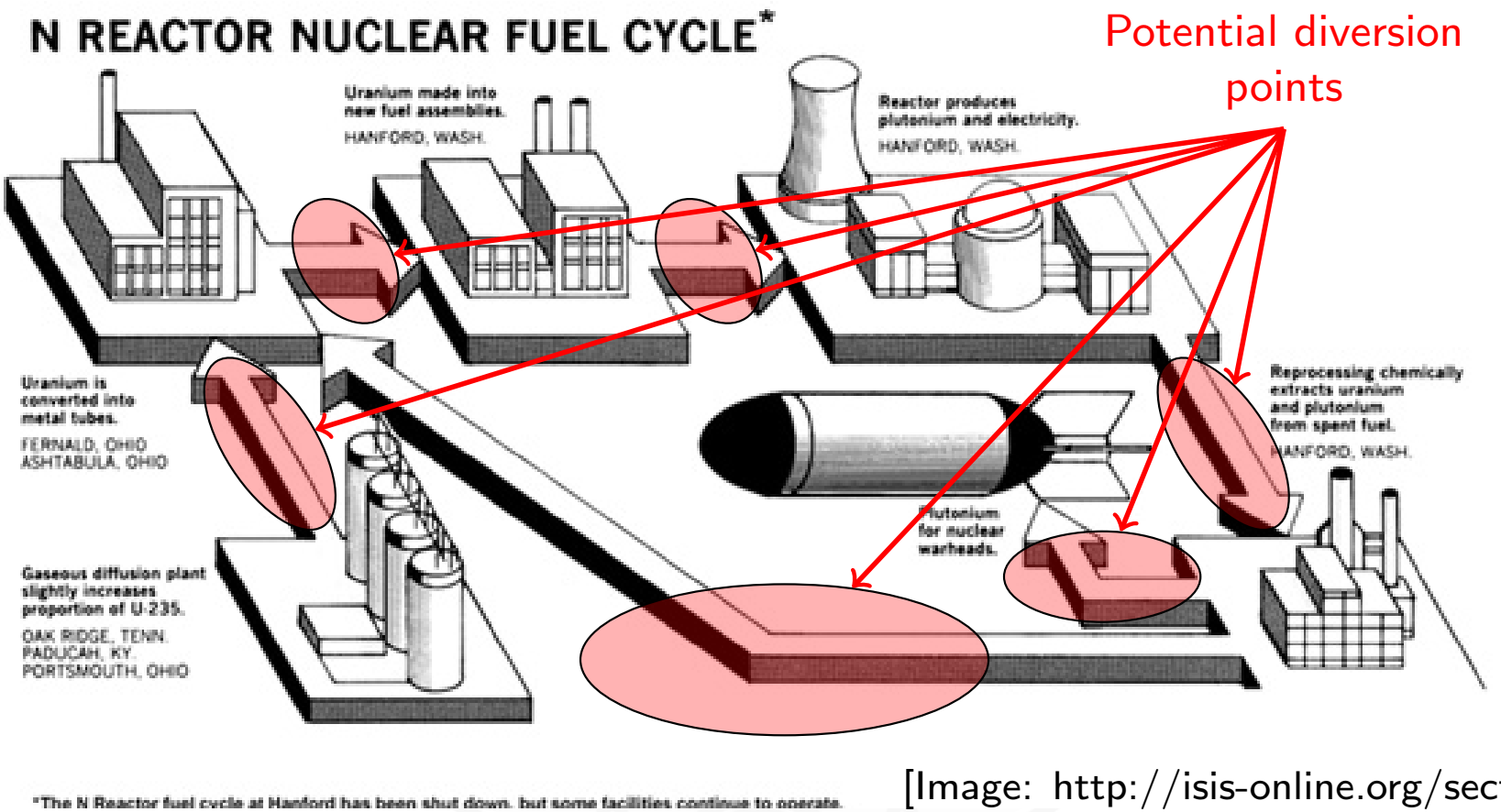
# Robust Monte Carlo Methods for Sequential Planning and Decision Making

Sue Zheng, Jason Pacheco, & John Fisher

Sensing, Learning, & Inference Group
Computer Science & Artificial Intelligence Laboratory
Massachusetts Institute of Technology

November 30, 2017

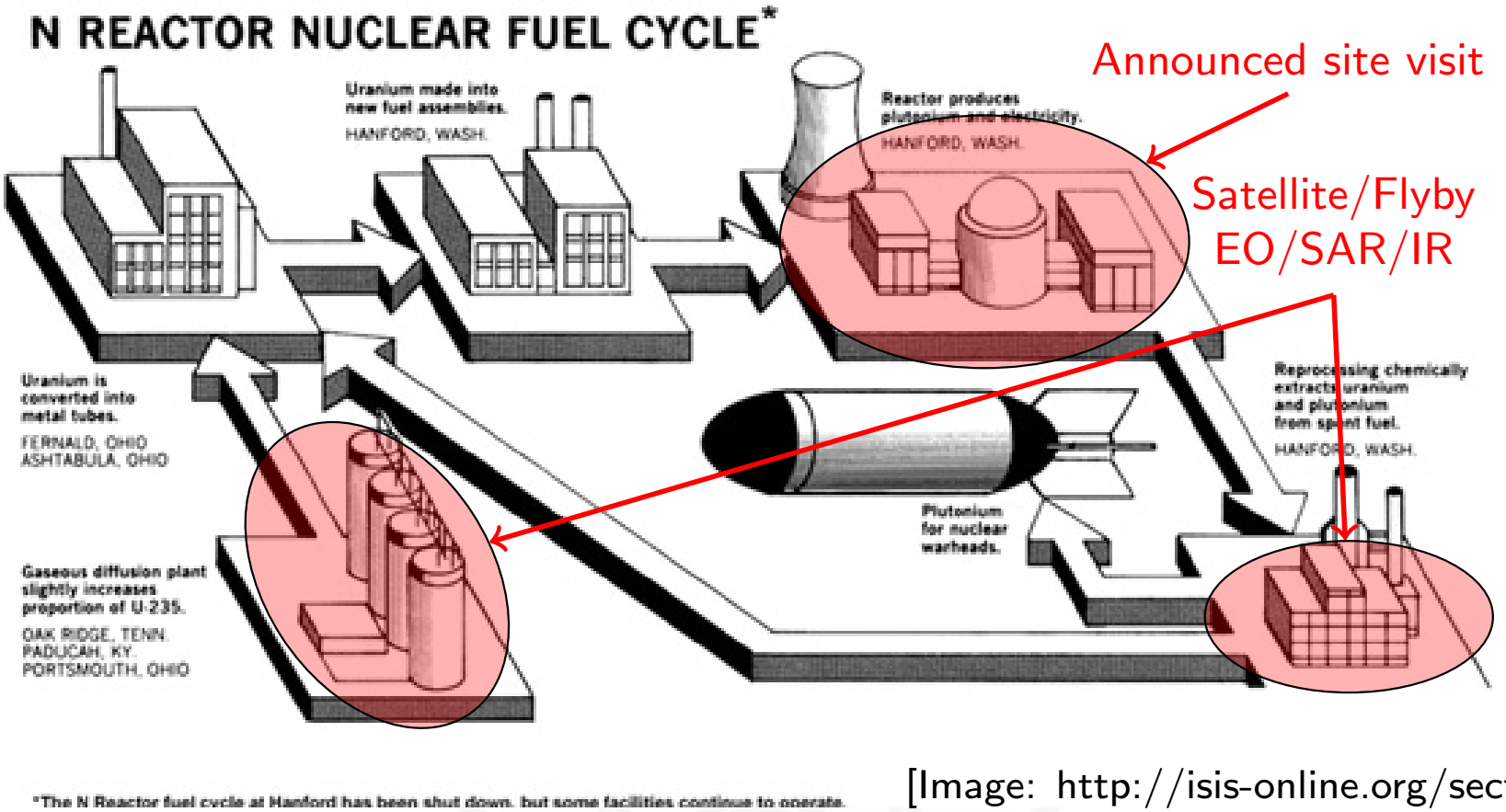
# Diversion Detection




[Image: http://isis-online.org/section-3]

Given sensor characteristics, observations, and ideal network, infer deviations (e.g., unknown network of material diversion).

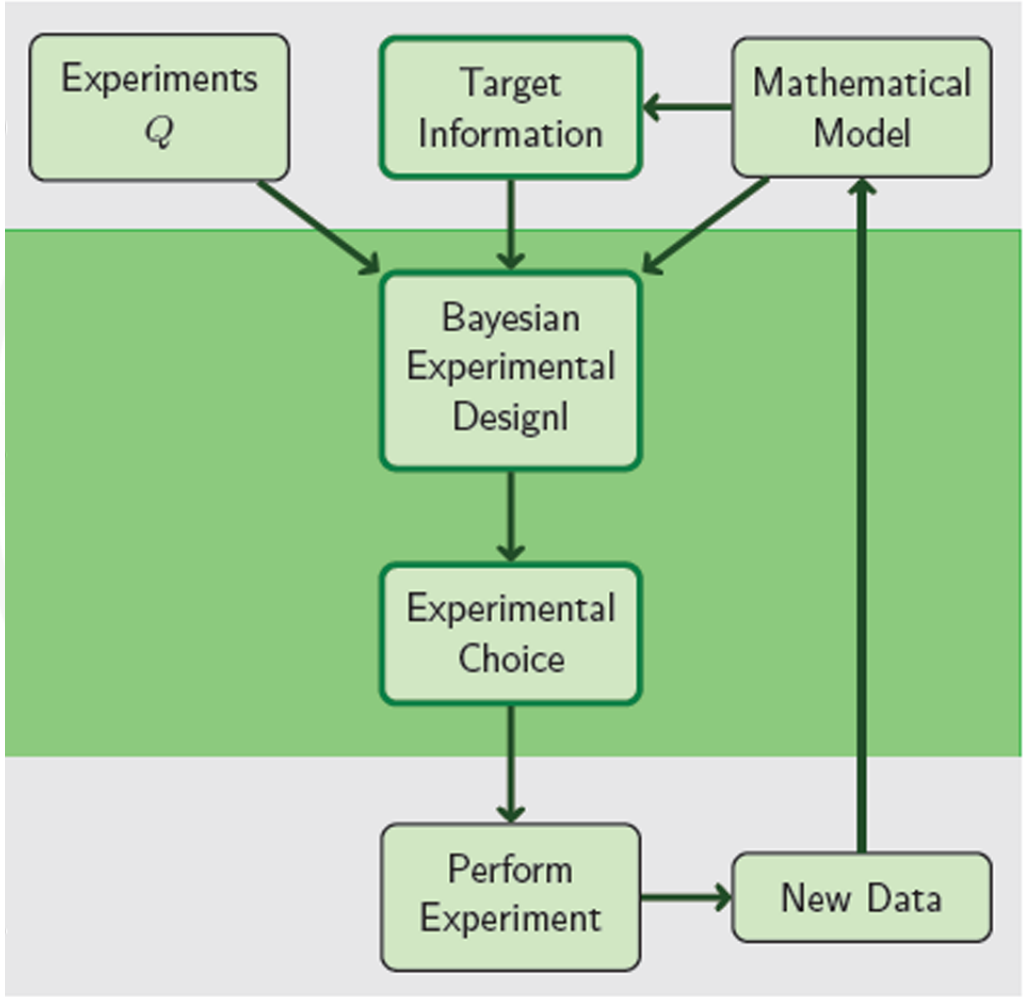
# Active Sensing and Decision Making




[Image: http://isis-online.org/section-3]

Announce site inspections to learn causal network structure.
Modify sensor configurations to reduce uncertainty of inferences.

