Sequential Decision-Making: Theory and Applications in Public Health

Mengyan Zhang

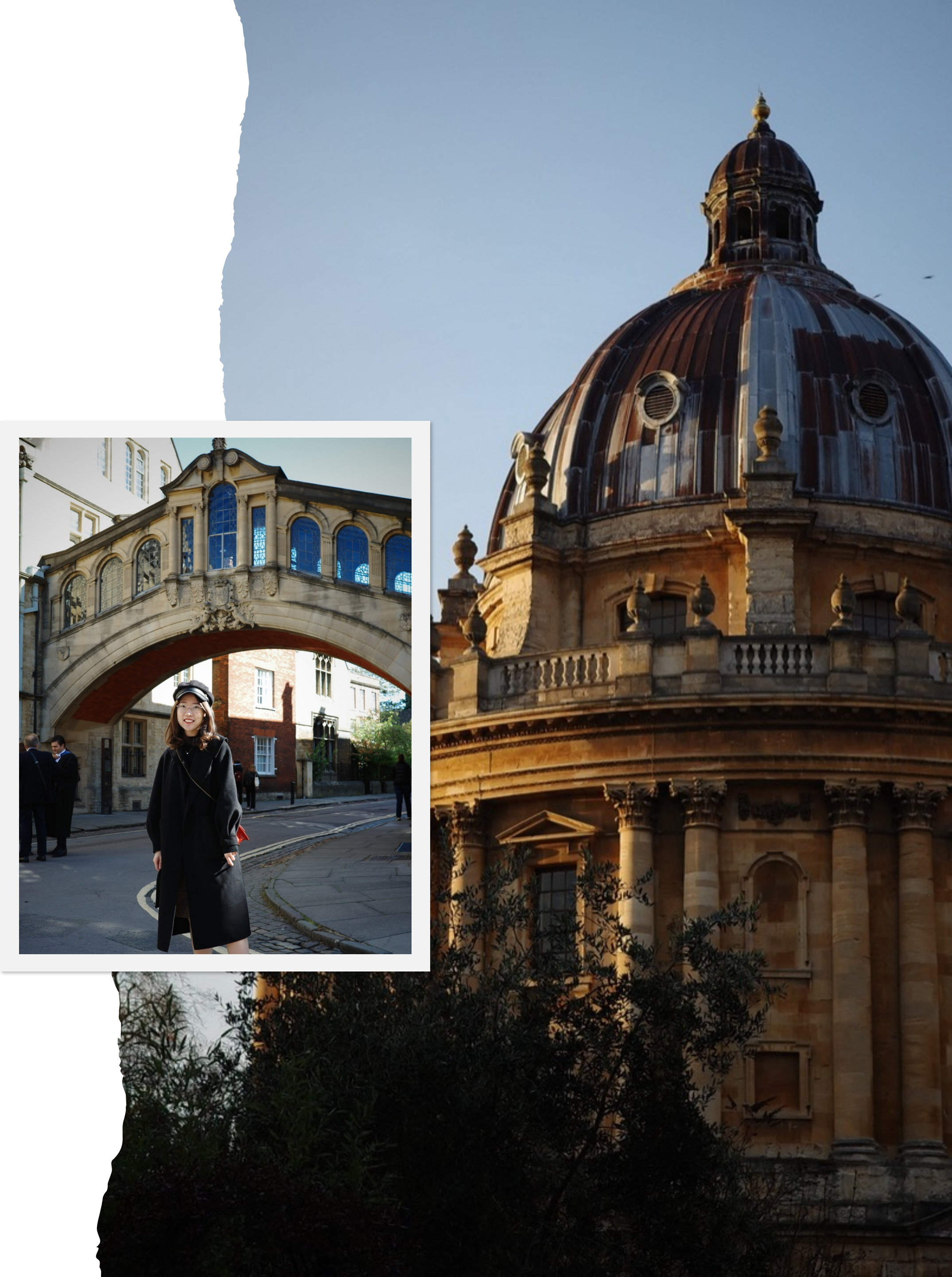
Department of Computer Science, University of Oxford

Machine Learning & Global Health Network

@LAS Group – ETH Zurich 15/2/2024

About me

- NCCR Automation Visit Fellowship
- Postdoc at the University of Oxford, supervised by Seth Flaxman
- Member of Machine Learning and Global Health
- PhD in Australian National University; Data61, CSIRO (Thesis: Adaptive Recommendations with Bandit Feedback)
- Research interest:
  - Sequential decision-making – imperfect feedback; causal decision-making
  - Applications in public health
- Hobbies: photography, nature, travel



Sequential Decision-Making & Public Health

Sequential Decision-Making

Bandit,
Bayesian Optimization,
Active Learning,
Reinforcement Learning

↔

Public Health

Vaccine Allocation,
Disease Surveillance,
Sensor Placement,
Personalized Treatment

Outline

Imperfect Feedback: Aggregated Feedback

Applications in Public Health

- Causal Decision Making
- Active Disease Surveillance
- Vaccine Allocation

Conclusion & Future Work

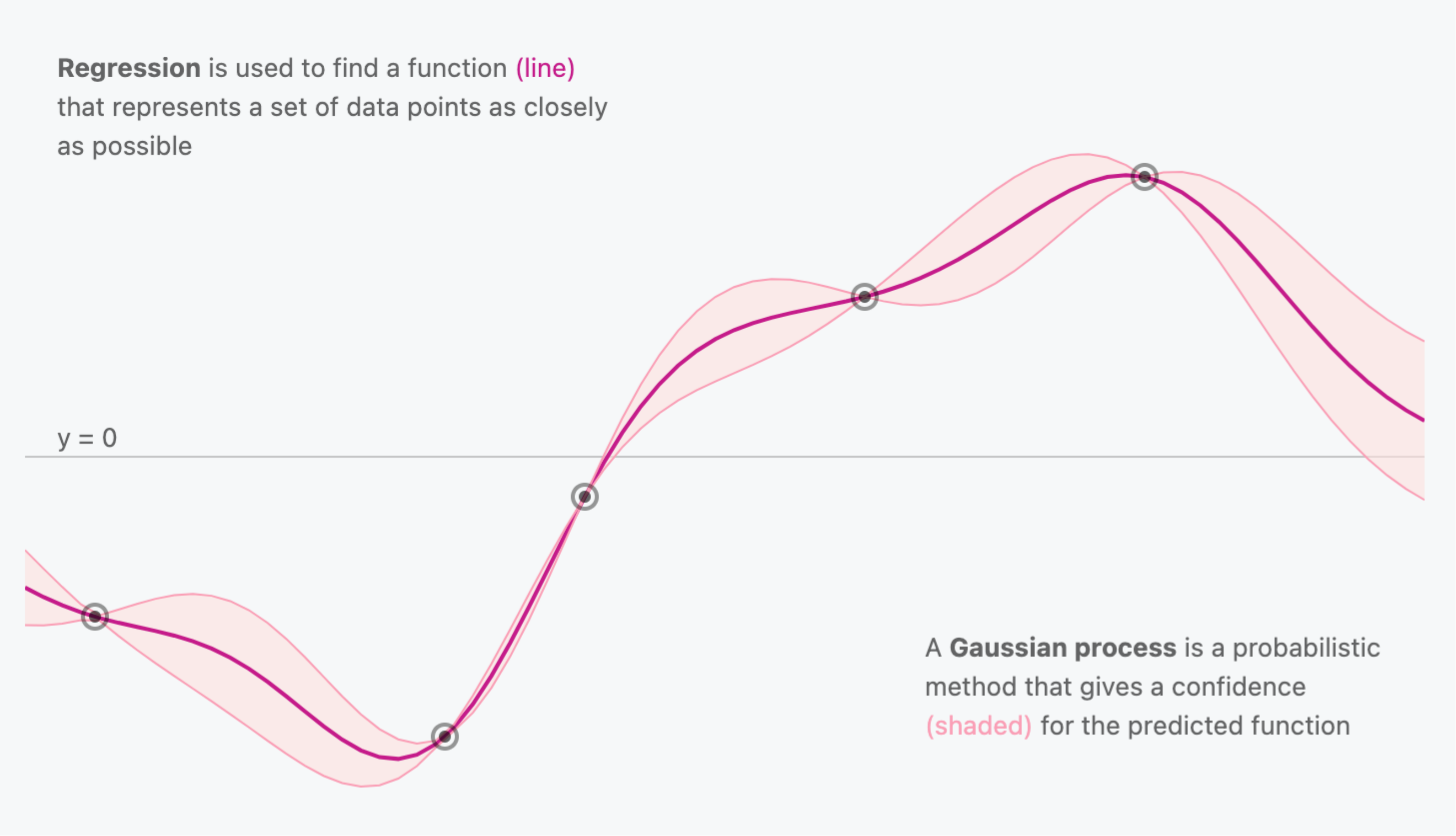
Why Aggregated Feedback?

Application	Arm	Reward	Goal: to design a policy such that…	Why Aggregated?
DNA Design	DNA sequences	average protein expression level in a mixed culture	identify DNA sequences with the highest protein expression level with a given budget	expensive; search space is large
Census Querying	Respondent	average age of respondents inside queried area	identify the region with the highest average age with a fixed amount of querying	privacy concerns
Radio Telescope	spatial-frequency coordinates of objects in the sky	average radio wave energy from the queried area	identify the region with the highest average radio energy with a fixed amount of querying	hardware constraint

