Deep Adaptive Design: Amortizing Sequential Bayesian Experimental Design

Adam Foster*   Desi R. Ivanova*   Ilyas Malik   Tom Rainforth

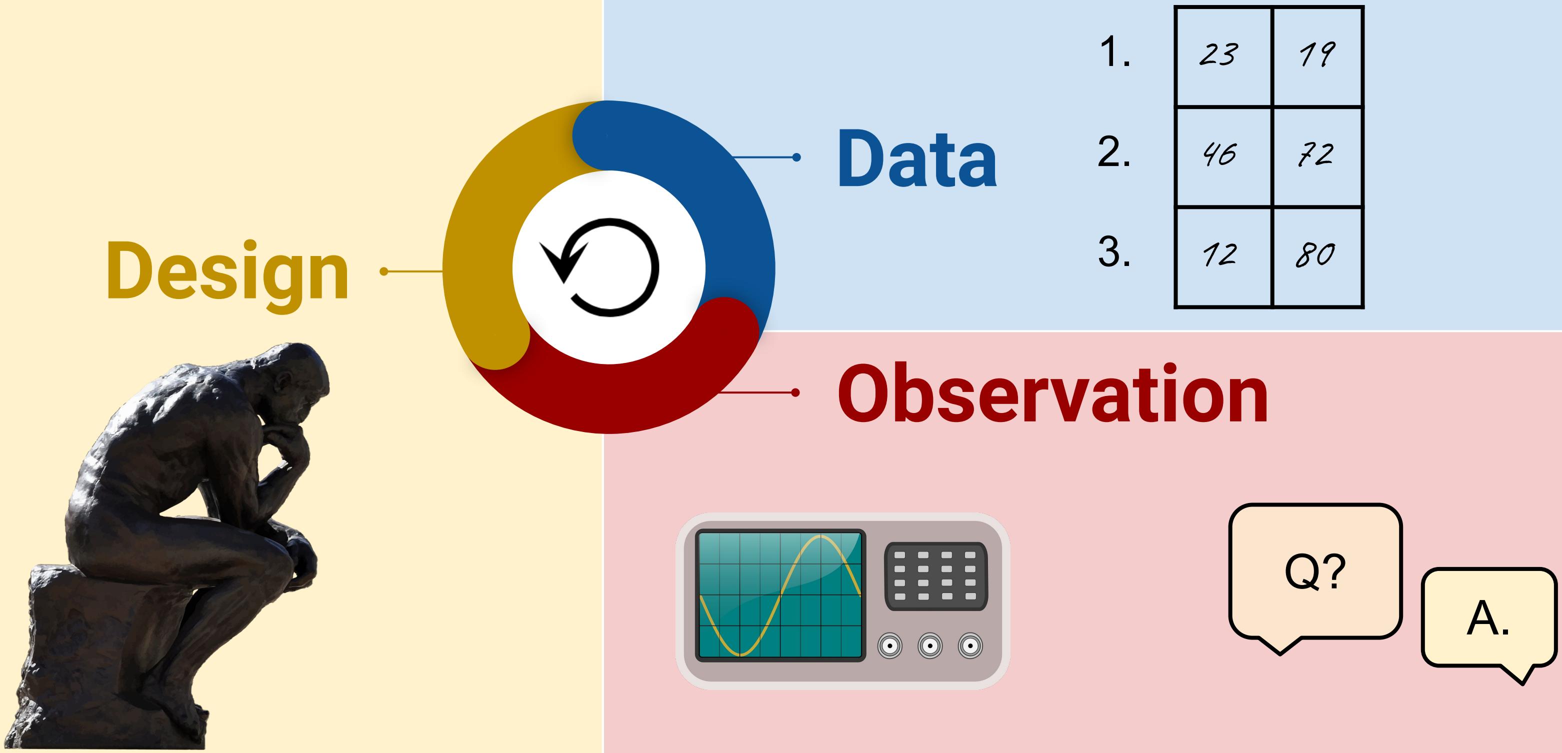
* equal contribution

UNIVERSITY OF OXFORD

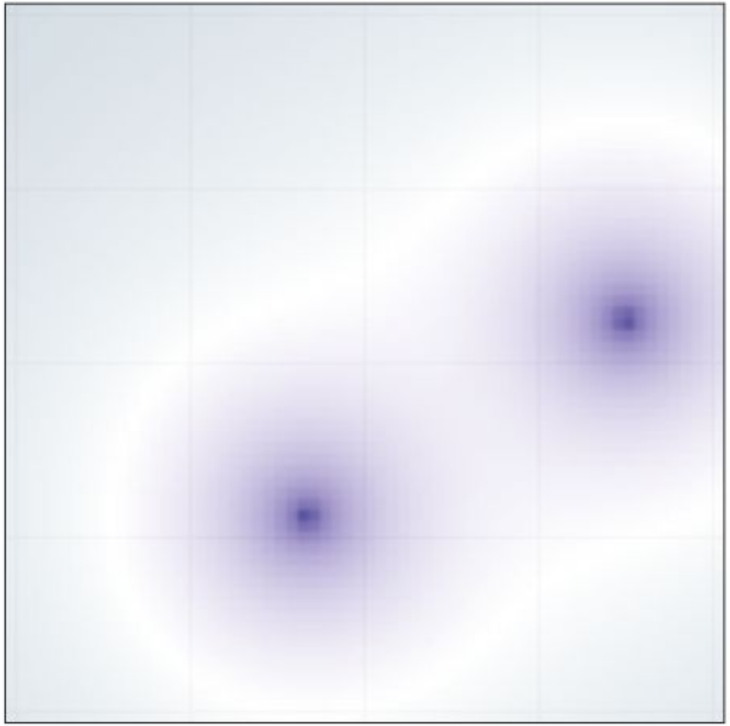
ICML 2021

Deep Adaptive Design (DAD)
enables
fast, adaptive experimentation.

