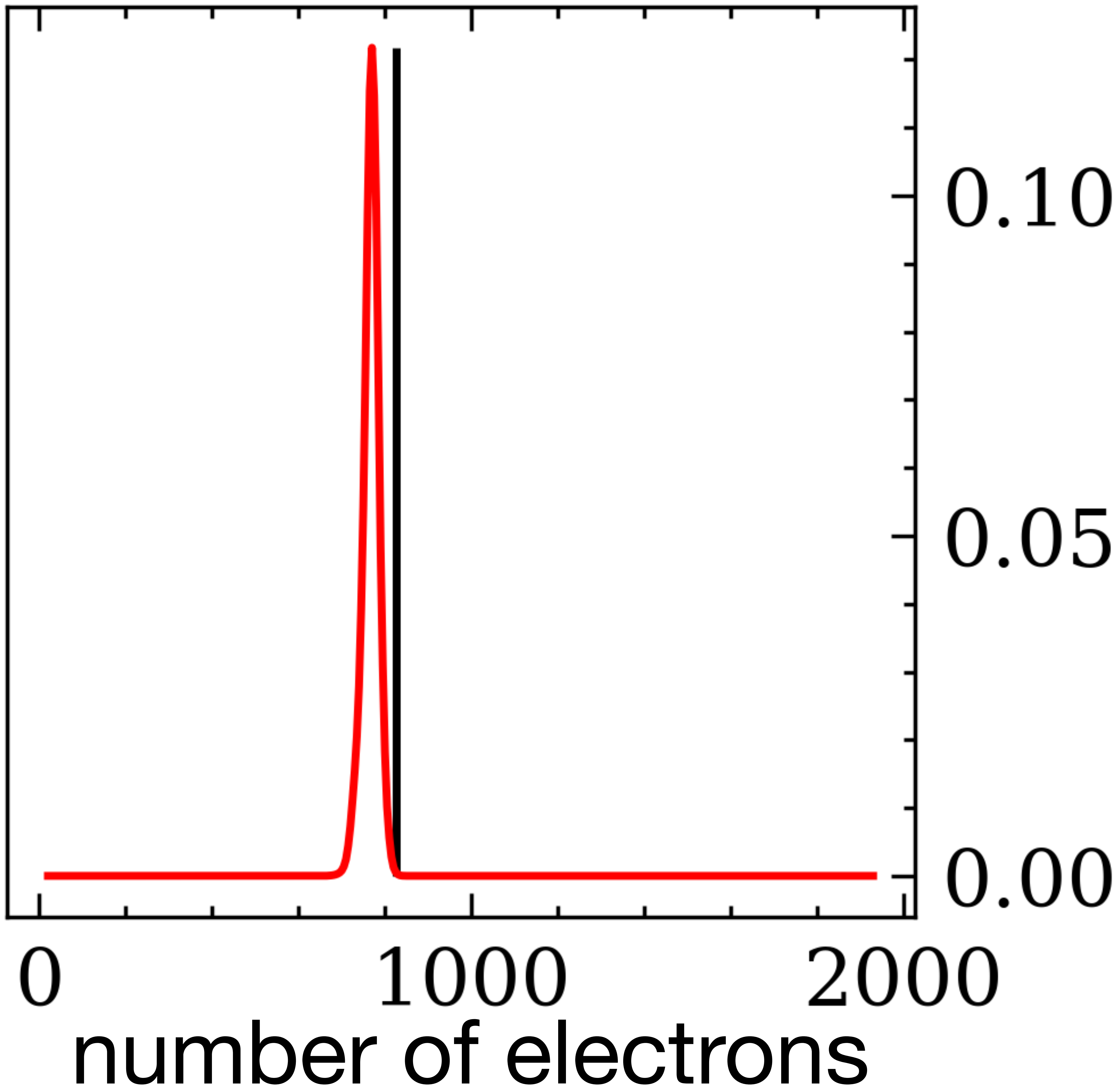
Summary

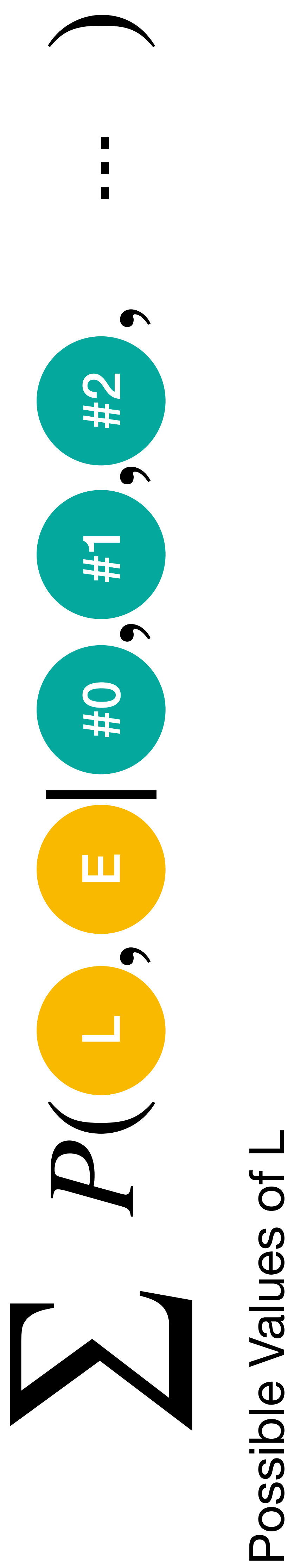
- Bayesian Networks are a
  - type of probabilistic graphical model
  - graph-based representation of the joint probability distribution as the basis for compactly encoding a high-dimensional distribution
  - simple, interpretable model that is designed for uncertainty quantification
    - *Probabilistic Graphical Models: Principles and Techniques*. Daphne Koller and Nir Friedman (2009).
- Bayesian network framework applicable to problems in particle physics
  - as demonstrated by an example with reconstruction of particle interactions in the context of dark matter direct-detection experiments
    - *Graphical Models are All You Need: Per-interaction reconstruction uncertainties in a dark matter detection experiment*. Christina Peters et al. (2022). NeurIPS Machine Learning and the Physical Sciences Workshop.
    - arXiv 2205.10305

Inferred Number of Electrons

Sensor Measurements







Inferred Number of Electrons - no interaction

Sensor Measurements