# Sequential Experiment Design




[Image: Liepe et al., 2013]

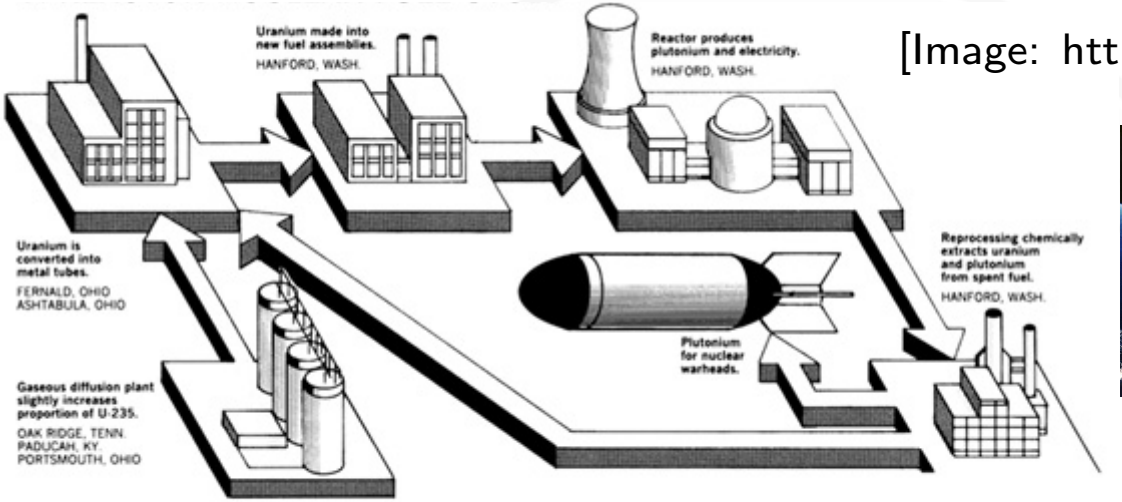
Experimental choice is a form of planning.

# Goal

**Develop algorithm for <u>sequential Bayesian inference and planning</u> with the following desiderata,**

- Information theoretic approach to planning
- Effective even for complex models
- Provable theoretical guarantees
- Maximally reuses computation in inference and planning phases

# Probabilistic Model & Inference




[Image: http://isis-online.org/section-3]

$\underbrace{\qquad\qquad\qquad\qquad}_{\text{Unknown } x}$ $\underbrace{\qquad\qquad}_{\text{Observation } y}$

Joint probability model: $p(x, y) = p(x)p(y|x)$

Prior belief in diversion network

Likelihood of observations given structure

Posterior belief in network structure given observations:

$$p(x|y) = \frac{p(x)p(y|x)}{p(y)}$$

# Decision-Driven Observations

| Unknown | Decision | Observation |
|---|---|---|
| $X$ | $d = 1$ | $Y \mid X \sim p(Y \mid X; d = 1)$ |
| | $d = 2$ | $Y \mid X \sim p(Y \mid X; d = 2)$ |
| | $\vdots$ | $\vdots$ |
| | $d = D$ | $Y \mid X \sim p(Y \mid X; d = D)$ |

Configuration variable $d = \{1, \ldots, D\}$ controls observation model.
E.g., announcing inspection of site A affects sensor observations.

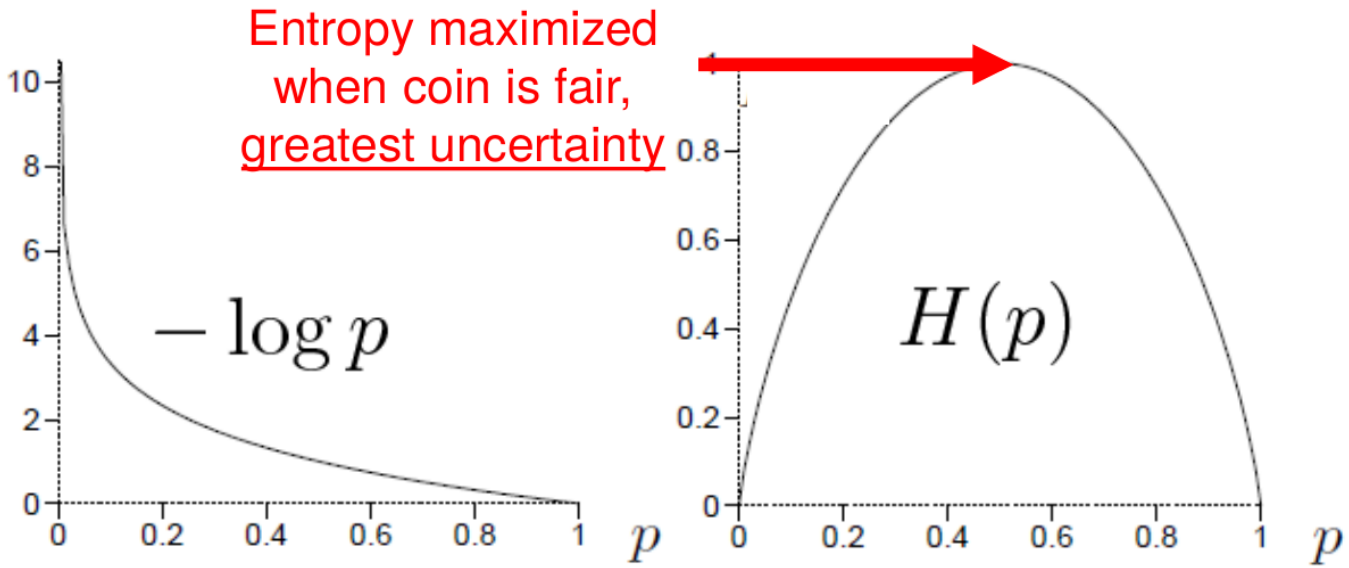
Choose configuration to **reduce posterior uncertainty**

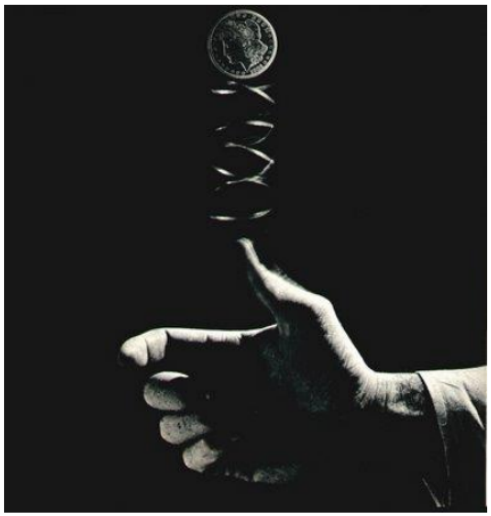
# Entropy

$$H(X) = \mathbb{E}[-\log p(X)]$$

Encodes uncertainty, more reliable than variance in many cases (multimodality)

**Coin flip example:** $X \sim \text{Bernoulli}(p)$



# Mutual Information

*What would be uncertainty of X if we knew Y?*

$$I(X;Y) = H(X) - H(X|Y)$$

Prior uncertainty

Expected posterior uncertainty

I will drop explicit dependence on configuration variable $d$ when clear from context:

$$I(X;Y|d) \Rightarrow I(X;Y)$$

# Information Theoretic Planning

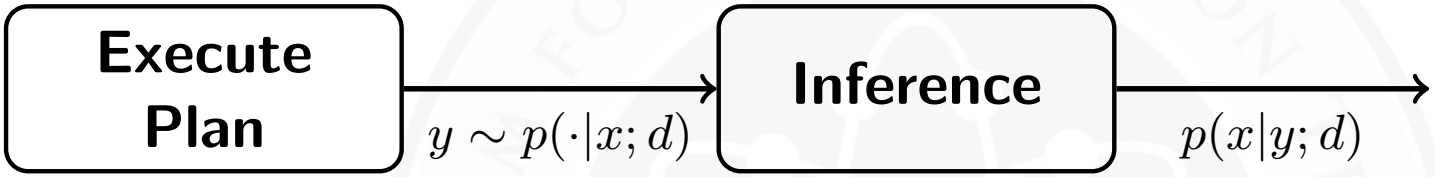
**Execute Plan** $\xrightarrow{y \sim p(\cdot|x;d)}$