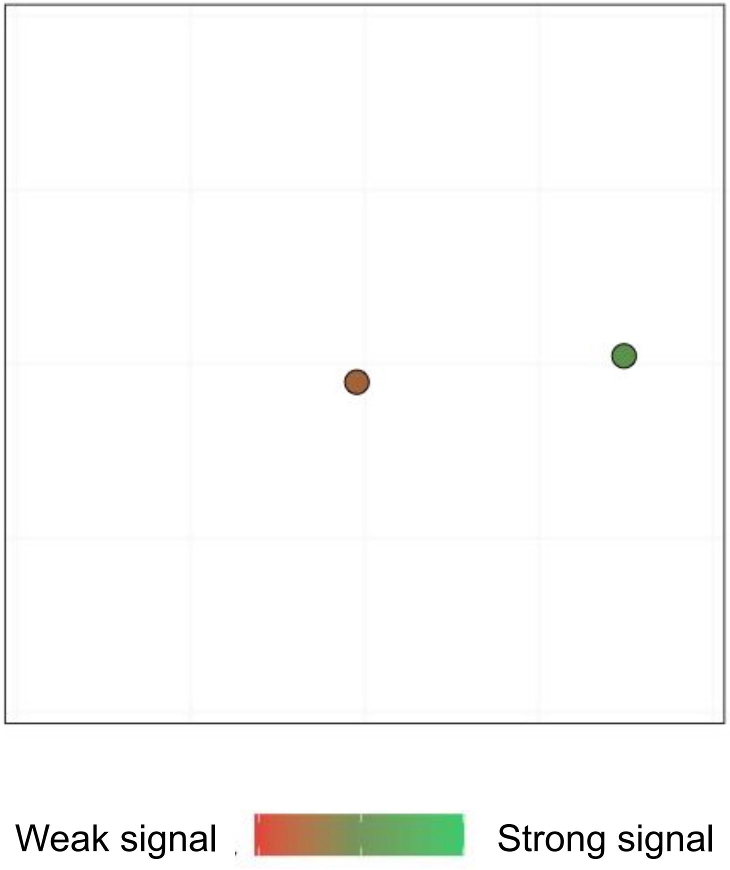
Adaptive experimentation



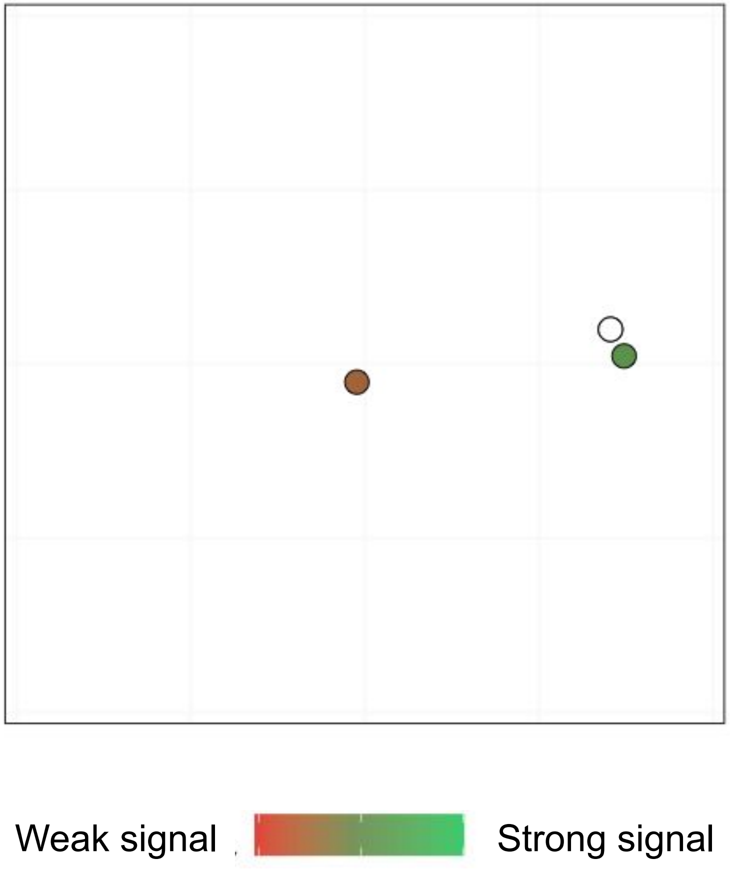
Example: discovering hidden sources



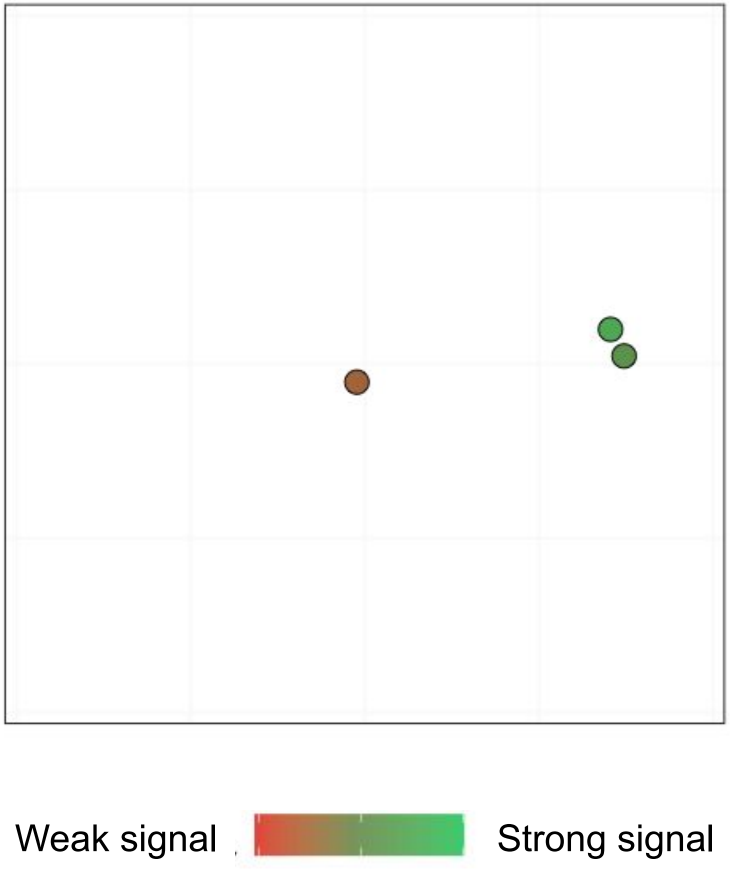
Gather initial data



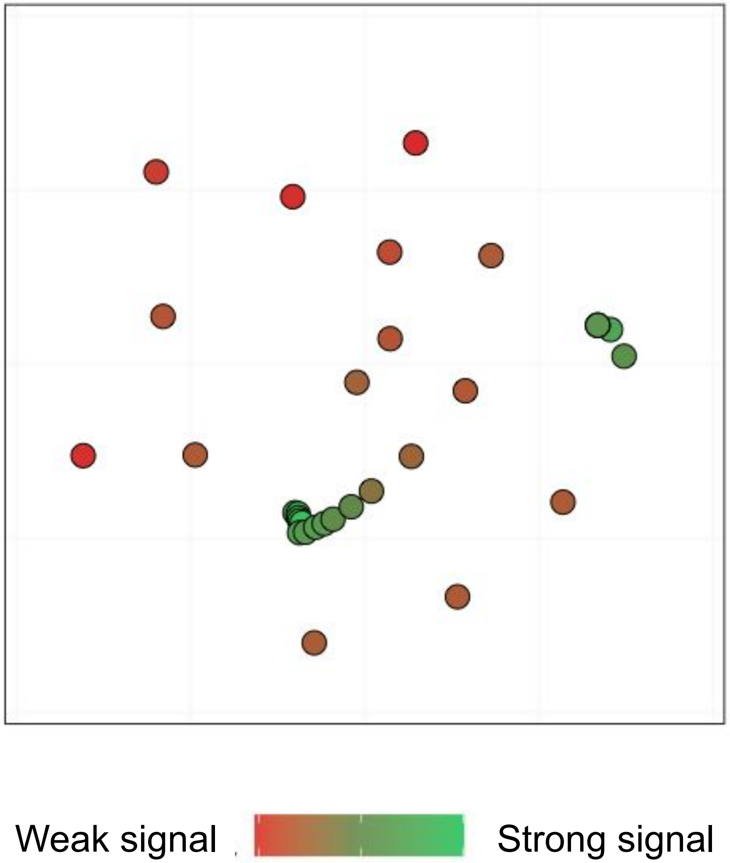
Use past data to select the next design



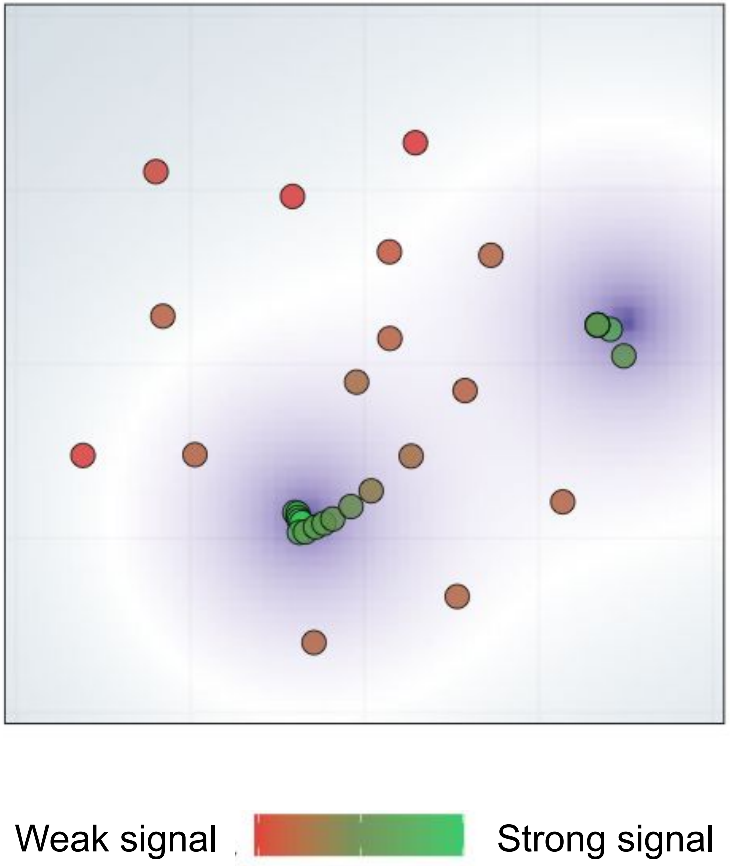
Observe at new design



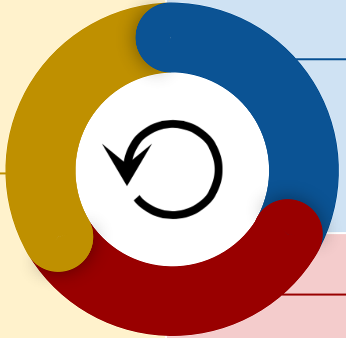
Iterate



Use final data to infer source locations