Reward Smoothness – e.g. Gaussian Process

$$f(\mathbf{x}) \sim \mathcal{GP}\left(\mu(\mathbf{x}), k\left(\mathbf{x}, \mathbf{x}'\right)\right)$$

$$\mu(\boldsymbol{x}) = \mathbb{E}[f(\boldsymbol{x})] \quad \text{and} \quad k(\boldsymbol{x}, \boldsymbol{x}') = \mathbb{E}[(f(\boldsymbol{x}) - \mu(\boldsymbol{x}))(f(\boldsymbol{x}') - \mu(\boldsymbol{x}'))].$$



https://distill.pub/2019/visual-exploration-gaussian-processes/

Acquisition Function: e.g. GP-UCB

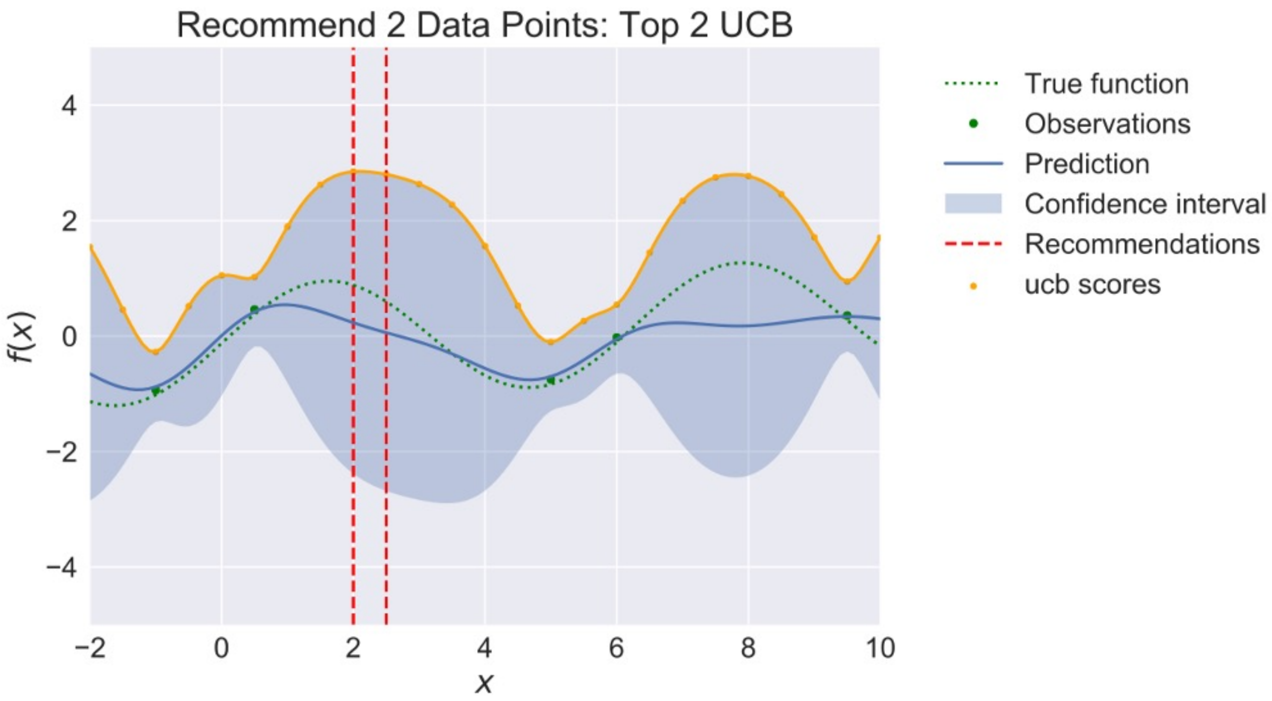
Gaussian Process Upper Confidence Bound (GP-UCB) [1]

Posterior mean: exploitation

Posterior standard deviation: exploration

$$\text{argmax}_{\mathbf{x}_i \in \mathcal{K}} \left(\mu_t(\mathbf{x}_i) + \beta_t \sigma_{t-1}(\mathbf{x}_i) \right)$$

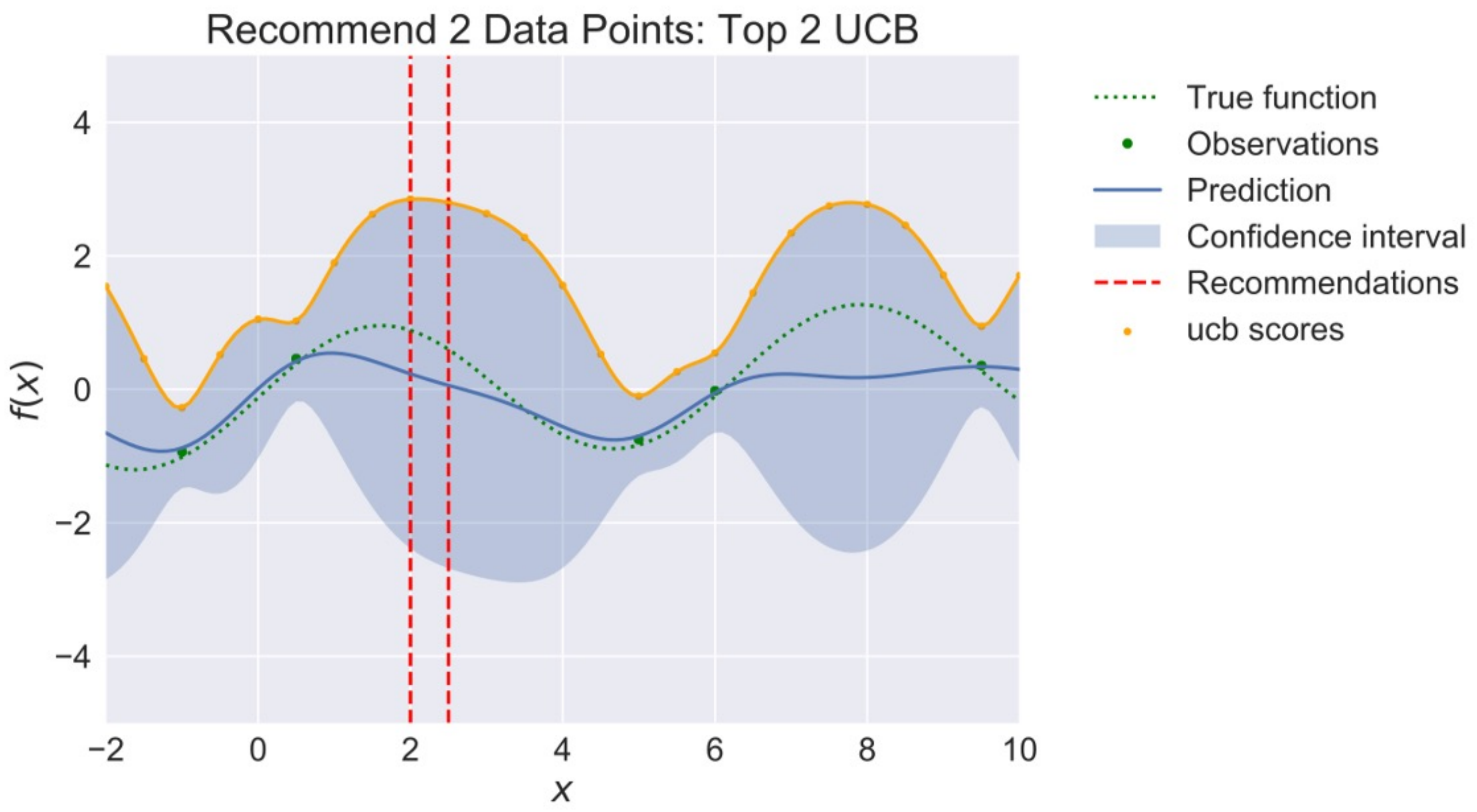
Balancing term



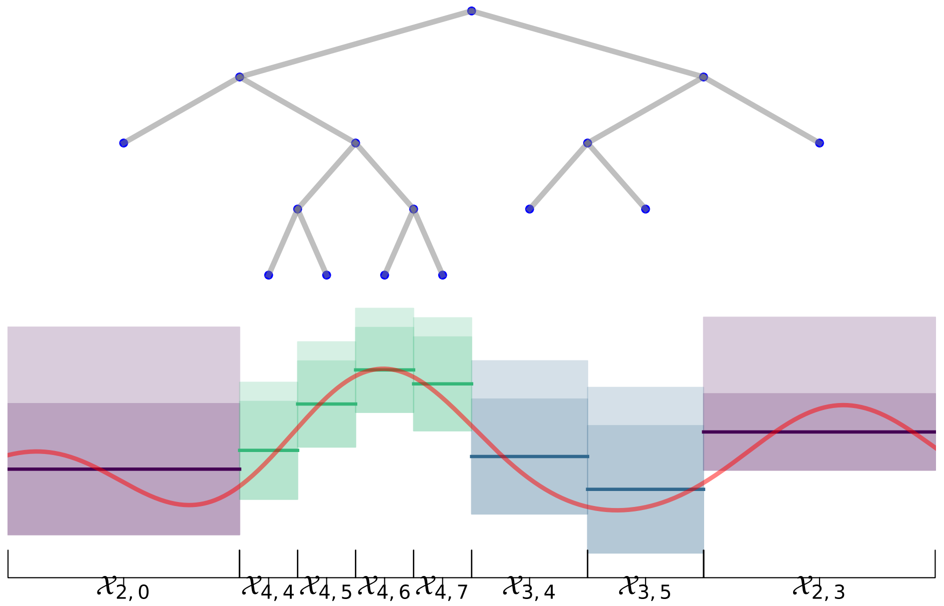
[1] Srinivas, N.; Krause, A.; Kakade, S. M.; and Seeger, M., Gaussian process optimization in the bandit setting: No regret and experimental design. ICML 2009.

Computational cost for continuous space

$$\text{argmax}_{\mathbf{x}_i \in \mathcal{K}} \left(\mu_t(\mathbf{x}_i) + \beta_t \sigma_{t-1}(\mathbf{x}_i)\right)$$



Problem Setting



- Arms: a leaf node, corresponding to a subset of continuous $[0,1]^d$
- Rewards: sampled from GP, only average reward for a node

$$r_t = \bar{F}(X_{h_t,i_t}) + \epsilon_t, \quad \bar{F}(X_{h_t,i_t}) := \frac{\sum_{\boldsymbol{x} \in \mathcal{C}_{h_t,i_t}} f(\boldsymbol{x})}{|\mathcal{C}_{h_t,i_t}|}$$

with **Representative points** $\mathcal{C}_{h,i} = \{\boldsymbol{x}_{h,i^s}\}_{1 \le s \le S}$, where $\boldsymbol{x}_{h,i^s} \in \mathcal{X}_{h,i}$.

- Aggregated Regret $R_N = f(\boldsymbol{x}^*) - \bar{F}(X_N)$.

[1] Gaussian Process Bandits with Aggregated Feedback. **Mengyan Zhang**, Russell Tsuchida, Cheng Soon Ong. AAAI 2022.

Gaussian Process Optimistic Optimisation (GPOO)

How to choose node and when to split?