- <u>Execute:</u> Draw new observation $Y = y$
- <u>Update:</u> Posterior belief

$$p(X|y) \propto p(X)p(y|X)$$

- <u>Select:</u> Information gain

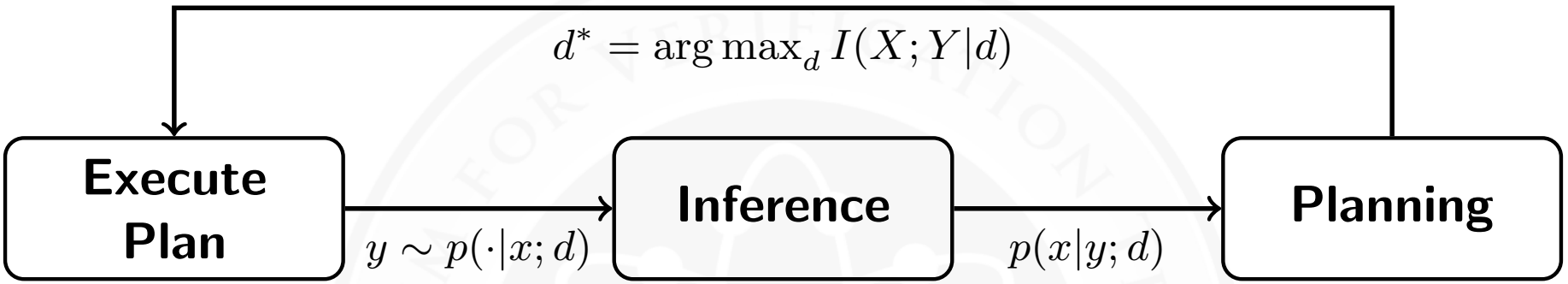
# Information Theoretic Planning




- Execute: Draw new observation $Y = y$
- Inference: Update posterior belief

$$p(X|y) \propto p(X)p(y|X)$$

# Information Theoretic Planning



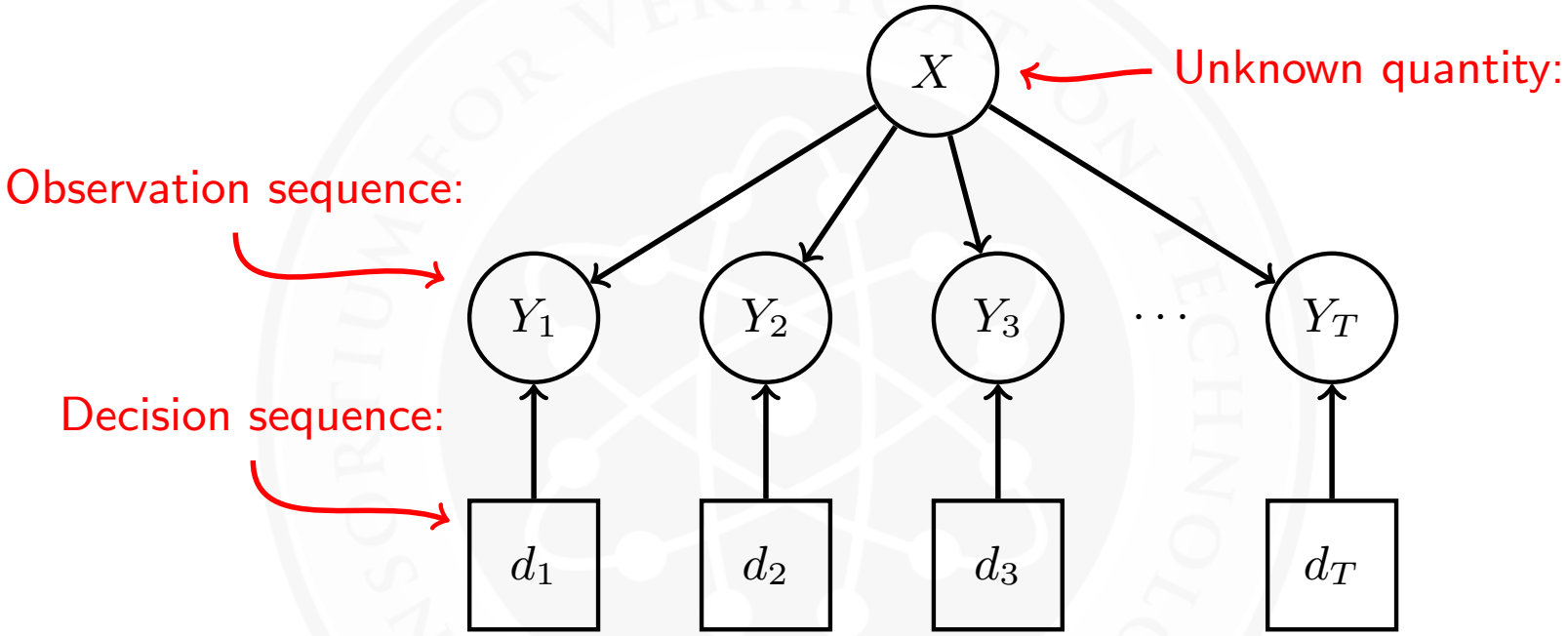

- Execute: Draw new observation $Y = y$
- Inference: Update posterior belief

$$p(X|y) \propto p(X)p(y|X)$$

- Planning: Choose most informative decision

$$d^* = \arg\max_d I(X;Y|d)$$

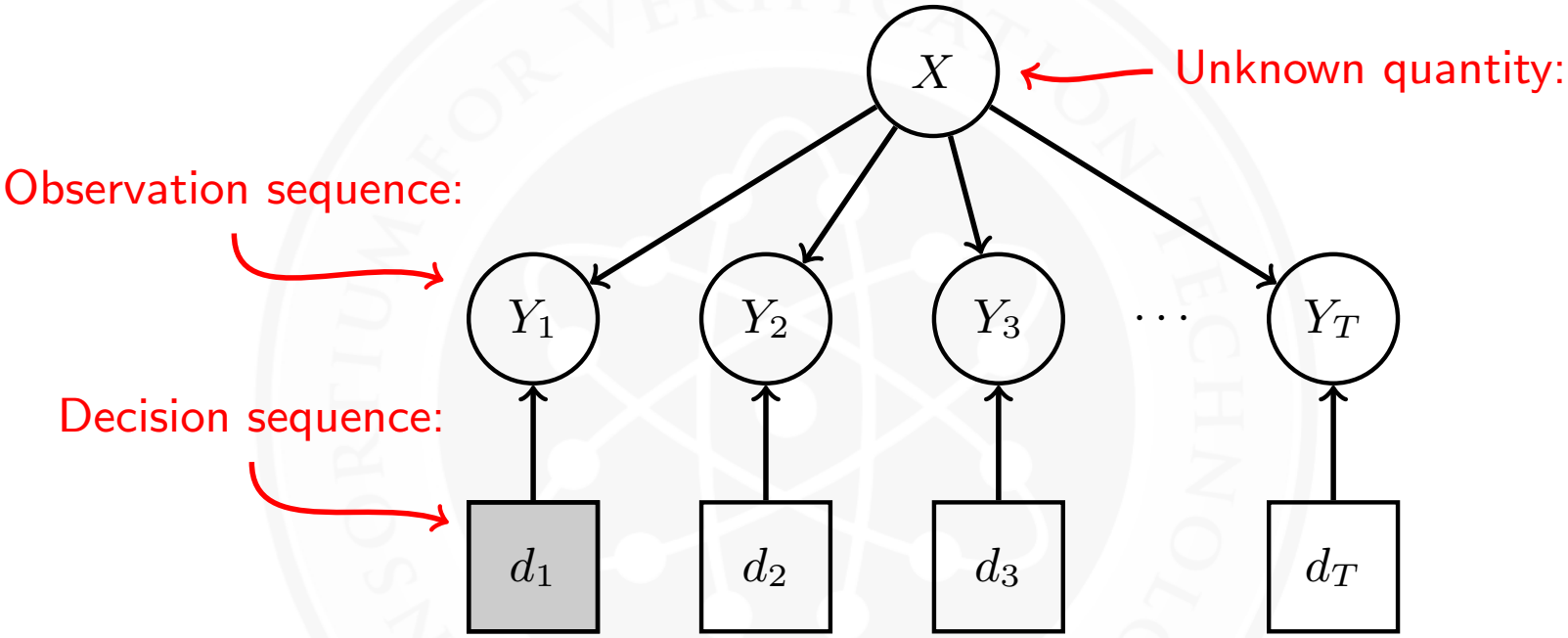
# Closed-Loop Greedy Planning




Condition on observed information during planning

$$d_t^{\text{greedy}} = \arg\max_d I(X; Y_t | y_1^{t-1}; d)$$

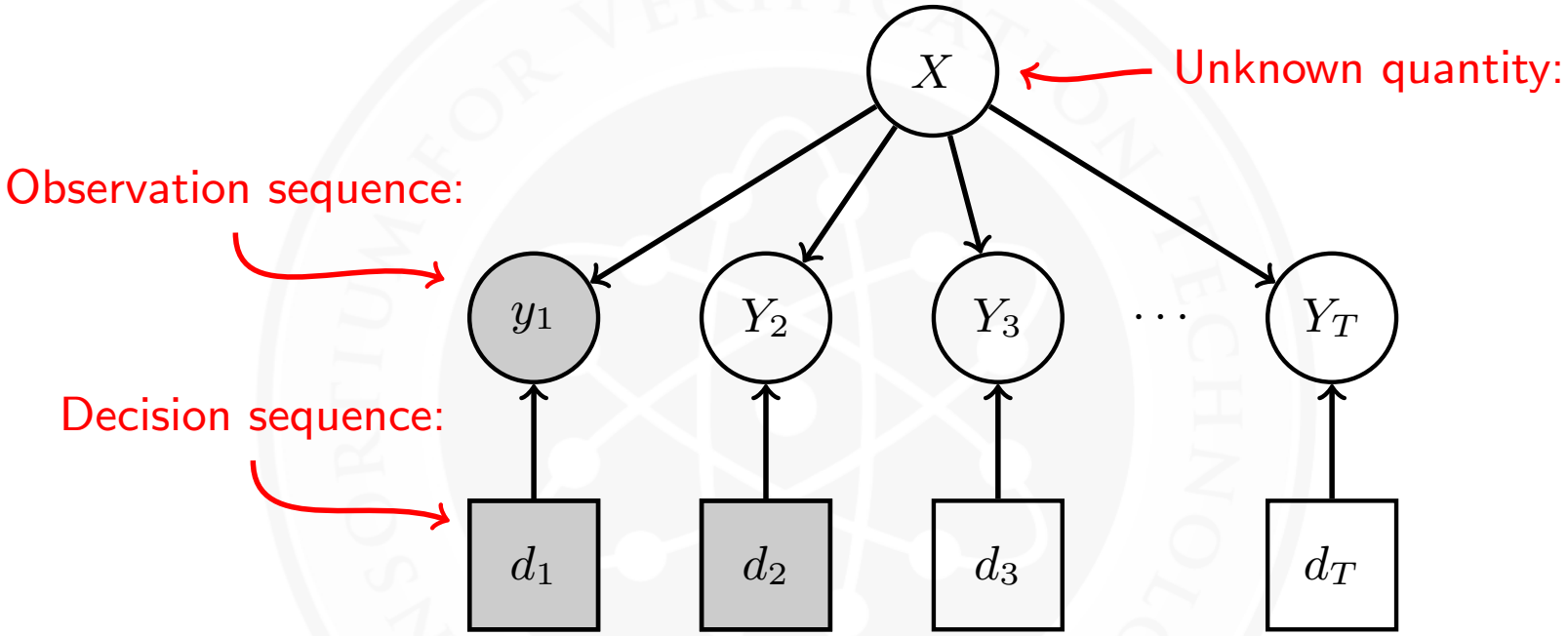
# Closed-Loop Greedy Planning




Condition on observed information during planning

$$d_t^{\text{greedy}} = \arg\max_d I(X; Y_t | y_1^{t-1}; d)$$

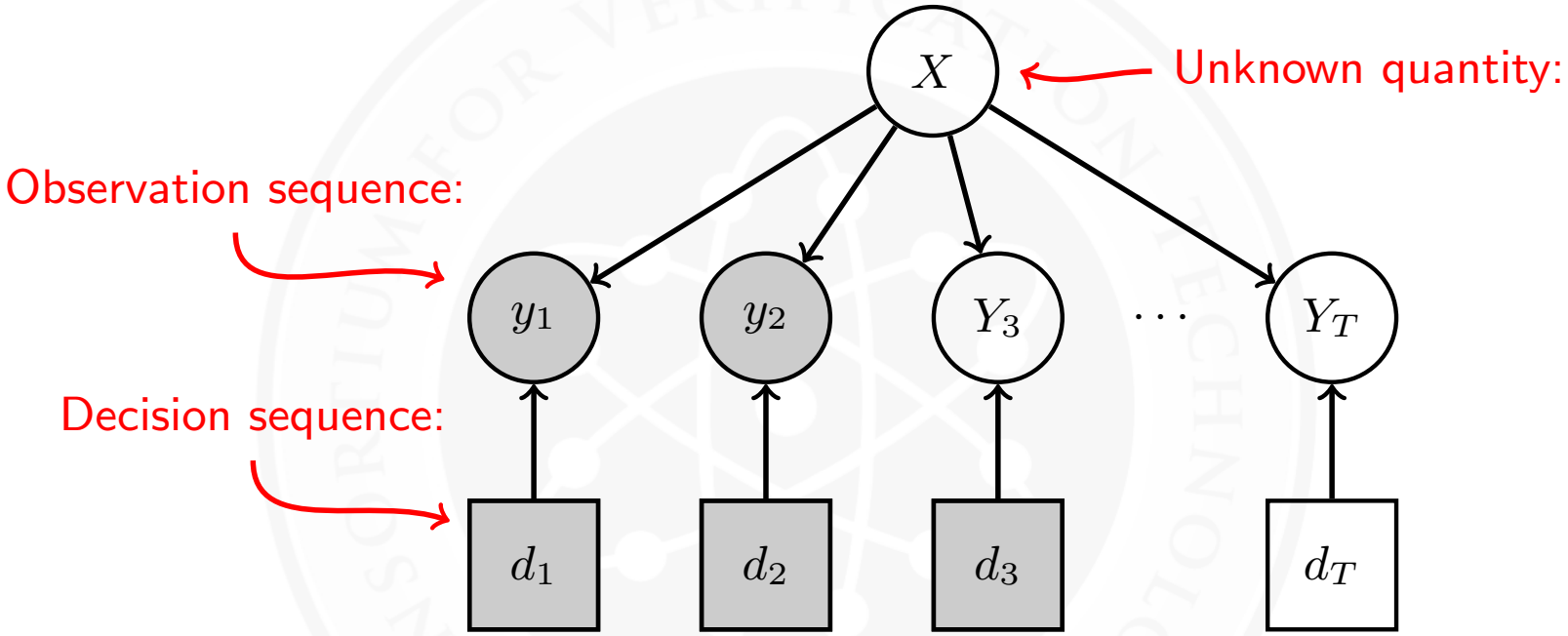
# Closed-Loop Greedy Planning




Condition on observed information during planning

$$d_t^{\text{greedy}} = \arg\max_d I(X; Y_t | y_1^{t-1}; d)$$

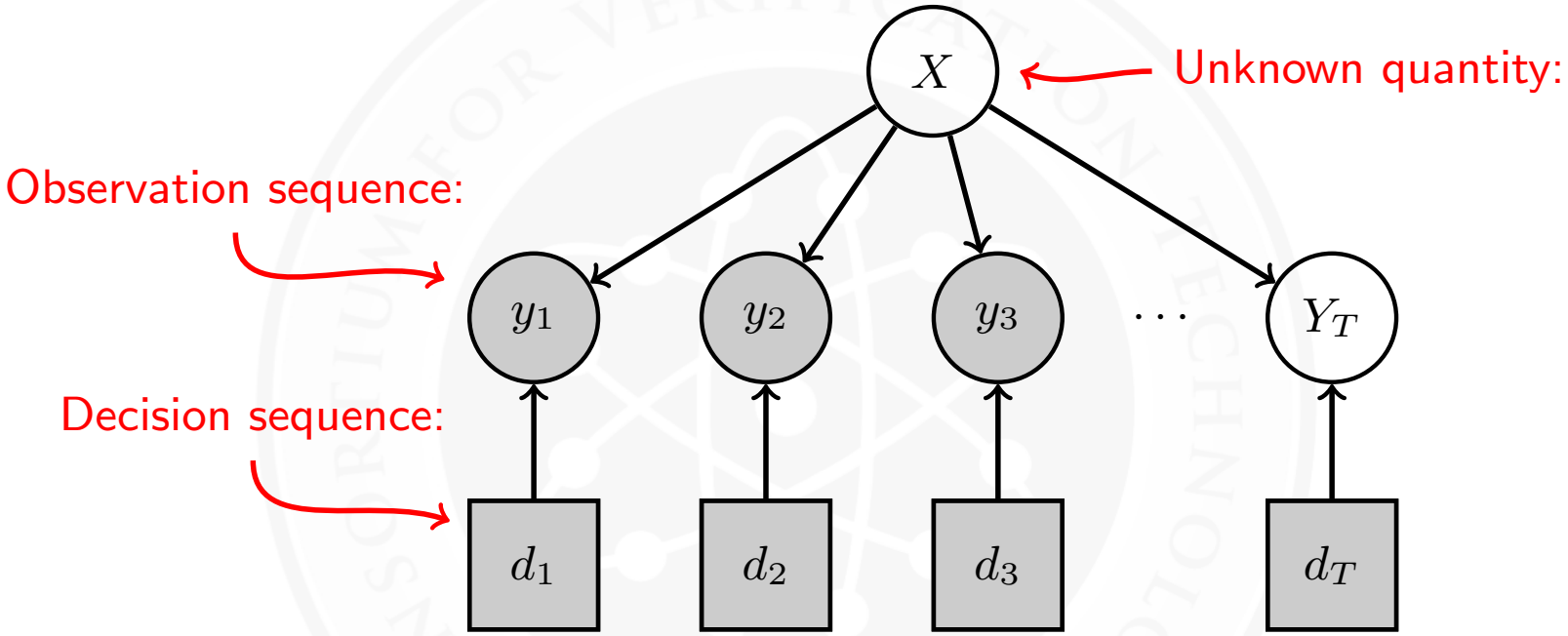
# Closed-Loop Greedy Planning




Condition on observed information during planning

$$d_t^{\text{greedy}} = \arg\max_d I(X; Y_t | y_1^{t-1}; d)$$