Bayesian experimental design



Design

Controlled by experimenter

ξ_t

Data

Past ξ's and y's

ξ_1	y_1
$\vdots$	$\vdots$
ξ_{t-1}	y_{t-1}

Observation

Outcome of doing the experiment at ξ_t

y_t produced with design ξ_t

Bayesian experimental design

Goal of the experiment: learn about *target of inference* θ

Bayesian prior $p(\theta)$

Likelihood model for the outcome of each experiment

$$y_t \sim p(y|\theta, \xi_t)$$

(Gelman et al., 2013; Kruschke, 2014)

Bayesian experimental design

Goal of the experimental design: choose designs so that the gathered data is informative about θ

Bayesian experimental design

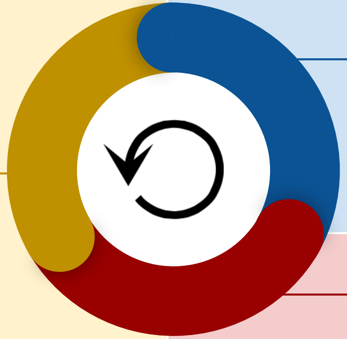
Choose ξ_t to maximize the **expected information gain (EIG)**

$$\mathbb{E}_{y_t|\xi_t}\left[\underbrace{H[p(\theta|\xi_{1:t-1}, y_{1:t-1})]}_{\text{Posterior entropy at } t\text{-}1} - \underbrace{H[p(\theta|\xi_{1:t}, y_{1:t})]}_{\text{Posterior entropy at } t}\right]$$

(Lindley, 1956)