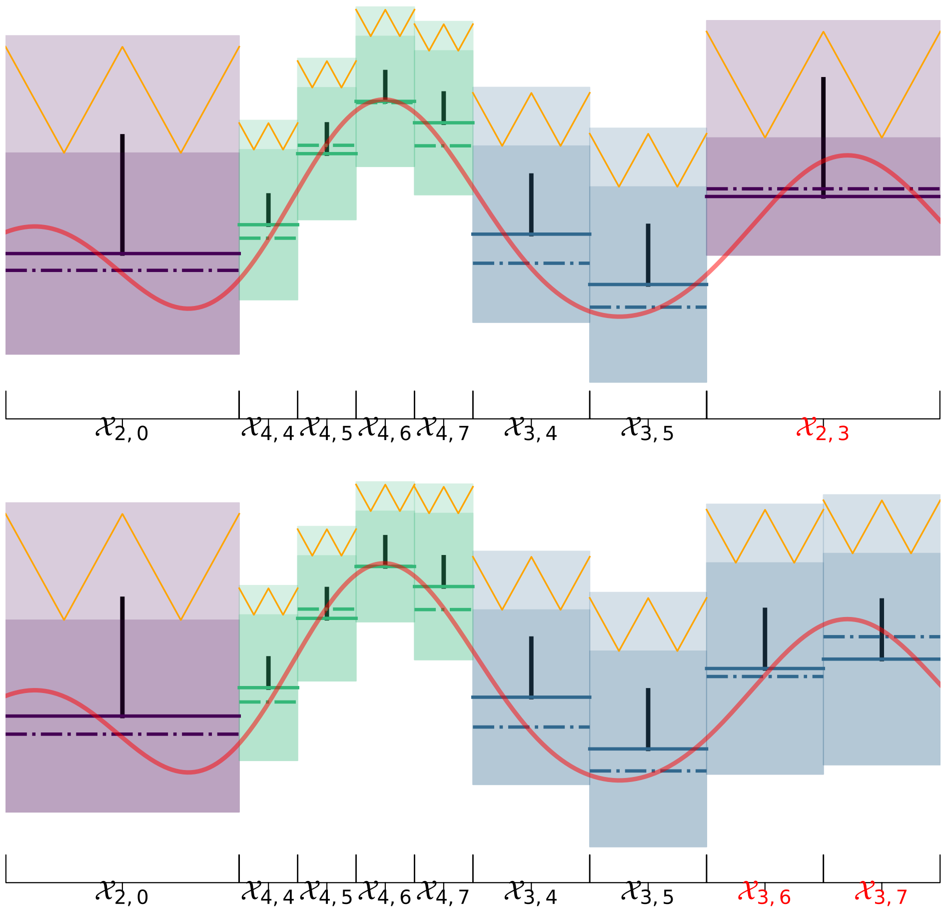
Assumption: **Decreasing Diameters:** $\sup_{\boldsymbol{x} \in \mathcal{X}_{h,i}} L\ell(\boldsymbol{x}_{h,i}, \boldsymbol{x}) \leq \delta(h)$ some decreasing sequence $\delta(h) > 0$.

- **Select leaf node with largest b-value:**

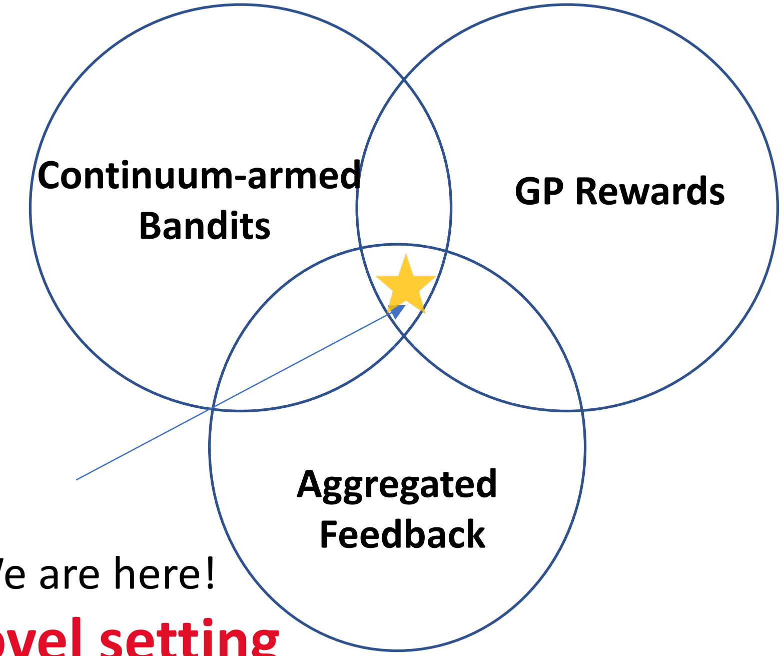
$b_{h,i}(t)$ = posterior mean + confidence interval + diameter $\delta(h_t)$

- **Expand:** if $\delta(h_t)$ > confidence interval

- **Return:** the node with the maximum posterior mean in the deepest non-leaf layer



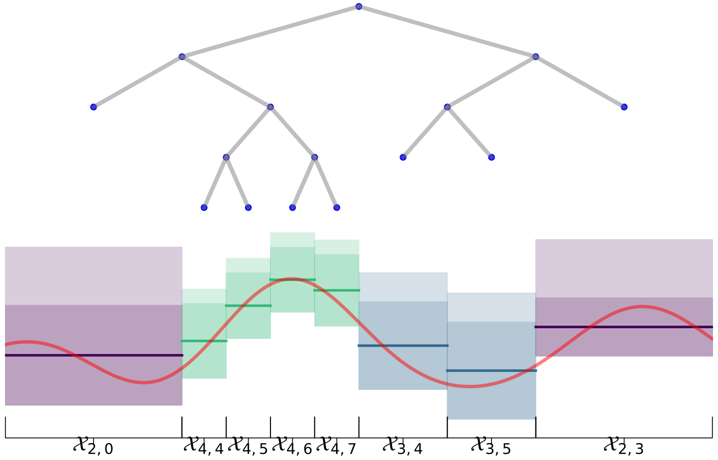
Contributions of GPOO



We are here!

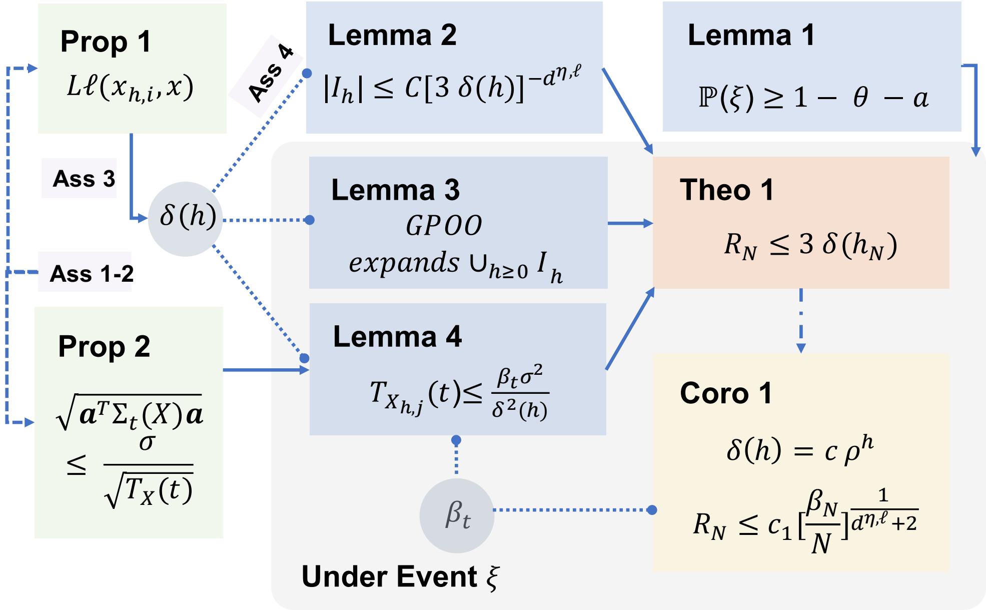
A novel setting

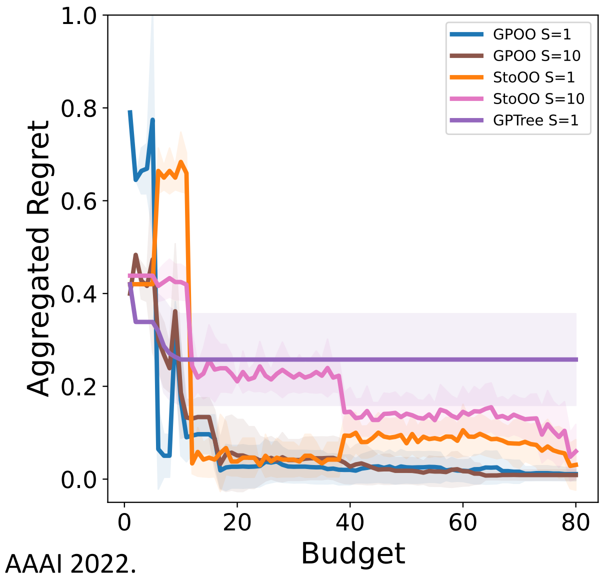
New Algorithm: **GPOO**



Theoretical results:

Upper bound on (aggregated) regret





Simulation results:

Outperform baselines

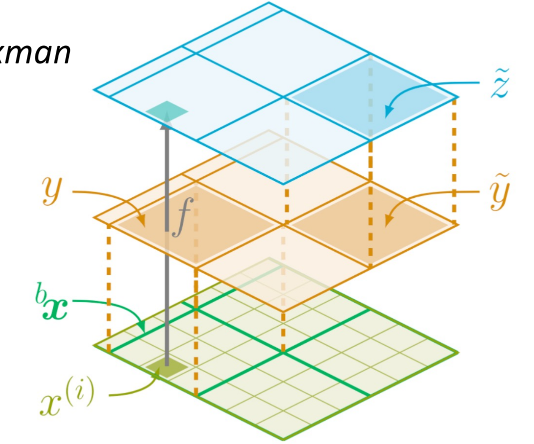
[1] Gaussian Process Bandits with Aggregated Feedback. **Mengyan Zhang**, Russell Tsuchida, Cheng Soon Ong. AAAI 2022.

Beyond GPOO?

- Can we select “cells” beyond the tree structure?
- Can we extend the average over representatives to a more complicated aggregation? -- conditional expectation
- Can we learn the optimal in terms of original functions, instead of aggregated values?