# Closed-Loop Greedy Planning




Condition on observed information during planning

$$d_t^{\text{greedy}} = \arg\max_d I(X; Y_t | y_1^{t-1}; d)$$

# Estimating Mutual Information

Mutual information typically lacks closed-form:

$$I(X;Y) = \mathbb{E}\left[\log \frac{p(x,y)}{p(x)p(y)}\right]$$

Mutual Information

Evidence

Evidence integrates every latent configuration:

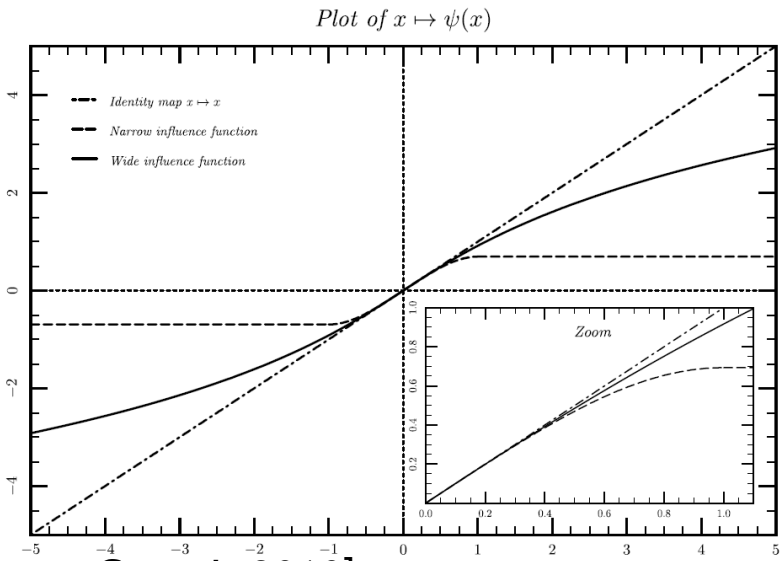
$$p(y) = \int p(x,y)dx$$

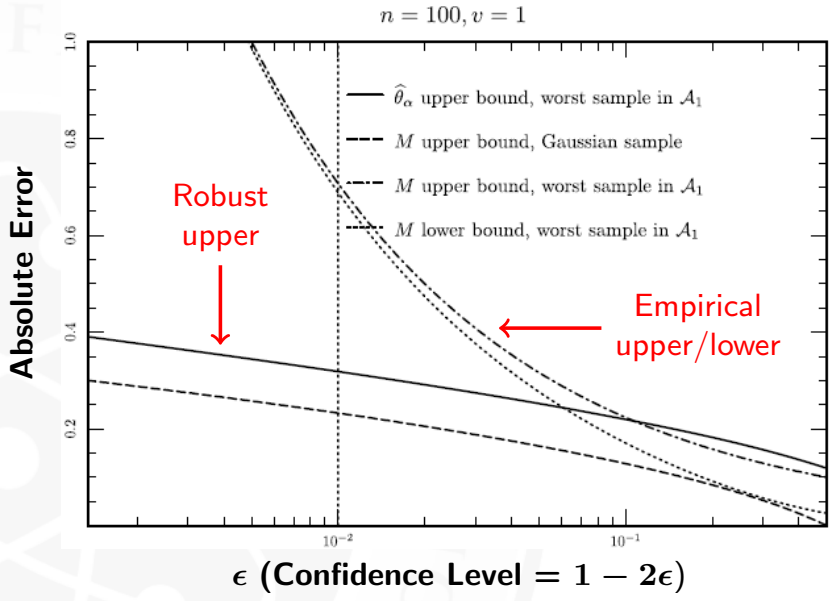
Can use Monte Carlo integration to estimate integrals from joint samples:

$$\{x_i, y_i\}_{i=1}^{N} \sim p(x,y)$$

Empirical estimate is sensitive to outliers in small sample regime

# Robust Estimation of Information



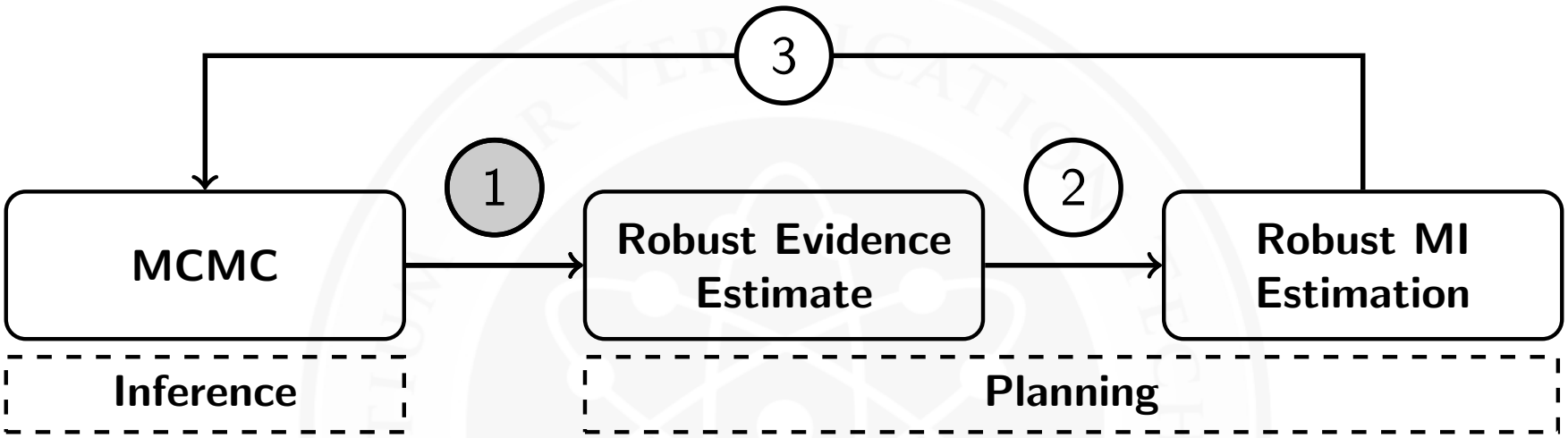

[Images: Catoni, 2010]

- Robust M-estimator is solution to root equation

$$\sum_i \psi(\alpha(\theta_i - \hat{\theta})) = 0 \text{ where } \theta_i = \log \frac{p(x_i, y_i)}{p(x_i)p(y_i)}$$

- Influence function $\psi$ reduces impact of outliers
- Provides quality guarantee in finite sample setting

# Integrated Inference & Planning



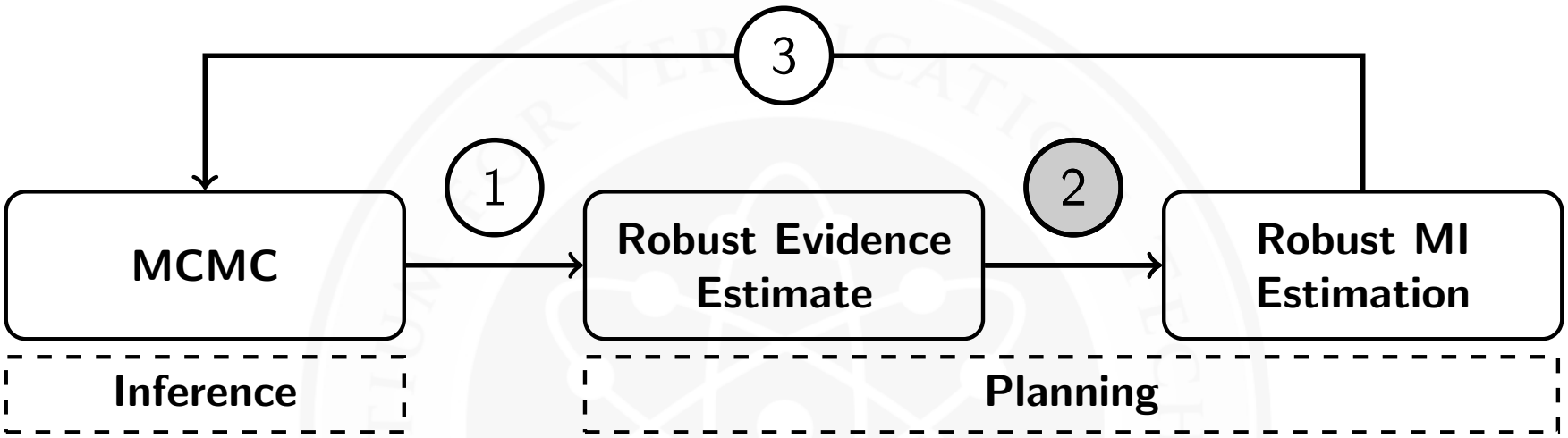

At time $t$ execute plan and draw observation:

$$y_t \sim p(y|x; d_t^{\text{greedy}})$$

Do posterior inference via MCMC samples

$$\{x_i\}_{i=1}^N \sim p(x|y_1^t)$$

# Integrated Inference & Planning




For each decision draw samples:

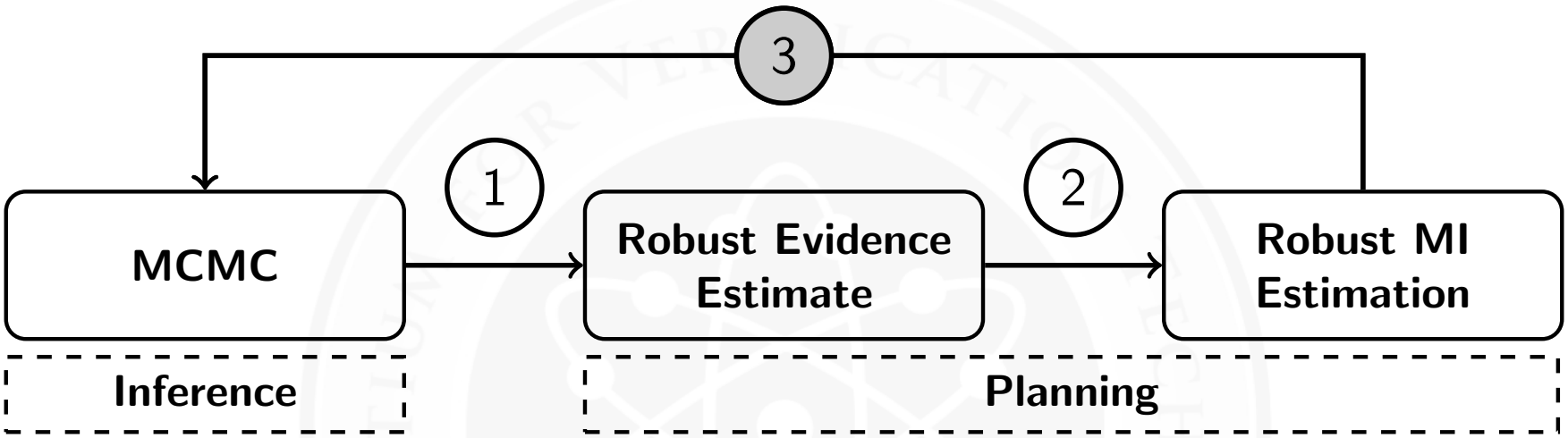
$$\{x_i\}_{i=1}^{N} \sim p(x|Y_{t+1}, y_1^t; d)$$

Robust estimation of model evidence:

$$\hat{p}(Y_{t+1}|y_1^t; d)$$

Can reuse MCMC samples with importance sampling

# Integrated Inference & Planning



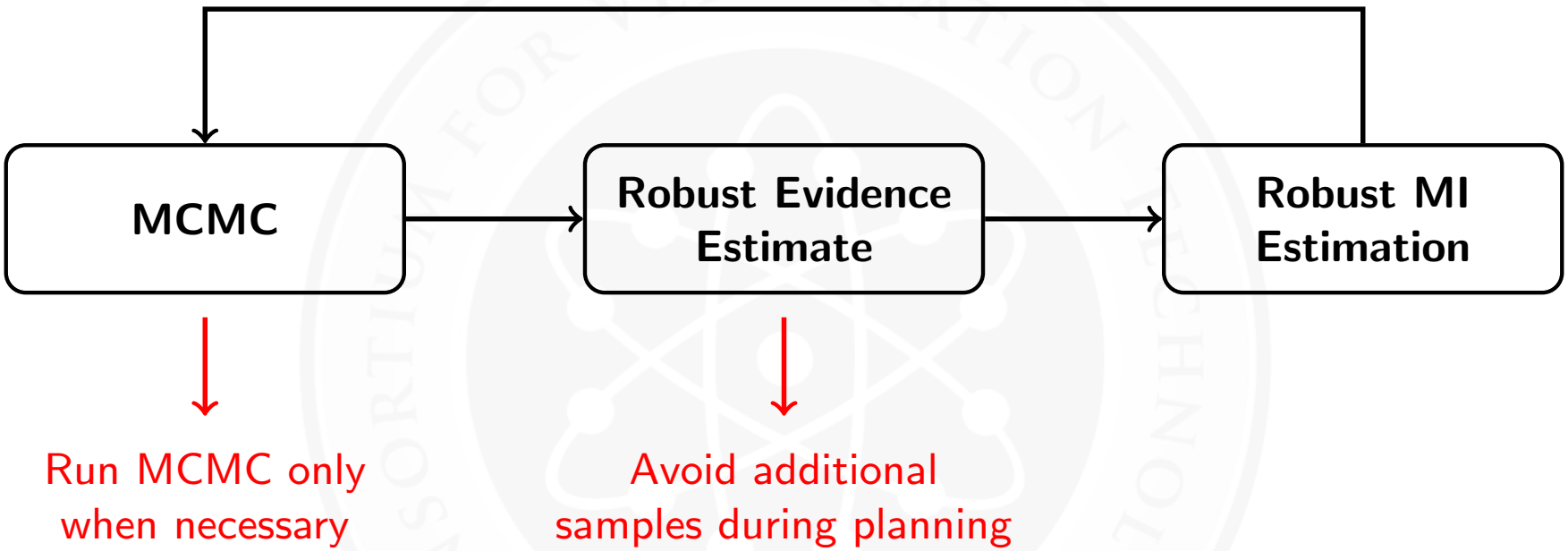

For each decision robust MI estimate:

$$\hat{I}(X; Y_{t+1}|y_1^t; d)$$

Greedy planning:

$$d_{t+1}^{\text{greedy}} = \arg\max_d \hat{I}(X; Y_{t+1}|y_1^t; d)$$

# Sequential Inference & Planning




Significantly reduces computation in inference and planning stages through sample reuse.

Algorithmic details are similar to a particle filter...

# Asymptotic Results

Theoretical bounds on estimators ensure high quality decisions.

$$\sqrt{N}(\hat{H} - H) \xrightarrow{d} \mathcal{N}(0, \cdot)$$

Establish central limit theorem as number of samples $N \to \infty$.

Show that estimators are consistent and approximately Normal with variance $\Theta(\frac{1}{N})$.

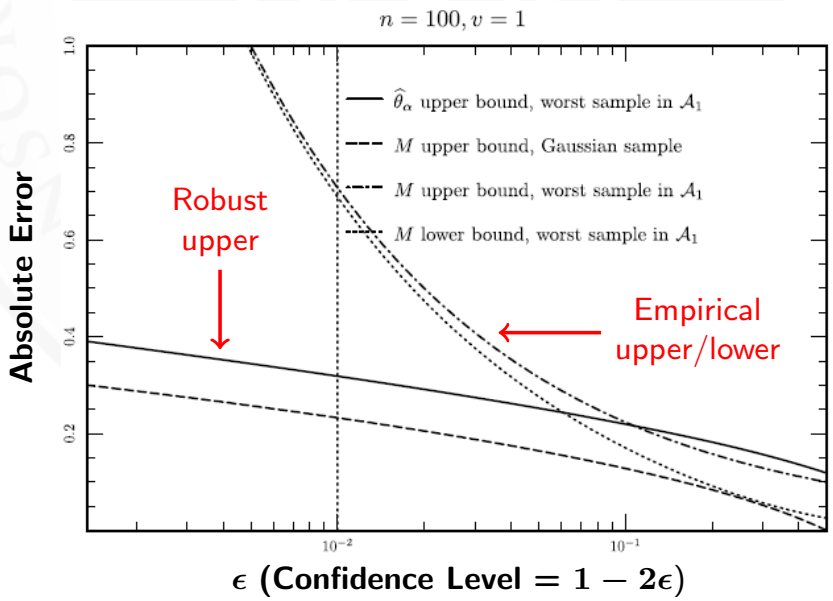
# Finite Sample Bounds

In practice, finite sample bounds are preferred over asymptotic results.

$$H(Y) + b - \text{const} < \hat{H}_Y < H(Y) + b + \text{const} \qquad \text{w.p.} \quad \geq 1 - 2\epsilon$$

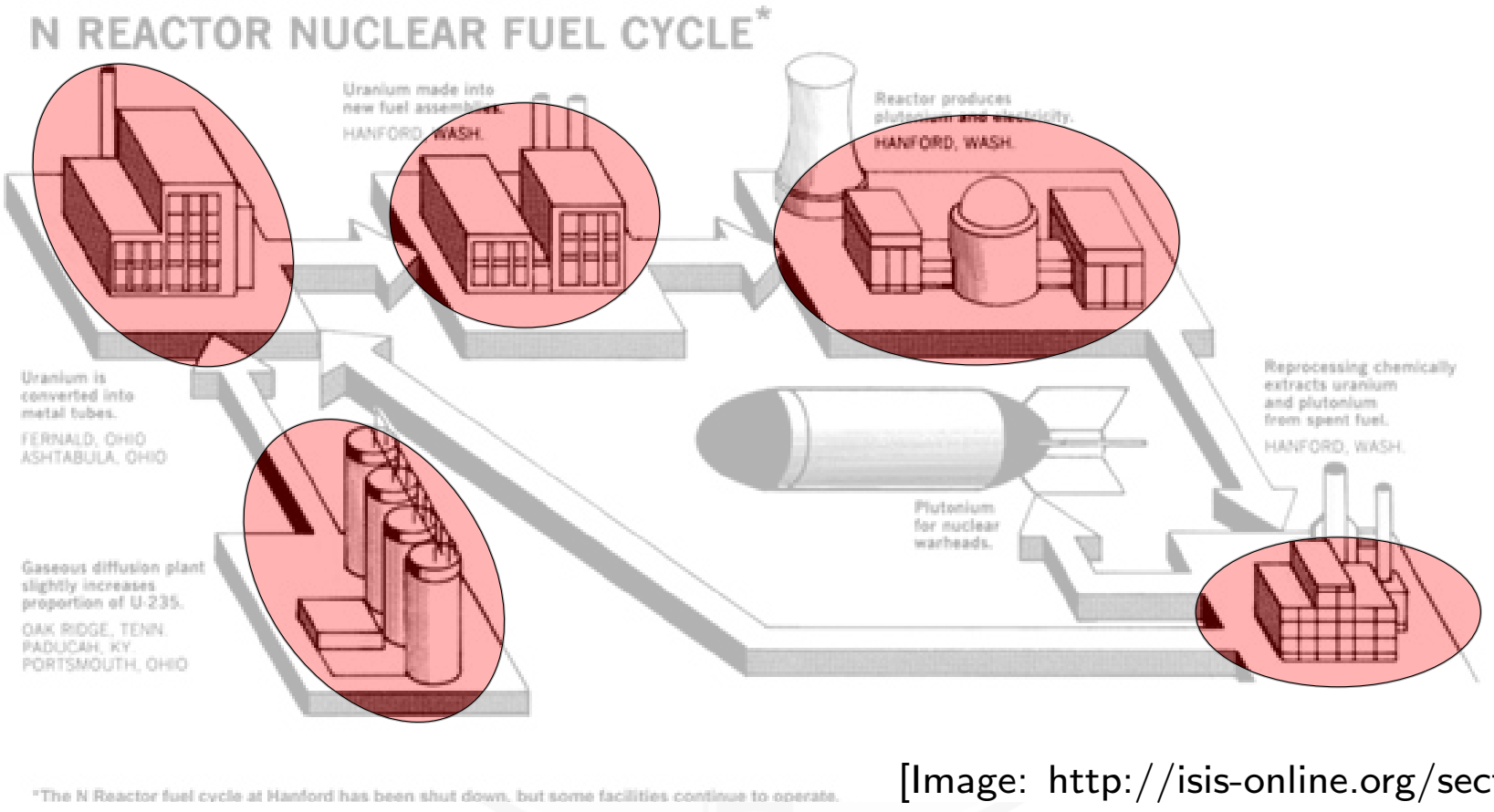
$$\text{where } b = \text{KL}(p(Y)||\hat{p}(Y))$$

Estimates are biased but deviation is bounded w.h.p. through use of robust estimator.