Slow adaptive experimentation



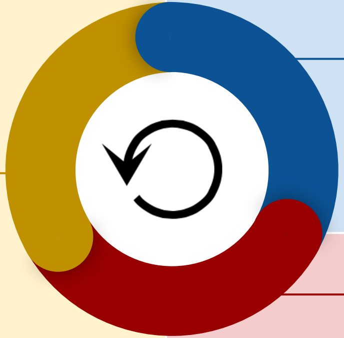
Design

Data

Observation

ξ_1	y_1
$\vdots$	$\vdots$
ξ_{t-1}	y_{t-1}

Slow adaptive experimentation



Design

1) Fit posterior

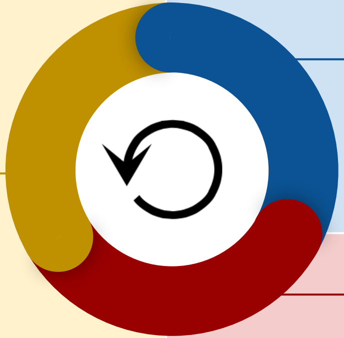
$$p(\theta|\xi_{1:t-1}, y_{1:t-1})$$

Data

ξ_1	y_1
$\vdots$	$\vdots$
ξ_{t-1}	y_{t-1}

Observation

Slow adaptive experimentation



Design

Data

ξ_1	y_1
$\vdots$	$\vdots$
ξ_{t-1}	y_{t-1}

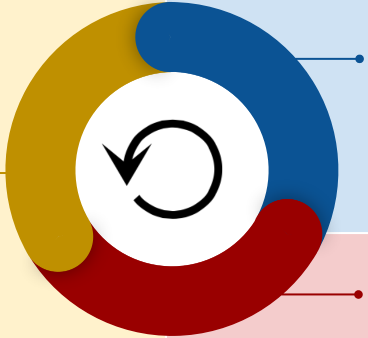
Observation

1) Fit posterior

$$p(\theta|\xi_{1:t-1}, y_{1:t-1})$$

2) Optimize EIG over ξ_t

Slow adaptive experimentation



Design

Data

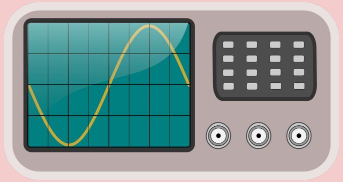
ξ_1	y_1
$\vdots$	$\vdots$
ξ_{t-1}	y_{t-1}

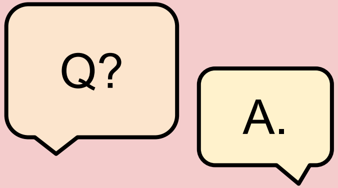
Observation

1) Fit posterior

$$p(\theta|\xi_{1:t-1}, y_{1:t-1})$$

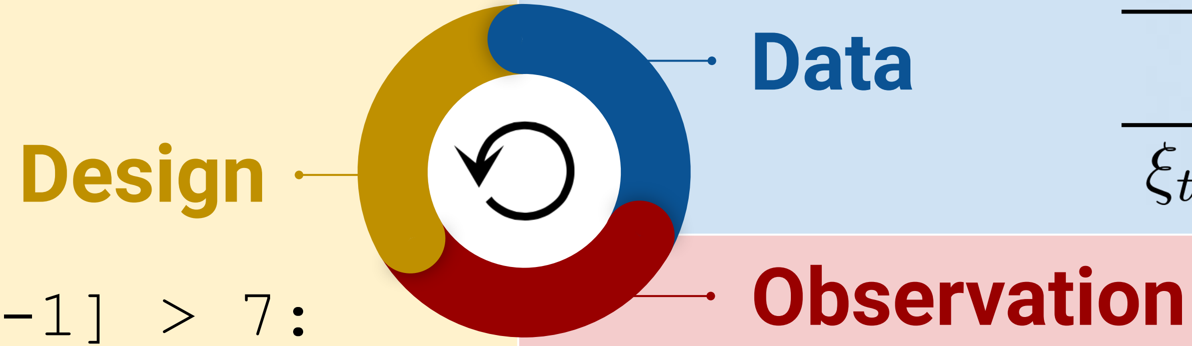
2) Optimize EIG over ξ_t





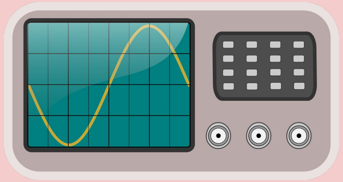
$$y_t \sim p(y|\theta, \xi_t)$$

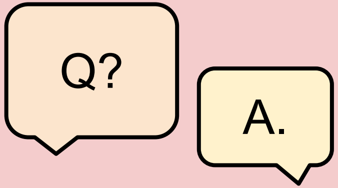
Fast heuristic experimentation



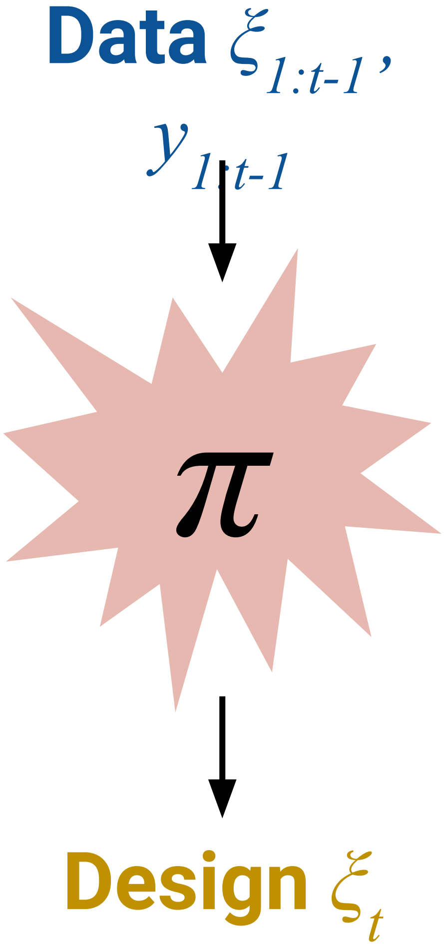
ξ_1	y_1
$\vdots$	$\vdots$
ξ_{t-1}	y_{t-1}

```
if y[-1] > 7:
    return 0
else:
    return 1
```