- Motivation: we’d like to query user groups by e.g. gender, age, region, etc (privacy, expensive data query)
  - Malaria incidence optimization (target: malaria incidence rate)
  - Survey design (target: survey of interest, non-response)
  - Treatment effect optimization (target: individual treatment effect)

Collaborated with Shahine Bouabid, Dino Sejdinovic, Seth Flaxman



Deconditional Bayesian Optimisation

- Optimise unknown function $f : \mathcal{X} \to \mathbb{R}$, where $\mathcal{X} \subset \mathbb{R}^d$
- Observe dataset $\mathcal{D}_1 = \{(x_j, y_j)\}_{j=1}^N$, where $y_j \sim \mathbb{P}_Y$ and $x_j \sim \mathbb{P}_{X|Y=y_j}$
- Goal: identify global optimal point $x_* = \arg\max_{x \in \mathcal{X}} f(x)$
- Sequentially query: $\tilde{y} \in \mathcal{Y}$ and observe noisy rewards $\tilde{z} = g(\tilde{y}) + \epsilon$

$$g(\tilde{y}) = \mathbb{E}_X[f(X)|Y = \tilde{y}] = \int_{\mathcal{X}} f(x)p(x|\tilde{y})dx$$

- Evaluation: expected simple regret, with budget M

$$R_M = \mathbb{E}[|f(\mathbf{x}_M) - f(\mathbf{x}_*)|]$$

Background: Deconditional Posterior

Conditional Mean Process (CMP): $f \sim \mathcal{GP}(m, k)$

Then $g \sim \mathcal{GP}(v, q)$

$$\nu(y) = \mathbb{E}_X[m(X) \mid Y = y] \quad q(y, y') = \mathbb{E}_{X,X'}\left[k(X, X') \mid Y = y, Y' = y'\right]$$

Deconditional Posterior:

$$\begin{bmatrix} f(\mathbf{x}) \\ \tilde{\mathbf{z}} \end{bmatrix} \mid \mathbf{y}, \tilde{\mathbf{y}} \sim \mathcal{N}\left(\begin{bmatrix} m(\mathbf{x}) \\ \nu(\tilde{\mathbf{y}}) \end{bmatrix}, \begin{bmatrix} \mathbf{K}_{\mathbf{x}\mathbf{x}} & \boldsymbol{\Upsilon} \\ \boldsymbol{\Upsilon}^\top & \mathbf{Q}_{\tilde{\mathbf{y}}\tilde{\mathbf{y}}} + \sigma^2 \mathbf{I}_M \end{bmatrix}\right)$$

Siu Lun Chau, Shahine Bouabid, and Dino Sejdinovic. Deconditional Downscaling with Gaussian Processes. NeurIPS 2021.

Deconditional Predictive Entropy Search

$$\mathcal{D}_1 = \{(x_j, y_j)\}_{j=1}^{N}, \text{ where } y_j \sim \mathbb{P}_Y \text{ and } x_j \sim \mathbb{P}_{X|Y=y_j}$$

$$\mathcal{D}_2^M = \{(\tilde{y}_j, \tilde{z}_j)\}_{j=1}^{M}$$

Pick in Y space, but optimise f(x)?

$$\begin{aligned}\tilde{y}_{M+1} &= \arg\max_{\tilde{y}\in\mathcal{Y}} \mathbb{I}(\tilde{z}, x_* | \tilde{y}, \mathcal{D}_1, \mathcal{D}_2^M) \\ &= \arg\max_{\tilde{y}\in\mathcal{Y}} H[\tilde{z}|\tilde{y}, \mathcal{D}_1, \mathcal{D}_2^M] - \mathbb{E}_{x_*|\mathcal{D}_1, \mathcal{D}_2^M}[H[\tilde{z}|x_*, \tilde{y}, \mathcal{D}_1, \mathcal{D}_2^M]]\end{aligned}$$

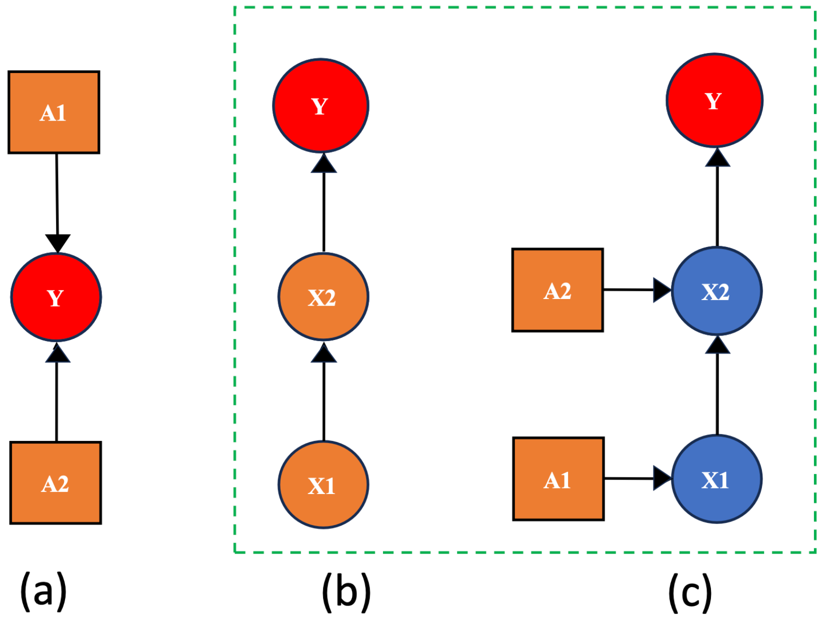
Potentially Future Work

- Sequentially query D1
- Partial monitoring
- Cumulative regret setting: information directed sampling
- Distributionally Robust BO (distributional context)

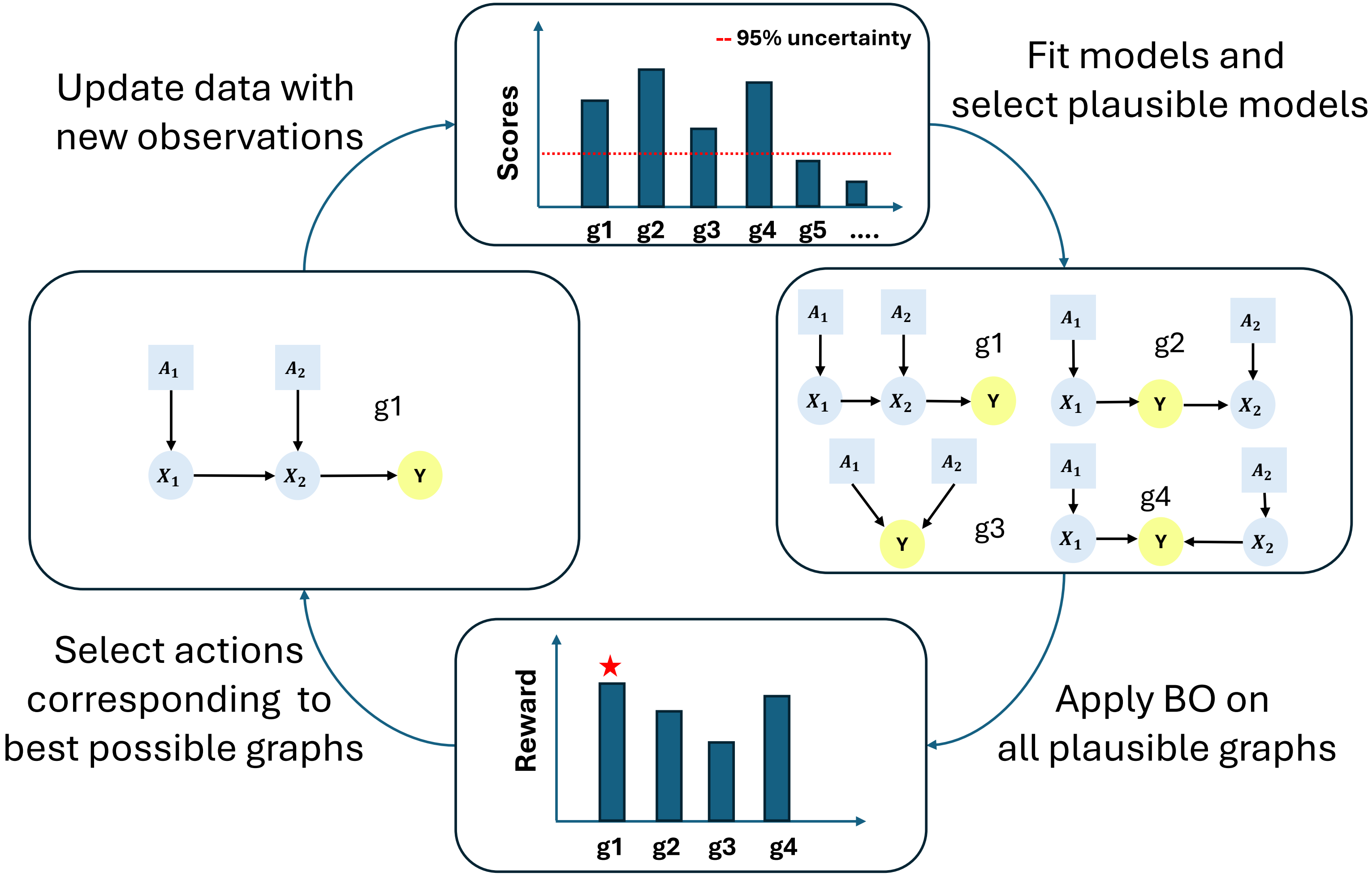
Applications in Public Health

Causal Bayesian Optimisation with Unknown Graph

- Structural Bandits – taking causal structure into account
- The causal graph can be unknown or partially known
- Balance: exploitation – exploration (surrogate model, causal structure)