[Image: Catoni, 2010]

# Diversions in Nuclear Fuel Cycle



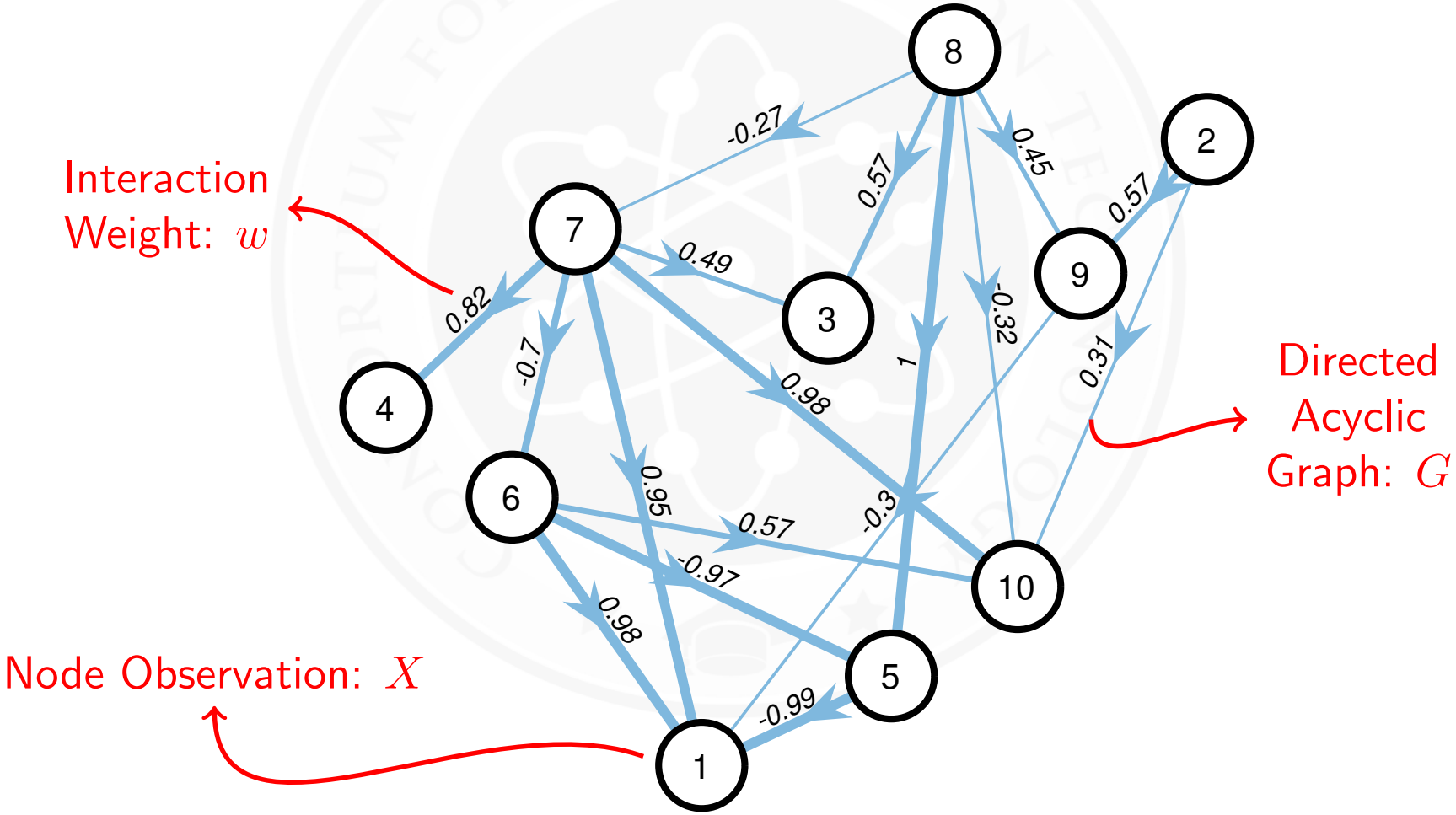

[Image: http://isis-online.org/section-3]

Identify sites for inspection announcement.

Performing an intervention to learn causal network structure.

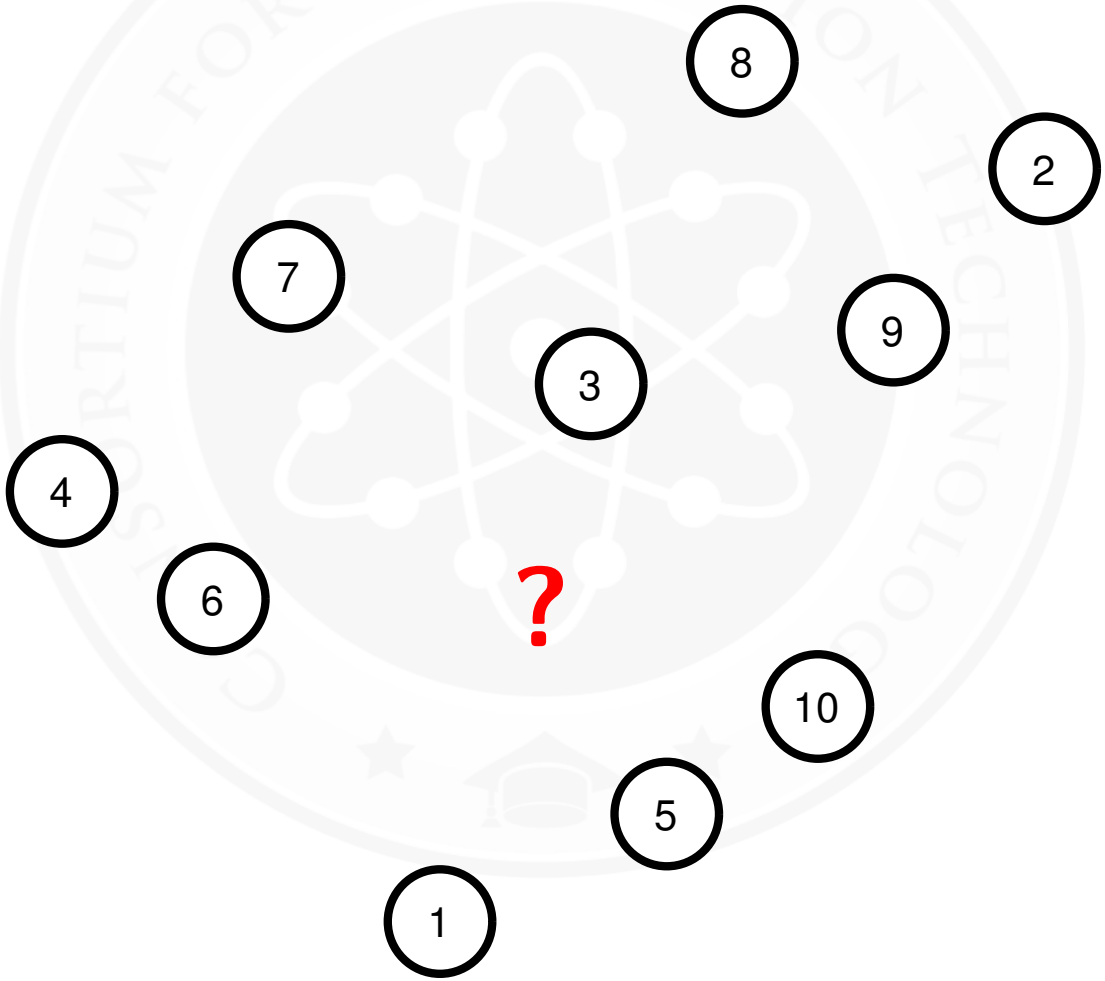
# Causal Network Inference

Nodes interact linearly according to *directed acyclic graph (DAG)* structure.



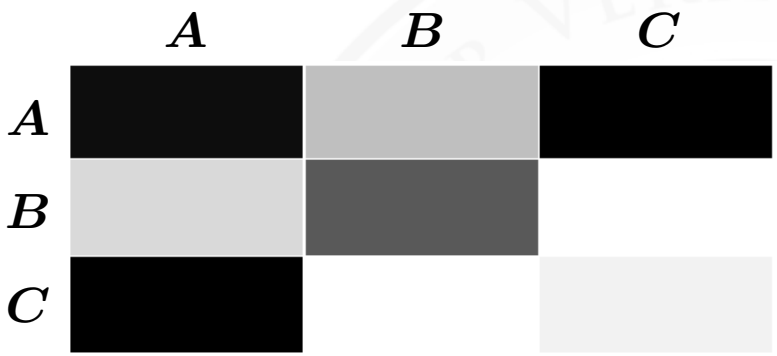
# Causal Network Inference

Graph structure and interaction weights unknown.