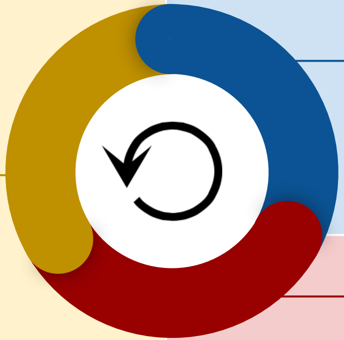


$$y_t \sim p(y|\theta, \xi_t)$$



Design Policy

Rethinking with policies



Design

Data

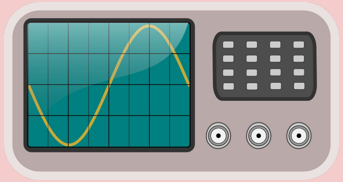
ξ_1	y_1
$\vdots$	$\vdots$
ξ_{t-1}	y_{t-1}

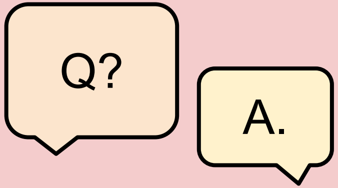
Observation

Apply policy π to data

$\xi_{1:t-1}, y_{1:t-1}$

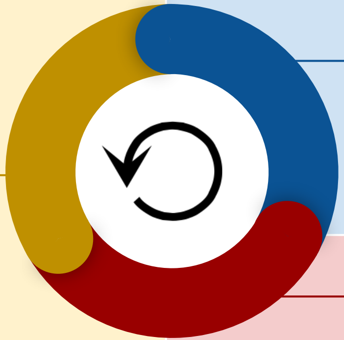
Output ξ_t





$$y_t \sim p(y|\theta, \xi_t)$$

Rethinking with policies



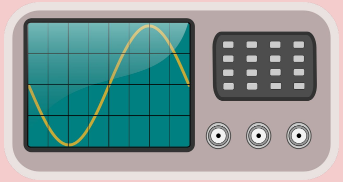
Design

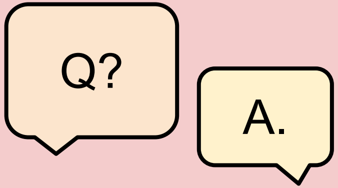
Apply policy π to data $h_{t\text{-}1}$

Output ξ_t

Data $h_{t-1} = \begin{array}{c|c} \xi_1 & y_1 \\ \hline \vdots & \vdots \\ \hline \xi_{t-1} & y_{t-1} \end{array}$