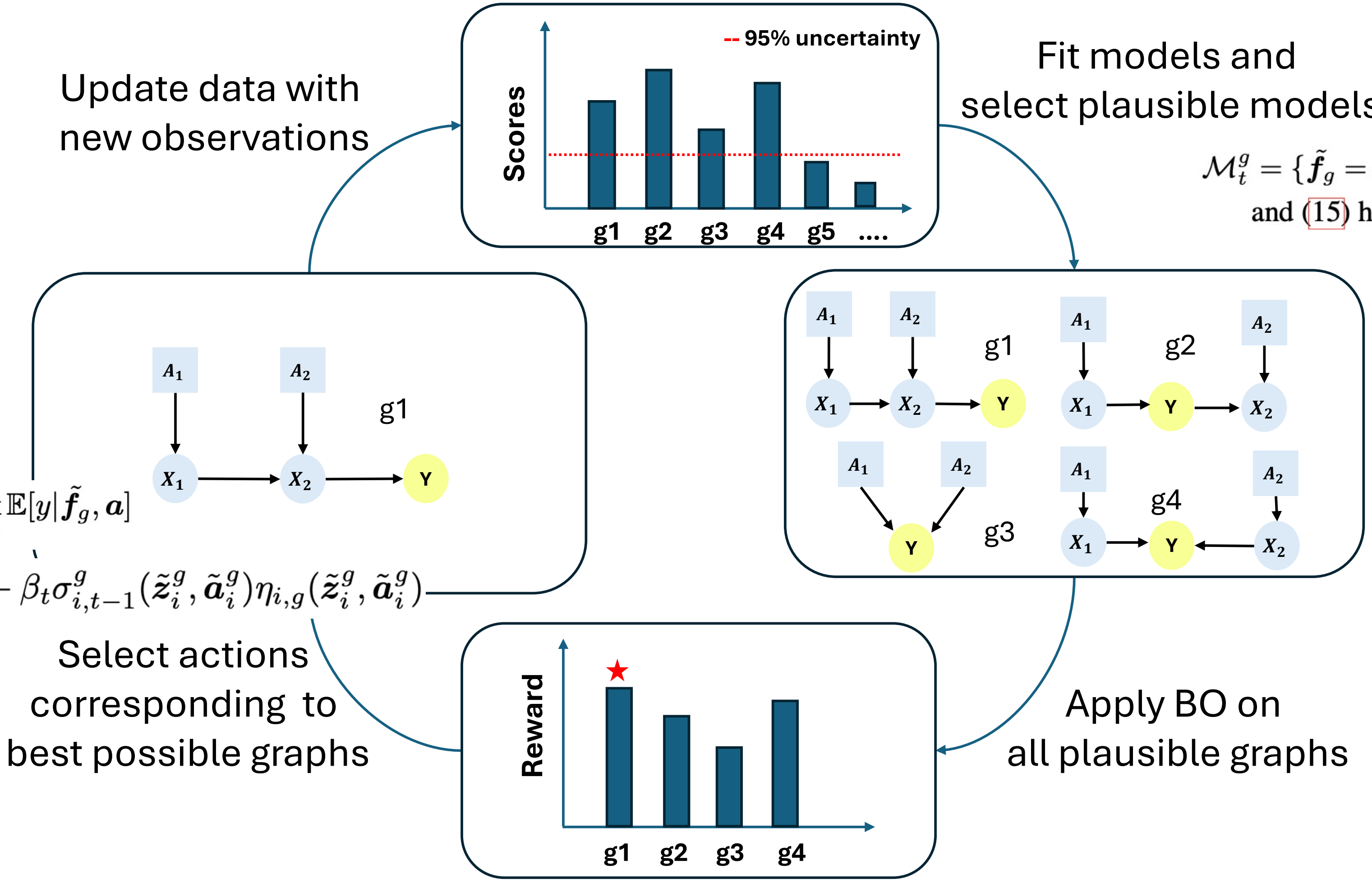
Causal Bayesian Optimisation with Unknown Graph



Collaborated with Sumantrak Mukherjee, Seth Flaxman, Sebastian Vollmer

Causal Bayesian Optimisation with Unknown Graph

$$S(V_i, \boldsymbol{Z}_i^g, \boldsymbol{A}_i^g|\mathcal{D}_t) = \int p(v_{i,1:t}|g,\boldsymbol{\theta}_i)\pi(\boldsymbol{\theta}_i|g)d\boldsymbol{\theta}_i$$



$$\mathcal{M}_t^g = \{\tilde{\boldsymbol{f}}_g = \{\tilde{f}_i^g\}_{i=0}^{|V|} \text{ s.t. } \forall_i : \tilde{f}_i^g \in \mathcal{H}_{k_i}, \|\tilde{f}_i^g\|_{k_i} \le \mathcal{B}_i, \text{ and (15) holds true } \forall a \in \mathcal{A}\}.$$

$$G_t = \{g|\forall_i \exists \tilde{f}_i^g \in \mathcal{M}_t^g, g \in G_{t-1}\}$$

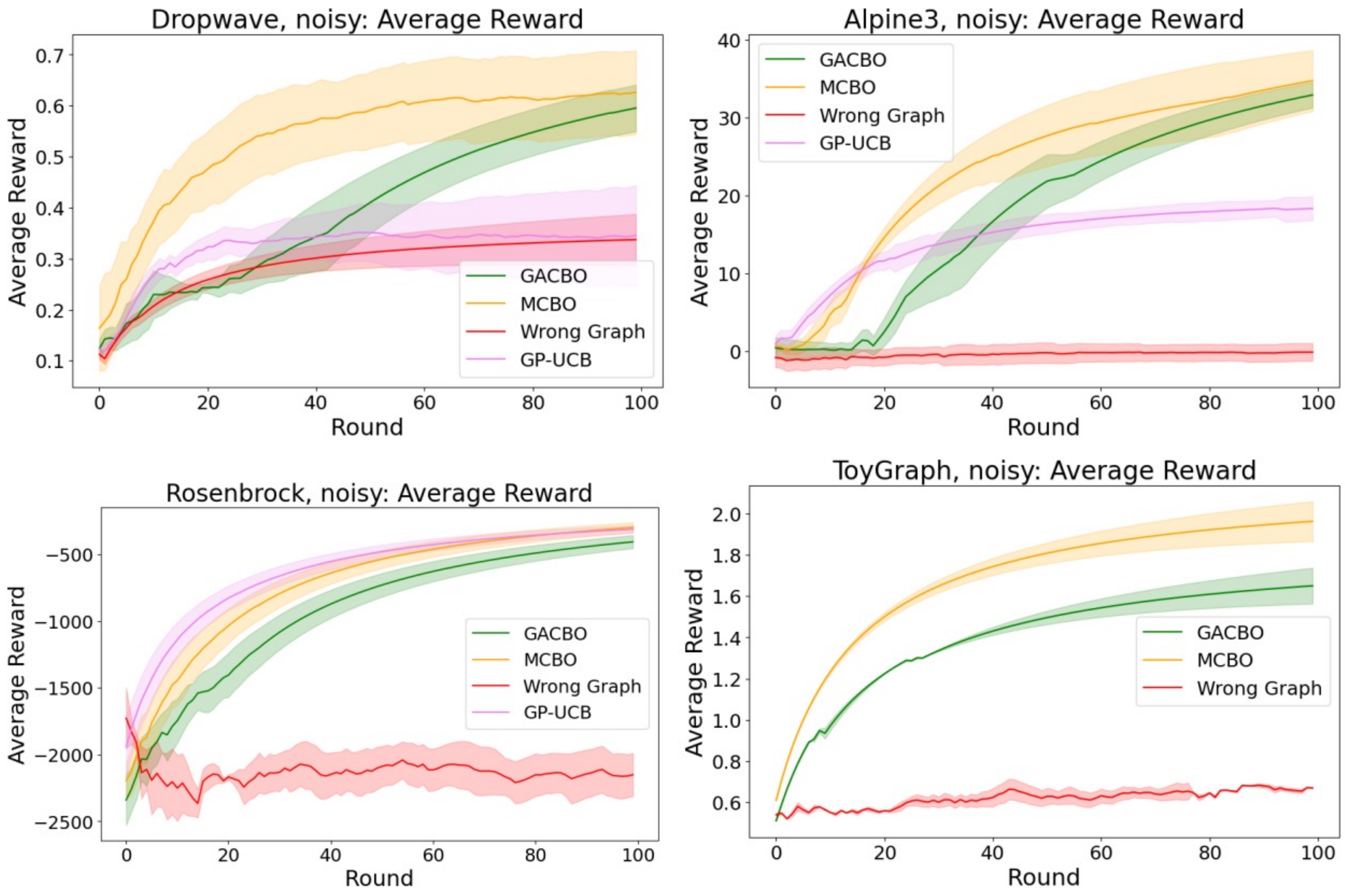
$$|\tilde{f}_{i,t}^g(\boldsymbol{z}_i^g, \boldsymbol{a}_i^g) - \mathbb{E}_{i,t}[\boldsymbol{z}_i, \boldsymbol{a}_i]| \le \beta_{i,t}\sqrt{\mathbb{V}_{i,t}[\boldsymbol{z}_i, \boldsymbol{a}_i]},$$

$$\mathbb{E}_{i,t}[\boldsymbol{z}_i, \boldsymbol{a}_i] = \mathbb{E}_{g\sim p(g|\mathcal{D}_t)}[\mu_{i,t-1}^g(\boldsymbol{z}_i^g, \boldsymbol{a}_i^g)];$$
$$\mathbb{V}_{i,t}[\boldsymbol{z}_i, \boldsymbol{a}_i] = \mathbb{V}_{g\sim p(g|\mathcal{D}_t)}[\mu_{i,t-1}^g(\boldsymbol{z}_i^g, \boldsymbol{a}_i^g)] + \mathbb{E}_{g\sim p(g|\mathcal{D}_t)}[(\sigma_{i,t-1}^g(\boldsymbol{z}_i^g, \boldsymbol{a}_i^g))^2].$$

$$\arg\max_{\boldsymbol{a}\in\mathcal{A}} \max_{g\in G_t} \max_{\boldsymbol{\eta}_g(\cdot)} \mathbb{E}[y|\tilde{\boldsymbol{f}}_g, \boldsymbol{a}]$$

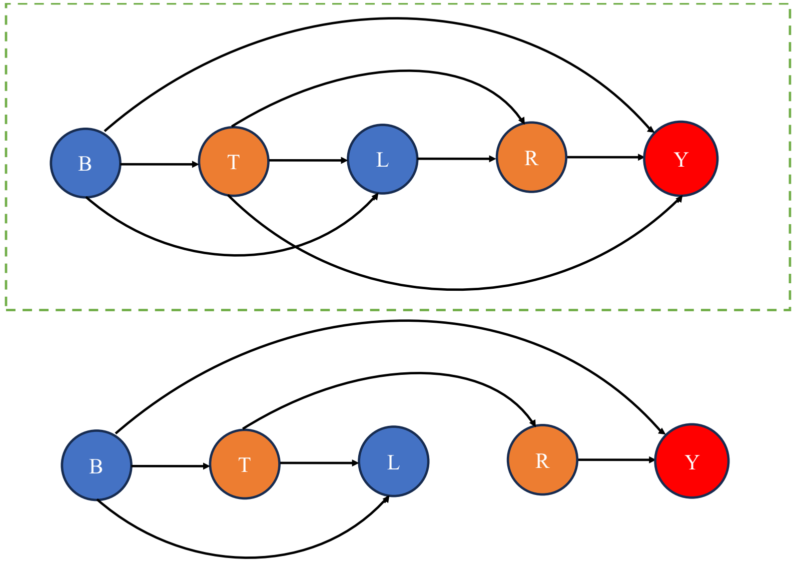
$$\mu_{i,t-1}^g(\tilde{\boldsymbol{z}}_i^g, \tilde{\boldsymbol{a}}_i^g) + \beta_t\sigma_{i,t-1}^g(\tilde{\boldsymbol{z}}_i^g, \tilde{\boldsymbol{a}}_i^g)\eta_{i,g}(\tilde{\boldsymbol{z}}_i^g, \tilde{\boldsymbol{a}}_i^g)$$