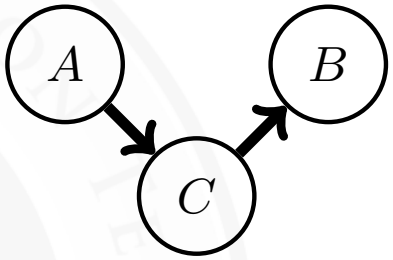


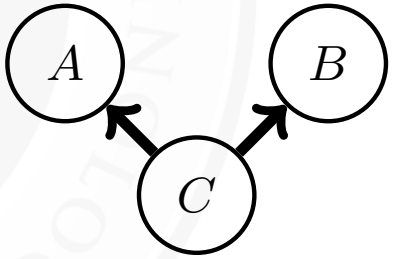
# Causal Network Inference

Covariance Matrix

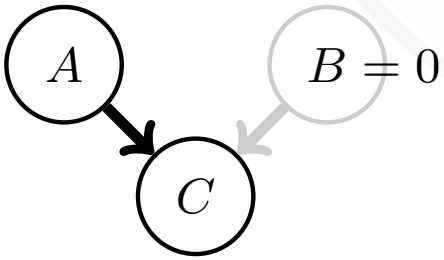



Possible Graphs

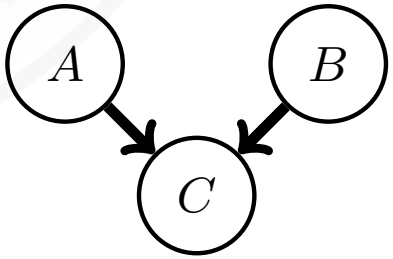


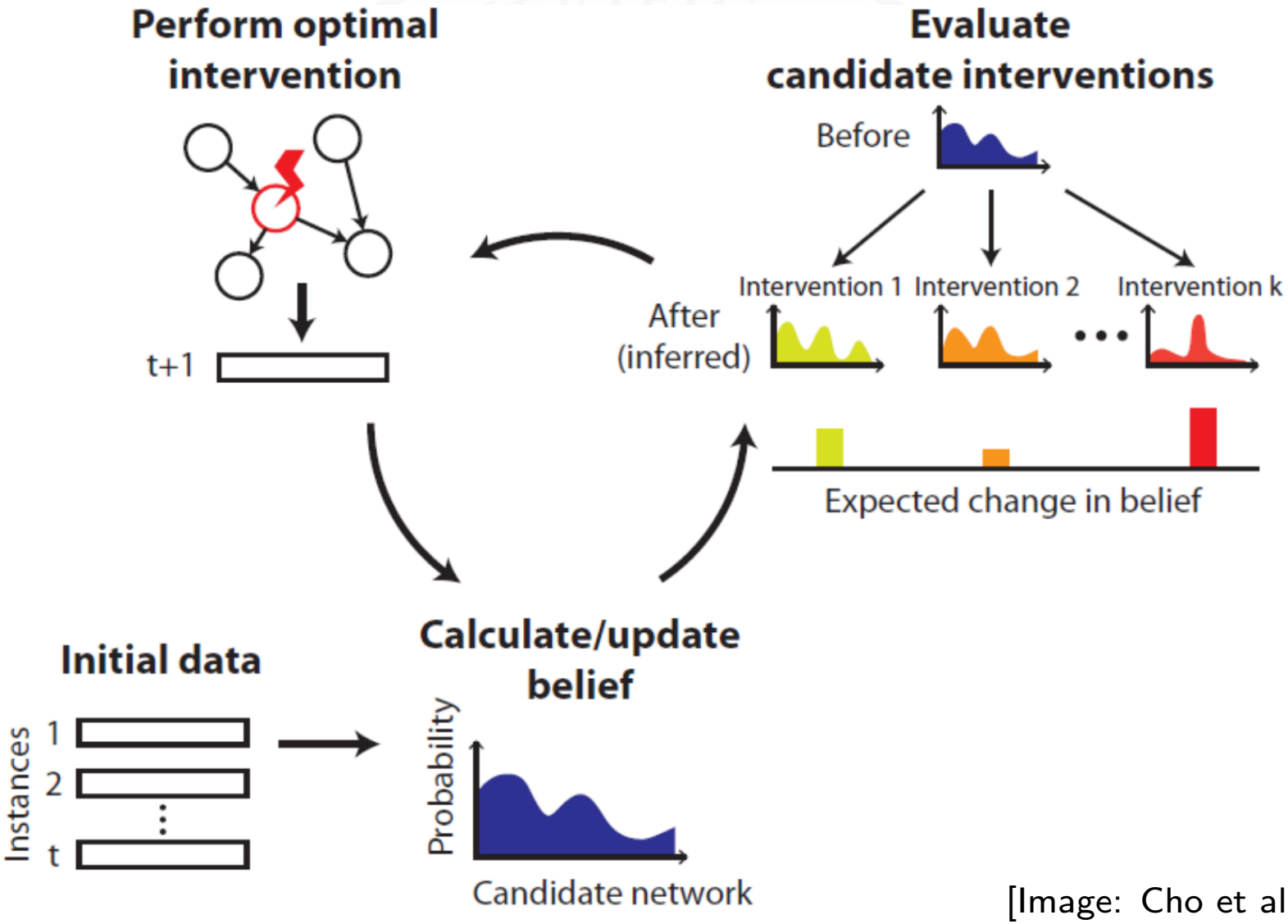

Cannot determine causality from correlations, need to perform <u>active interventions</u>




*Clamp* node to fixed value.



# Causal Network Inference

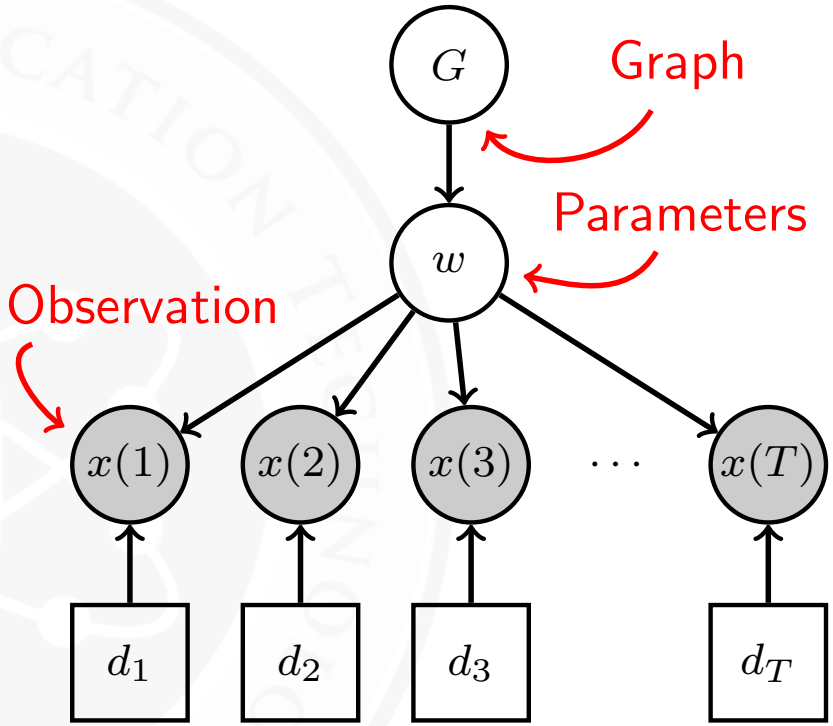



[Image: Cho et al., 2016]

# Causal Network Inference

## Model:

$$x_j | x_{\mathsf{Pa}(j)}, w_j, G \sim \mathcal{N}(w_j^\intercal x_{\mathsf{Pa}(j)}, \sigma_j^2)$$
$$w_j | G \sim \mathcal{N}(\cdot)$$
$$G \sim \text{Uniform-DAG}$$




## Planning:

Previous $t-1$ observations:

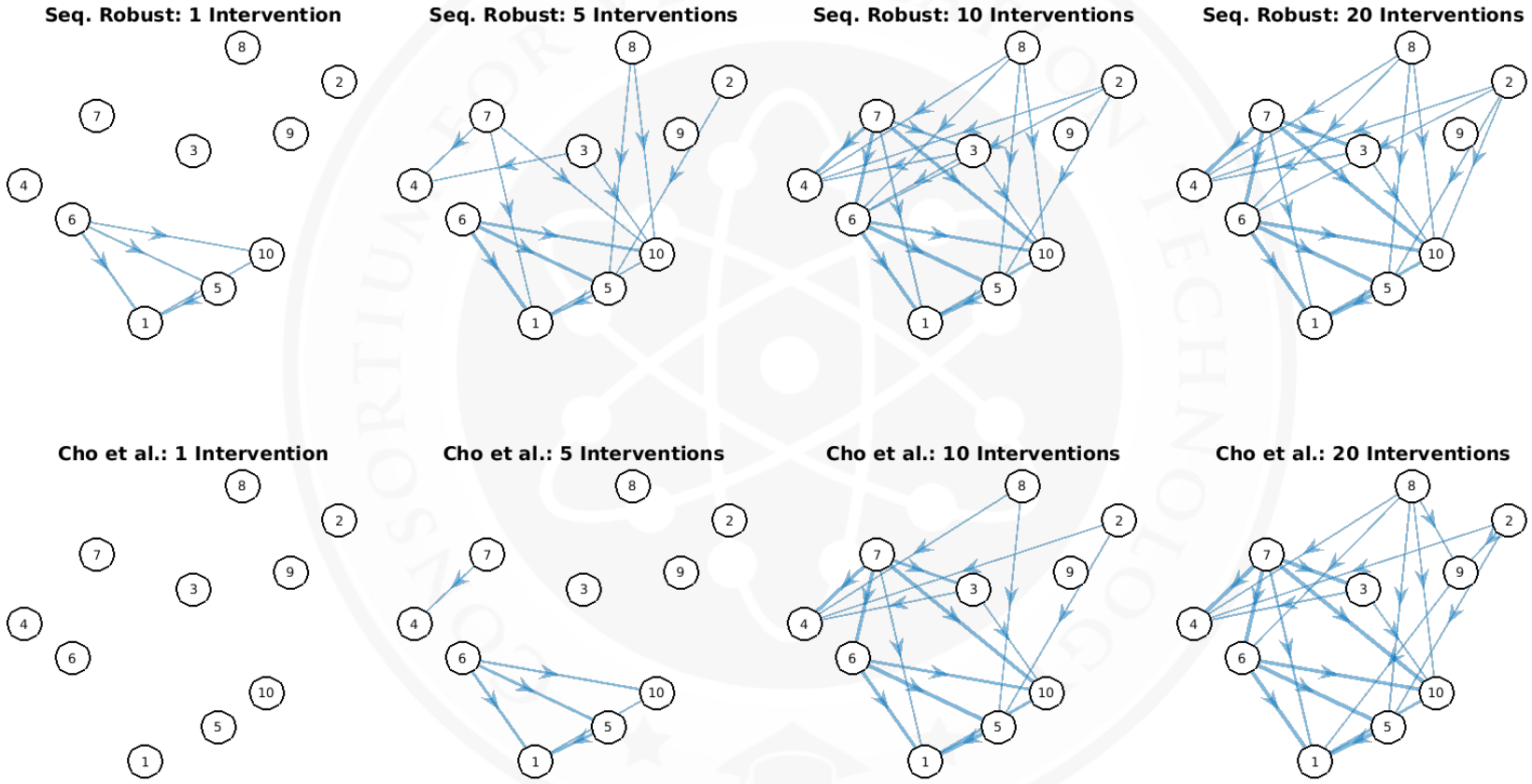
$$\mathcal{X} = \{x(1), \ldots, x(t-1)\}$$

Select intervention to maximize mutual information:

$$d^* = \arg\max_d I(G; X | \mathcal{X}, d)$$

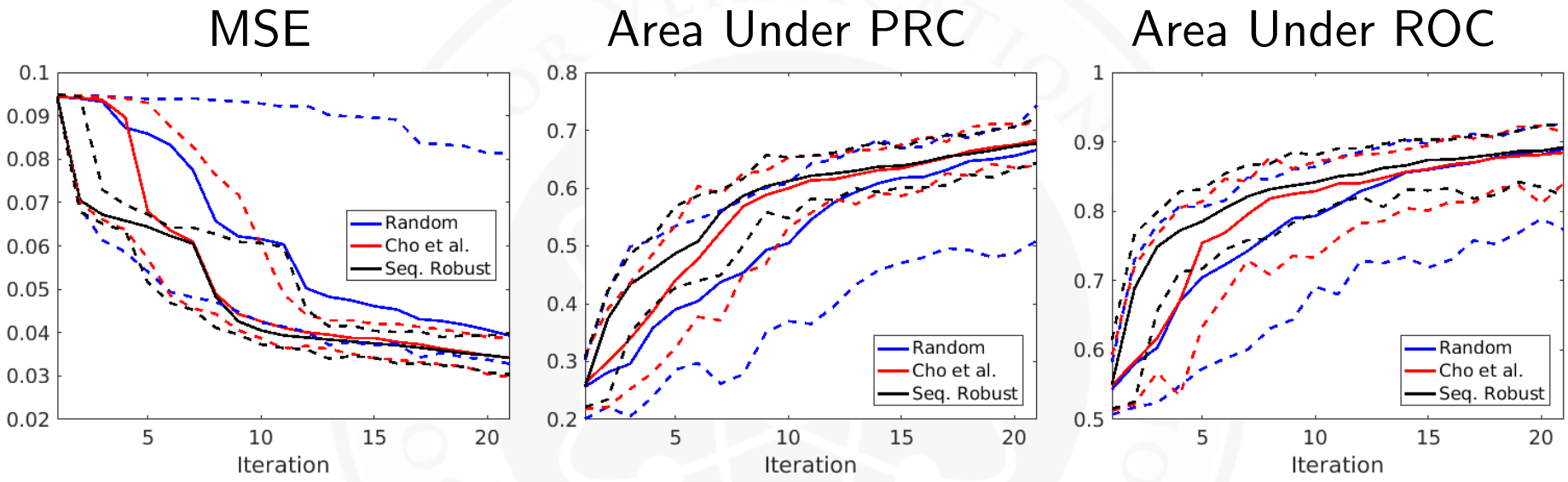
Clamp node $x_{d^*} = 0$ and observe remaining nodes.

# Causal Network Inference




Robust planning selects most informative interventions in early iterations.