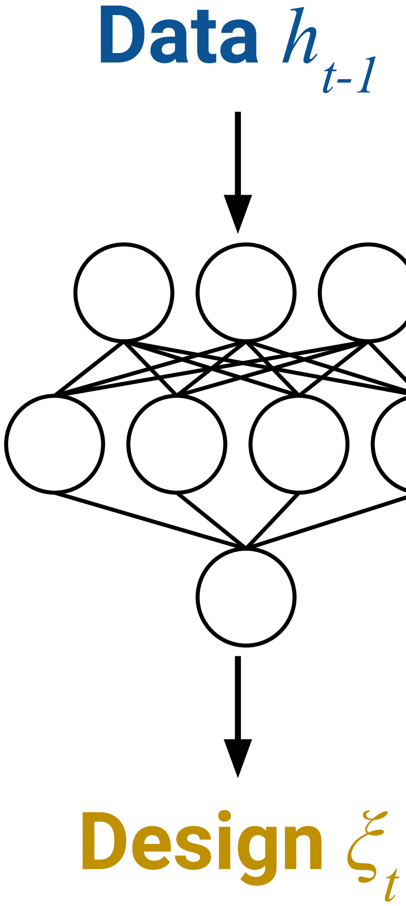
Observation





$$y_t \sim p(y|\theta, \xi_t)$$

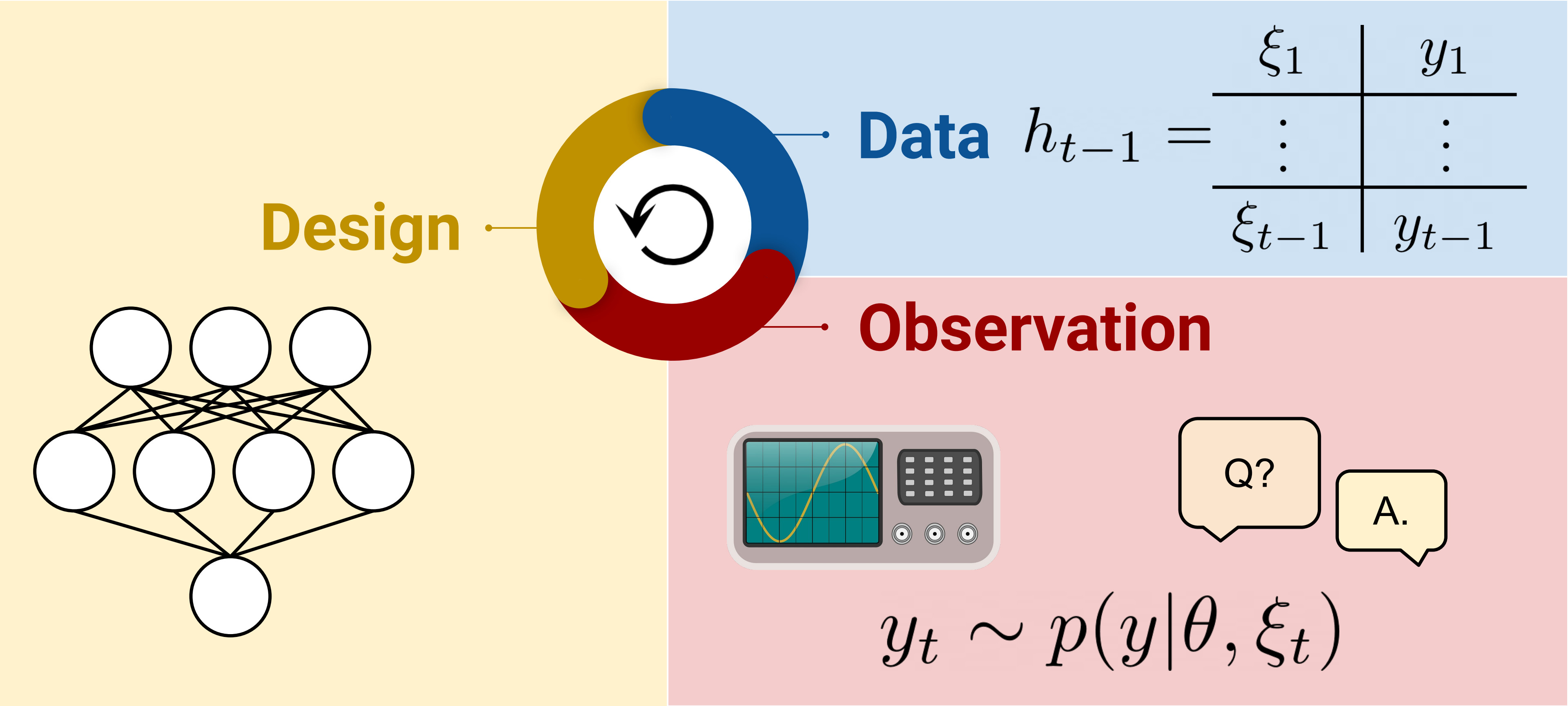
DAD: use a **neural network** as the **design policy**



Design Network

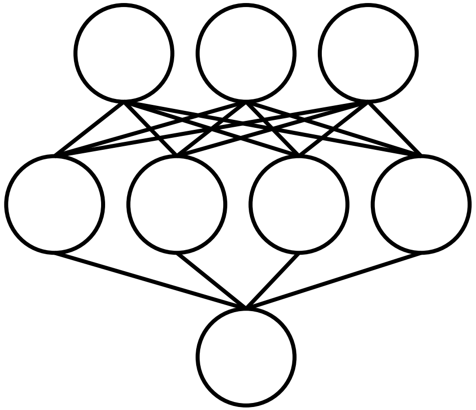
$$\pi_\phi$$

Design policy network

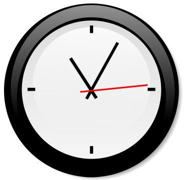


How to train the design network?

The design network

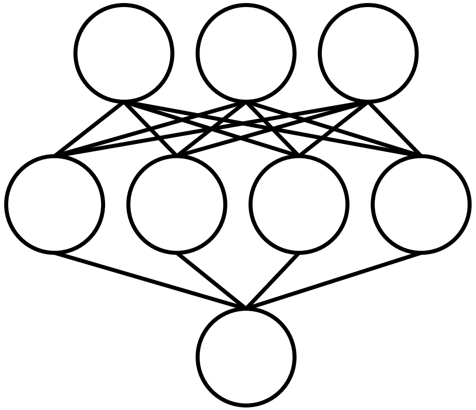


should **amortize** slow adaptive design



How to train the design network?