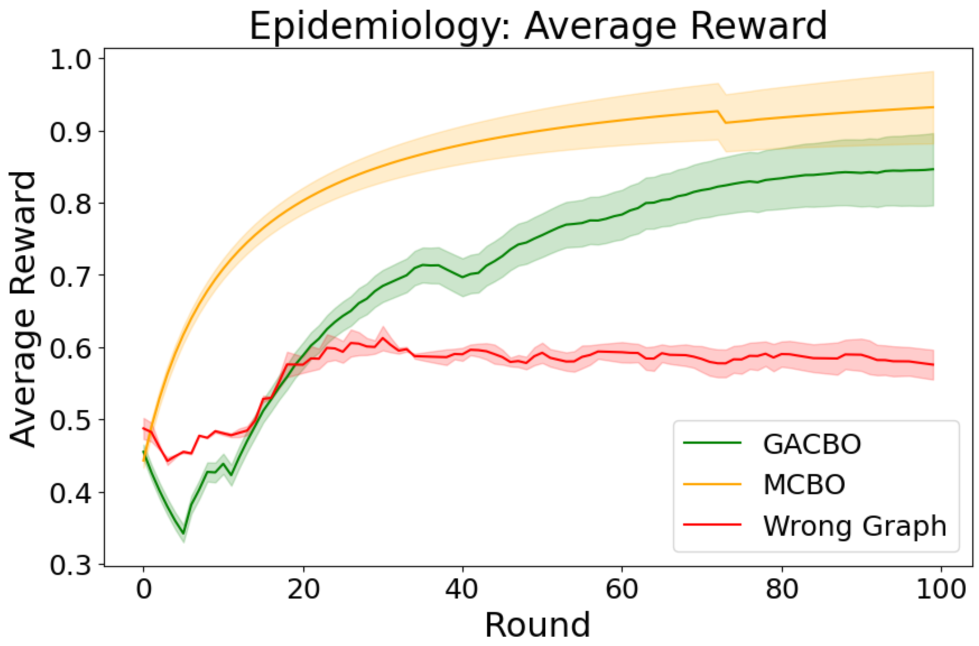
Causal Bandits with Unknown Graph



Applications: Epidemiology

Minimise HIV viral load – two possible treatments T,R





Extension

- Theoretical analysis
- Best arm identification setting
- More scalable approach
- With imperfect feedback

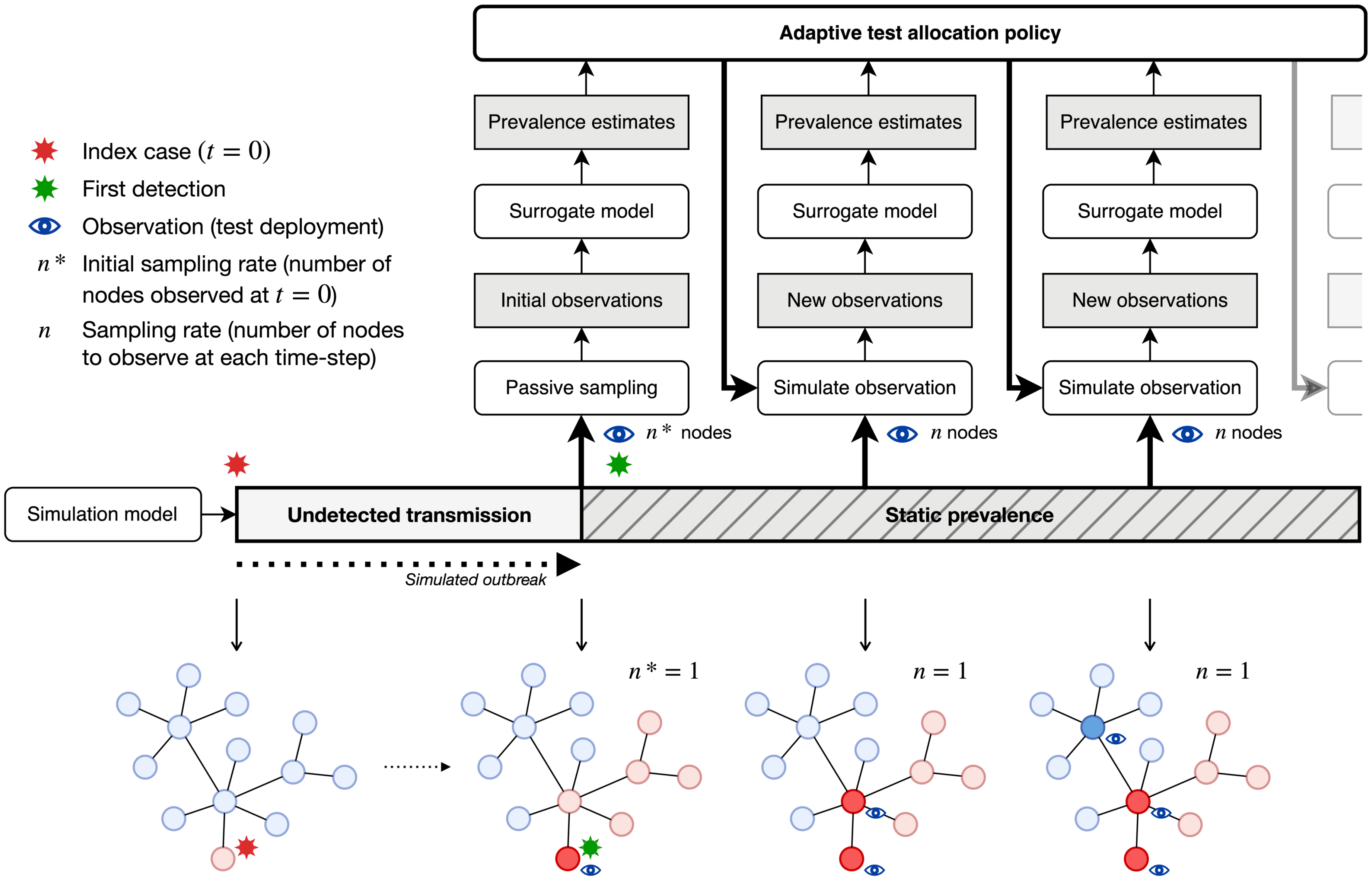
Collaborated with Joseph Tsui, Seth Flaxman, Elizaveta Semenova, Moritz Kraemer

Active Disease Surveillance

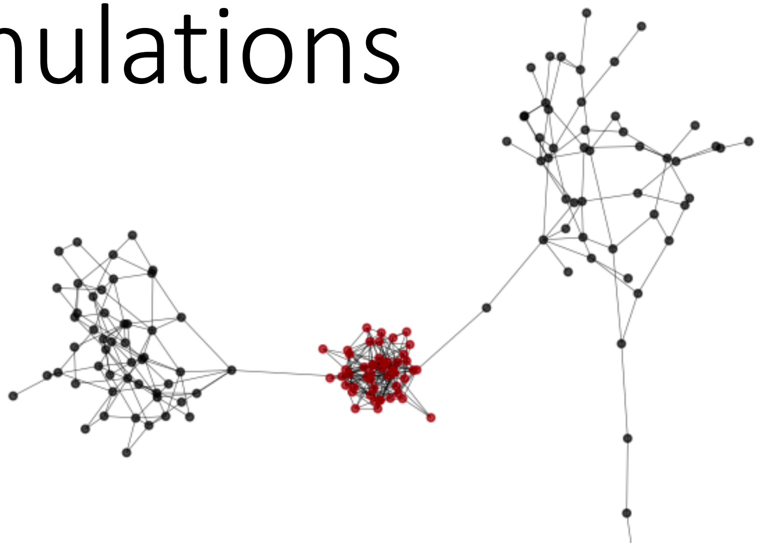
Node entropy

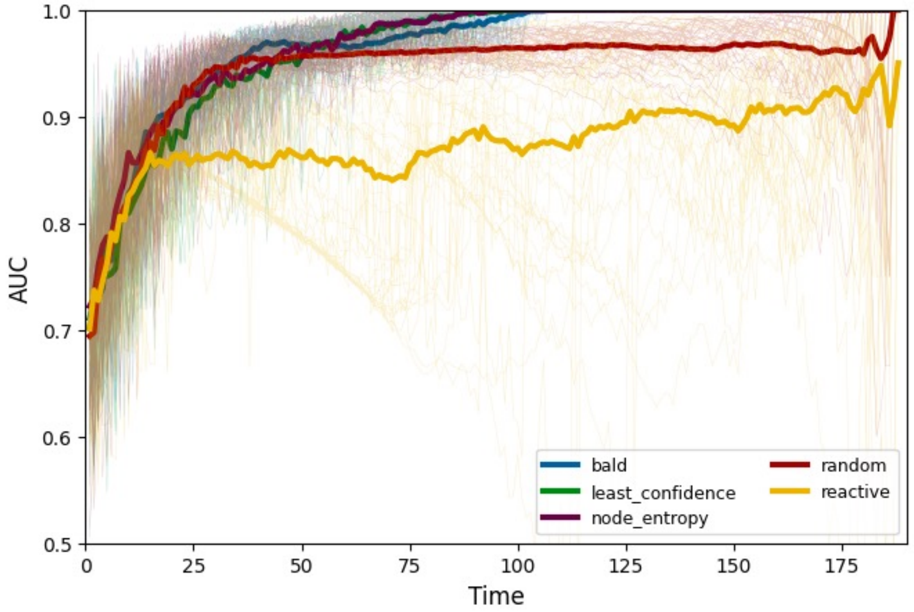
BALD $I[\boldsymbol{\theta}; y|\boldsymbol{x}, \mathcal{D}_t]$

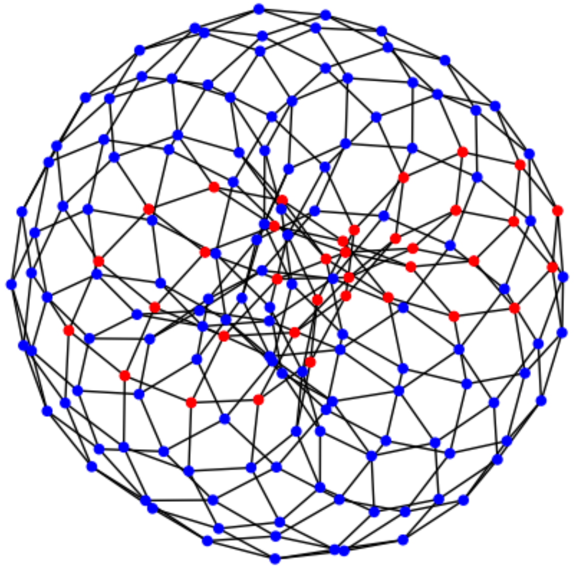
Conditional Auto-Regressive (CAR)