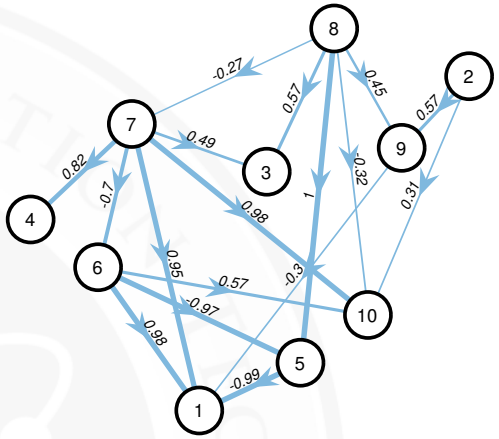
# Causal Network Inference

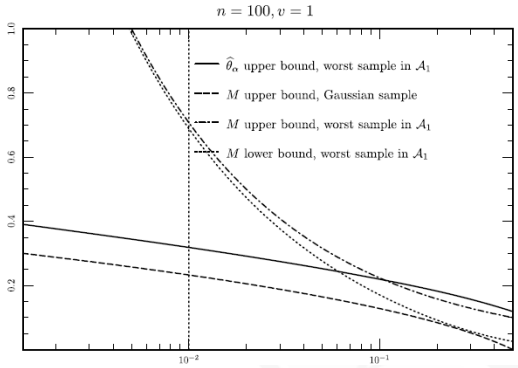



- Median (solid) best/worst (dashed) out of 50 runs
- More informative experiments than Random
- In early iterations chooses more informative experiments compared to Cho et al., 2016

# Summary

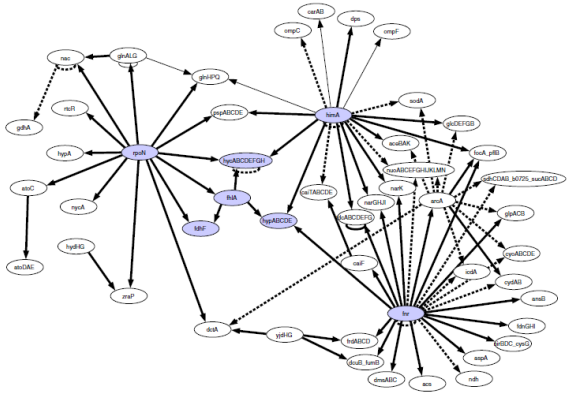
Sequential Bayesian inference and planning applies to complex models (more interesting applications...)



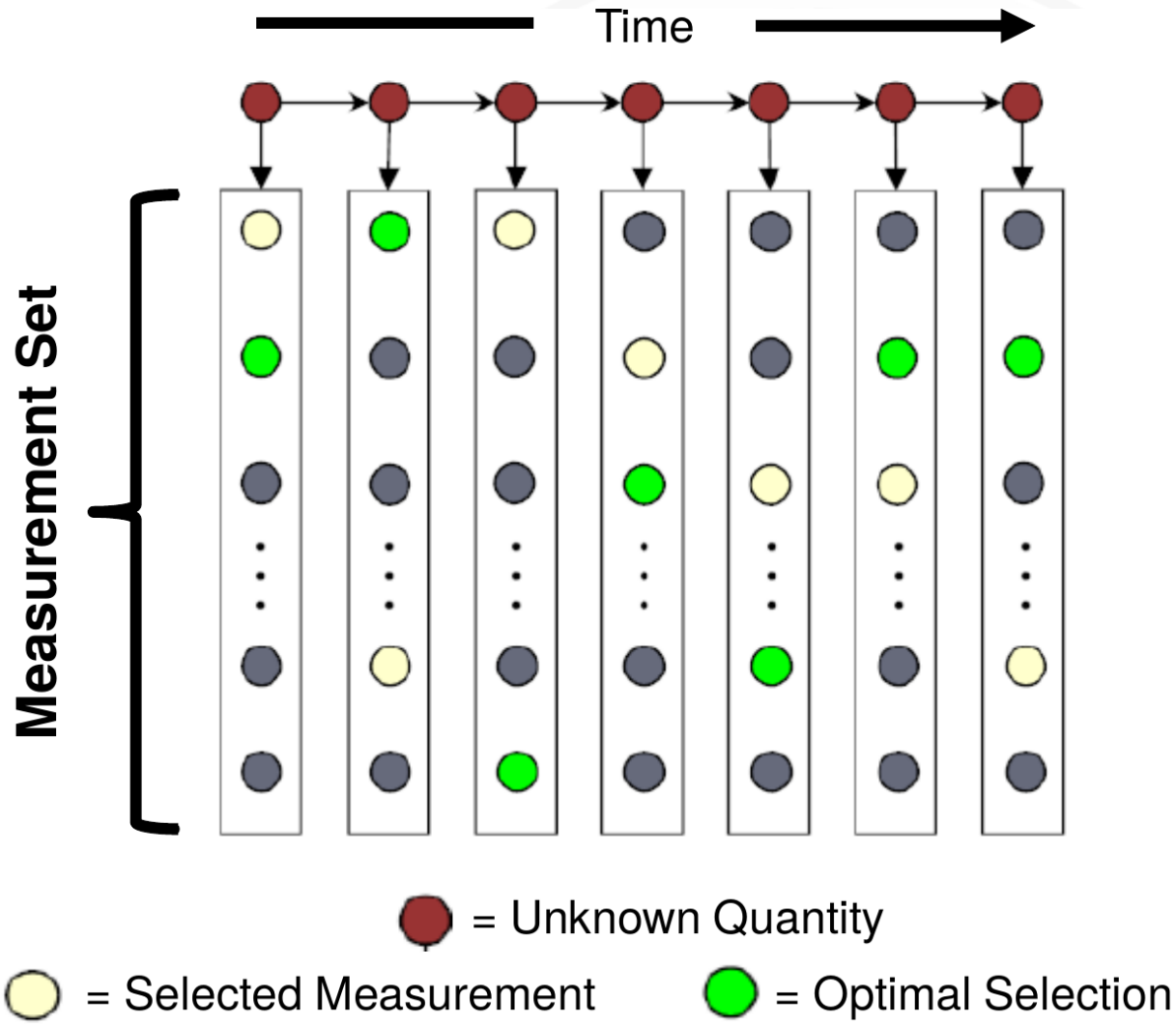




Theoretical guarantees on estimator quality for costly decisions.

Scale up to larger problems and bound probability of incorrect selection.



# Measurement Selection




- Sensor Types
  - Satellite
  - Flyby
  - Earth-based
- Sensing Modes
  - EO/SAR/IR
  - Hyperspectral
  - Radio Freq

# Plan Execution Costs

$$d^* = \arg\max_d I(X;Y|d) - \lambda R(d)$$

Information Reward

Information/Cost Tradeoff

Cost

- Plan is feasible if information justifies cost:

$$I(X;Y|d) > \lambda R(d)$$

- Possible costs for this application:
  - Sensor costs
  - Power consumption

# Monte Carlo Integration

Need to compute expected values of the form:

$$\mathbb{E}\left[f(x)\right] = \int p(x) f(x) dx$$

Draw samples from the distribution:

$$\{x^i\}_{i=1}^N \sim p(x)$$

Monte Carlo integration is **empirical mean**

$$\mathbb{E}\left[f(x)\right] \approx \frac{1}{N} \sum_{i=1}^{N} f(x^i)$$

Good statistical properties as $N \to \infty$, but problematic for finite samples

# Asymptotic and Finite Sample Bounds

## Asymptotic Bounds:

$$\sqrt{N}\hat{H}_Y \xrightarrow{d} \mathcal{N}\left(H(Y), \mathbb{E}_y\left[\frac{\sigma_x^2\left(p(Y|X)\right)}{Mp^2(Y)} + \sigma^2\left(\log p(Y)\right)\right]\right)$$

$$\sqrt{N}\hat{H}_{Y|X} \xrightarrow{d} \mathcal{N}\left(H(Y|X), \sigma^2\left(\log p(Y|X)\right)\right)$$

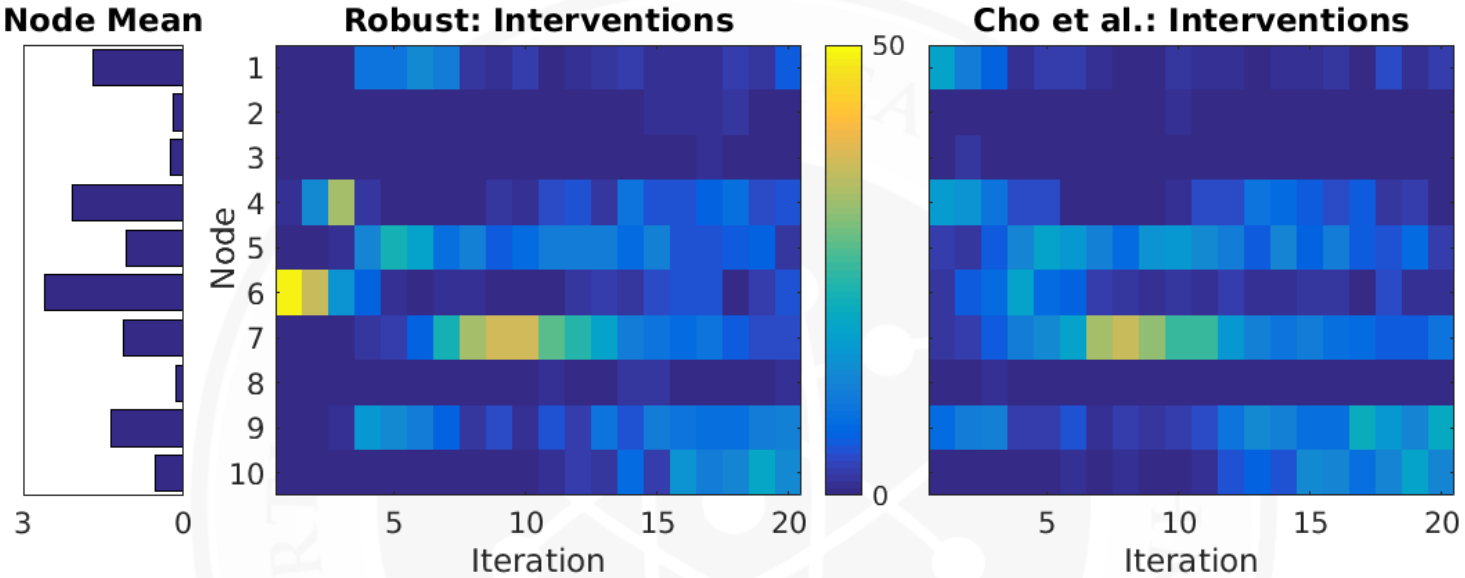
## Finite-Sample Bounds:

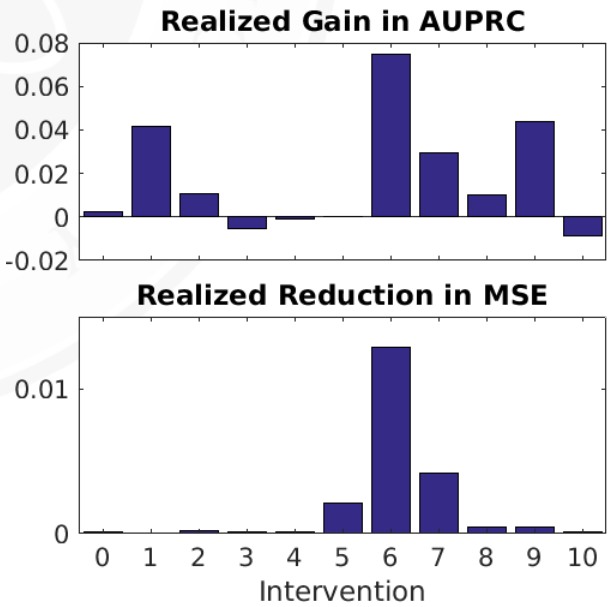
Assuming $N > 2(1 + \log \epsilon^{-1})$

$$H(Y) + b - c < \hat{H}_Y < H(Y) + b + c \qquad \text{w.p.} \quad \geq 1 - 2\epsilon$$

$$b = \mathbb{E}\left[\log \frac{p(Y)}{\hat{p}(Y;X)}\right] \qquad c = \frac{1 + \log \epsilon^{-1}}{1 - (1 + \log \epsilon^{-1})/N}\sqrt{\frac{\sigma^2}{2N}}$$

# Causal Network Inference




- Robust is more consistent in intervention selection.
- Average realized gain at $t = 1$ indicates intervention at node 6 is optimal.