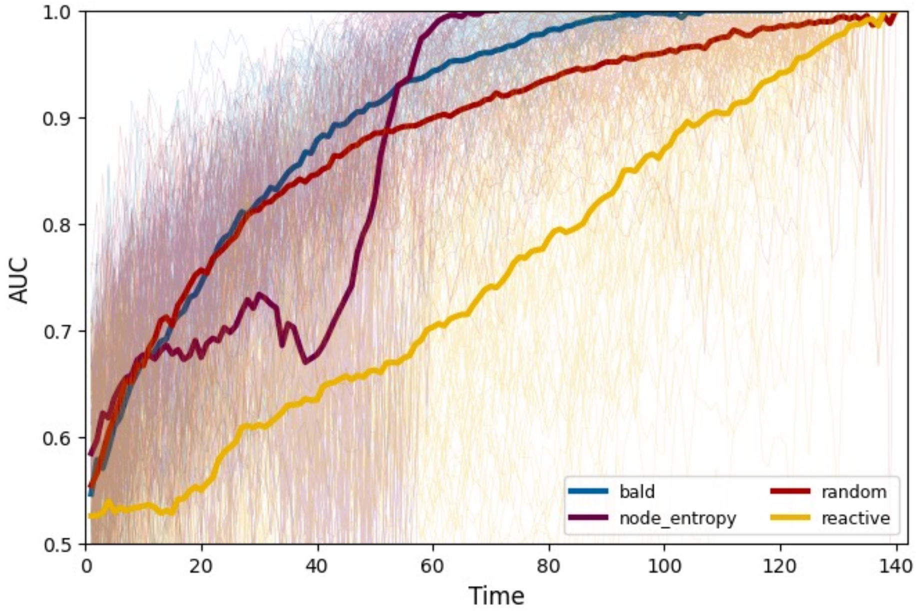


Simulations



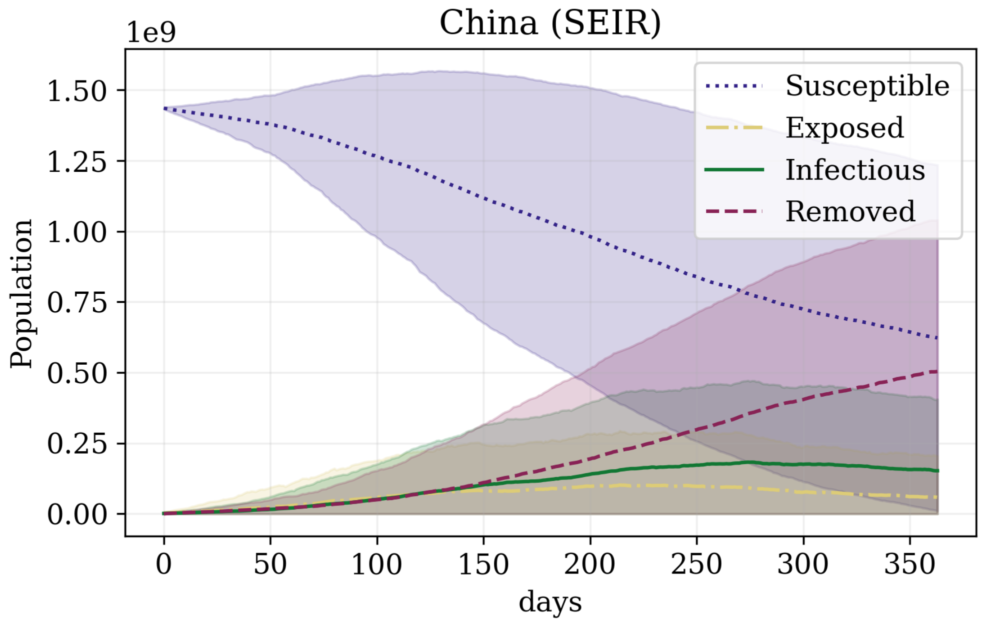




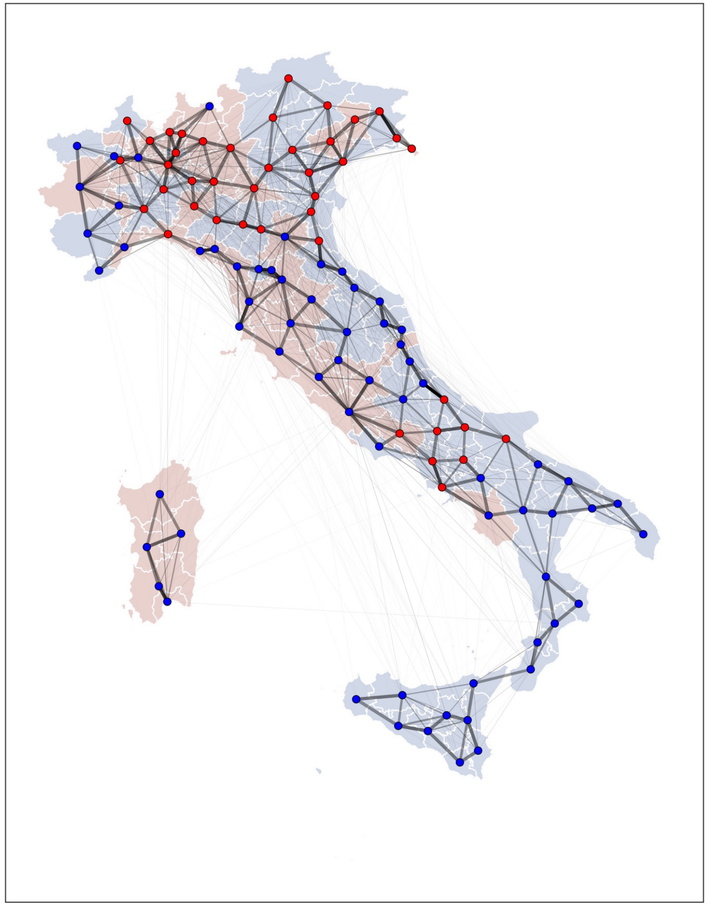


Next Steps

- Extend to GP on graph
- Extend to Dynamic Setting (prevalence changes wrt time) – SEIR Model



Italy network with thinning (only top 20%), with simulated prevalence (red indicated infected, blue uninfected)

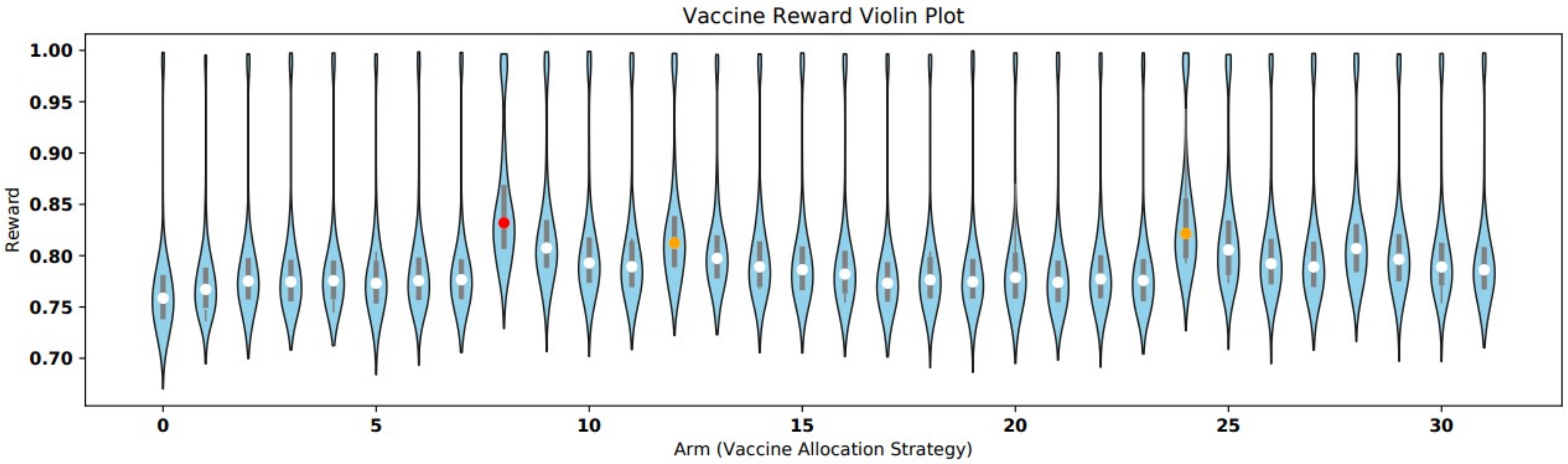


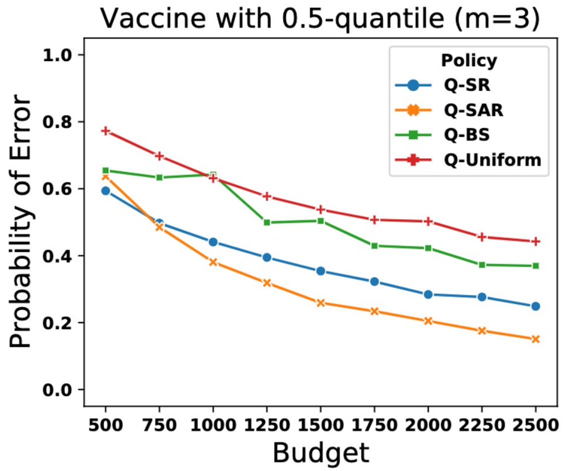
Pepe, E., Bajardi, P., Gauvin, L., Privitera, F., Lake, B., Cattuto, C., & Tizzoni, M. (2020). COVID-19 outbreak response, a dataset to assess mobility changes in Italy following national lockdown. Scientific Data 7, 230 (2020). https://data.humdata.org/dataset/covid-19-mobility-italy

https://github.com/kraemer-lab/pandemic_simulation_framework?tab=readme-ov-file

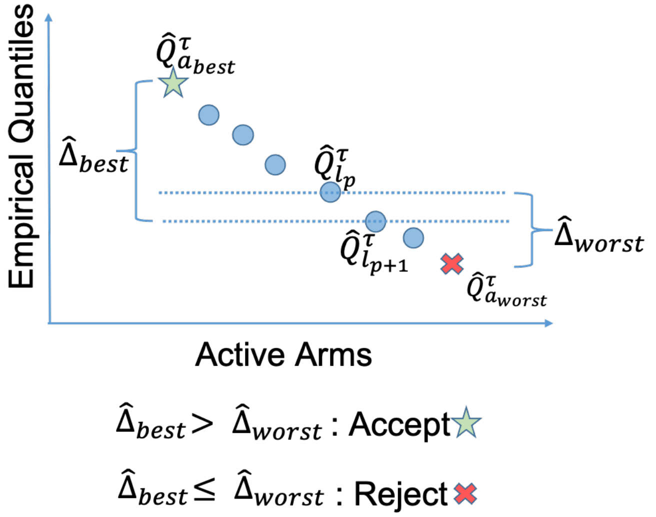
Applications: vaccine allocation

- Identify optimal strategies (highest **median** reward) for vaccine allocation
- **Arm**: vaccine allocation strategy (Allocate 100 vaccine doses to 5 age groups -- all combinations as arms)
- **Reward**: proportion of individuals that did not experience symptomatic infection





Contributions on Quantile BAI



New Algorithm: Quantile-based Successive Accepts and Rejects (Q-SAR)

New Concentration inequalities

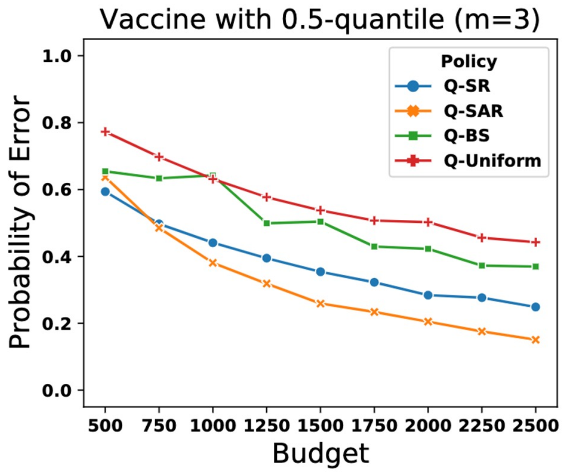
$$\mathbb{P}\left(Q^{\tau} - \hat{Q}^{\tau}_{n} \geq d^{l}_{n,\gamma}\right) \leq \exp(-\gamma)$$

$$\mathbb{P}\left(\hat{Q}^{\tau}_{n} - Q^{\tau} \geq d^{r}_{n,\gamma}\right) \leq \exp(-\gamma)$$

Probability of error

$$e_N := \mathbb{P}\left(\mathcal{S}^{N}_{m} \neq \mathcal{S}^{*}_{m}\right) \leq 2K^2 \exp\left(-\frac{N-K}{\overline{\log}(K) H^{\tau}}\right)$$

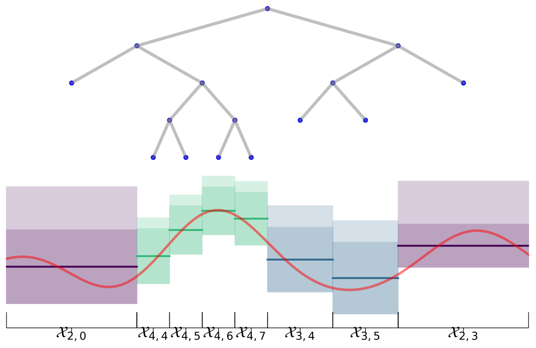
Experiments on vaccine allocation

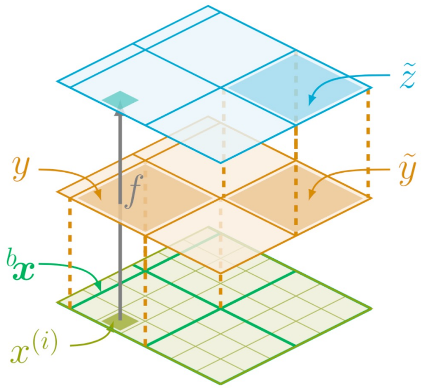


[1] Quantile Bandits for Best Arms Identification. **Mengyan Zhang**, Cheng Soon Ong. International Conference on Machine Learning 2021.

Conclusion & Future Work

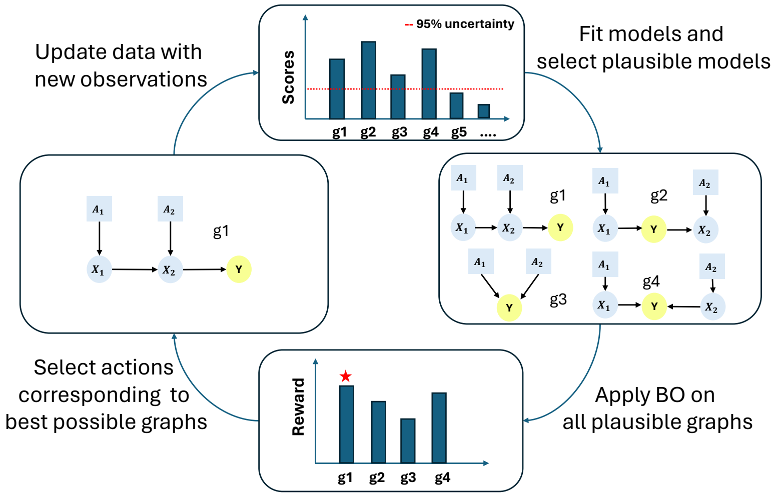
- Aggregated Feedback





  - Patrial monitoring
  - Cumulative regret setting
  - Instrumental Variable regression
  - Distributionally Robust BO

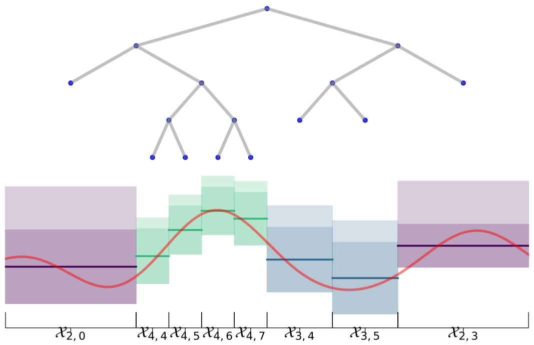
- Causal Decision making

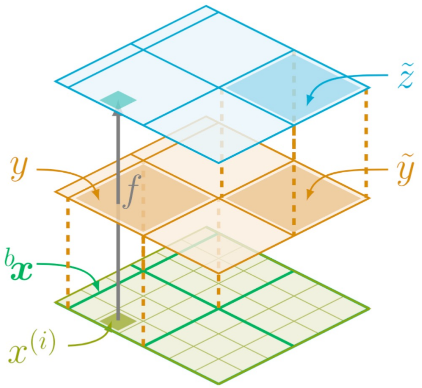


  - Theoretical analysis
  - Best arm identification setting
  - More scalable approach
  - With imperfect feedback

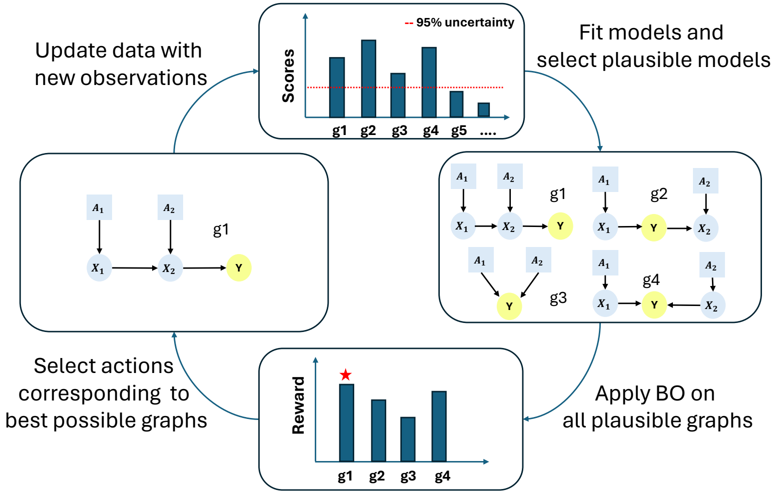
Conclusion & Future Work

- Aggregated Feedback

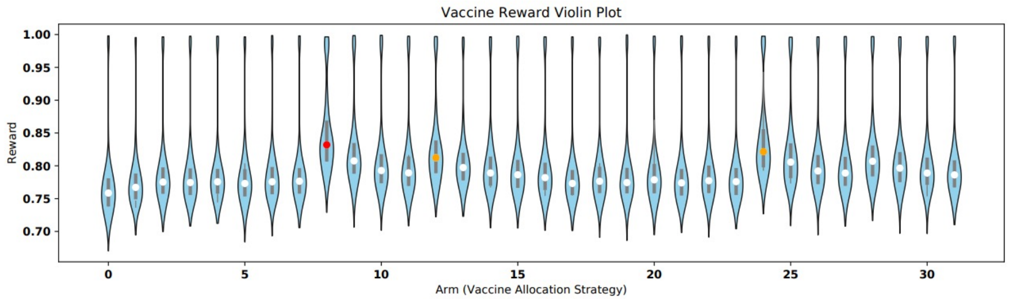


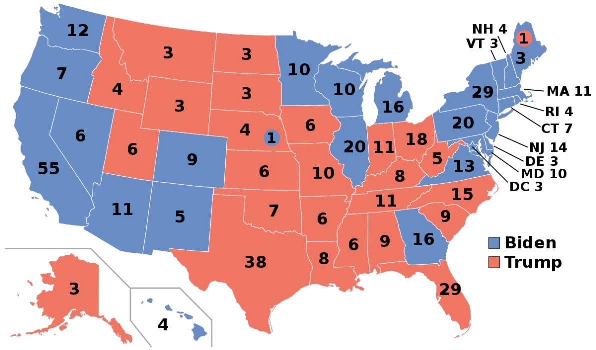


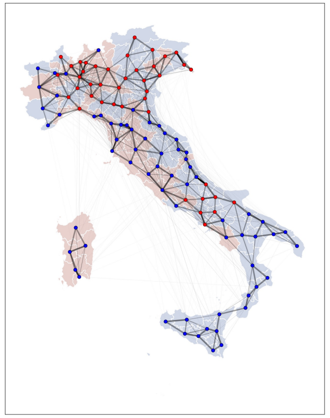
- Causal Decision making

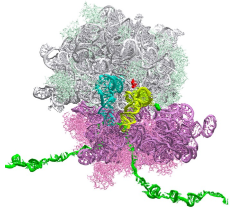


- … and public health/policy applications









Thanks for listening!

and looking forward to chatting with you :)

Mengyan Zhang
mengyan.zh@outlook.com
https://mengyanz.github.io/