The design network



should **maximize** the expected sum of information gains

$$\mathbb{E}\left[\sum_{t=1}^{T} \mathrm{EIG}_t\right]$$

A unified objective

Theorem 1

The total objective for the design policy is

$$\mathcal{I}_T(\pi_\phi) = \mathbb{E}\left[\sum_{t=1}^{T} \mathrm{EIG}_t\right] = \mathbb{E}_{p(\theta)p(h_T|\theta,\pi_\phi)}\left[\log \frac{p(h_T|\theta,\pi_\phi)}{p(h_T|\pi_\phi)}\right]$$

where

$$p(h_T|\theta,\pi) = \prod_{t=1}^{T} p(y_t|\theta,\xi_t)$$

A unified objective

Theorem 1

The total objective for the design policy is

$$\mathcal{I}_T(\pi_\phi) = \mathbb{E}\left[\sum_{t=1}^{T} \mathrm{EIG}_t\right] = \mathbb{E}_{p(\theta)p(h_T|\theta,\pi_\phi)}\left[\log \frac{p(h_T|\theta,\pi_\phi)}{p(h_T|\pi_\phi)}\right]$$

where

$$p(h_T|\theta,\pi) = \prod_{t=1}^{T} p(y_t|\theta,\xi_t)$$

The DAD approach

Use a **neural net policy** π_ϕ and optimize the parameters ϕ to max $\mathcal{I}_T(\pi_\phi)$

! $\mathcal{I}_T(\pi_\phi)$ is not tractable

The DAD approach

Use a **neural net policy** π_ϕ and optimize the parameters ϕ to max $\cancel{\mathcal{I}_T(\pi_\phi)}$

! $\cancel{\mathcal{I}_T(\pi_\phi)}$ is not tractable

Optimize a tractable lower bound $\mathcal{L}_T(\pi_\phi)$

Lower bounding $\mathcal{I}_T(\pi_\phi)$

Sequential Prior Contrastive Estimation

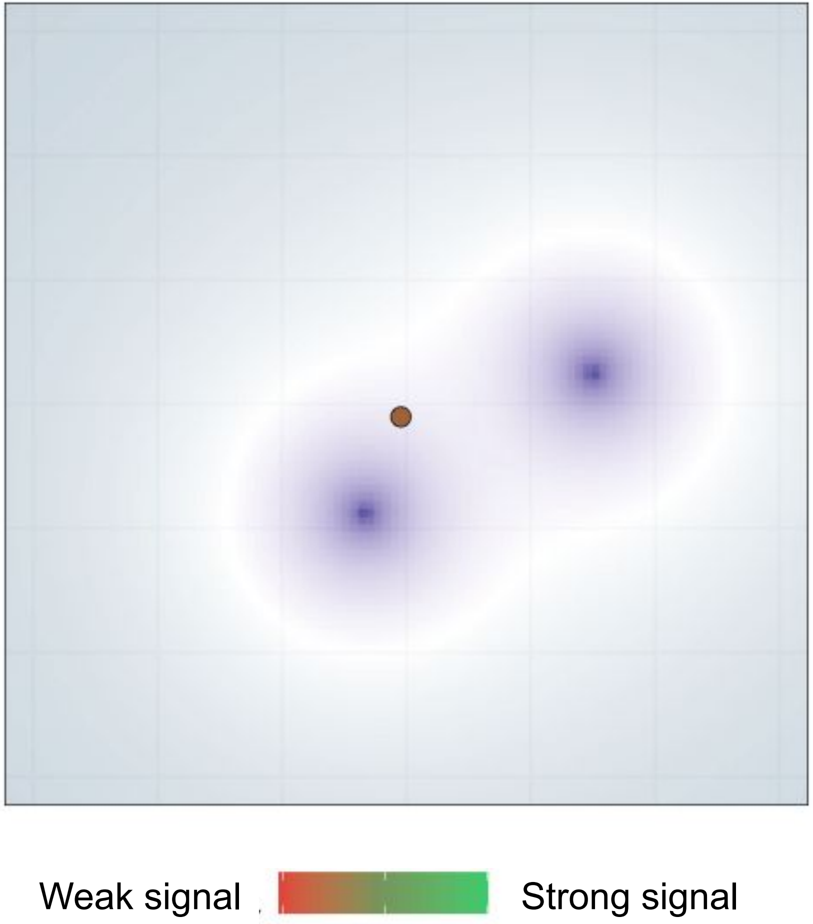
$$\mathcal{L}(\pi_\phi; L) = \mathbb{E}_{p(\theta_0)p(h_T|\theta_0,\pi_\phi)p(\theta_{1:L})}\left[\log \frac{p(h_T|\theta_0,\pi_\phi)}{\frac{1}{L+1}\sum_{\ell=0}^{L} p(h_T|\theta_\ell,\pi_\phi)}\right]$$

✓ Unbiased estimates

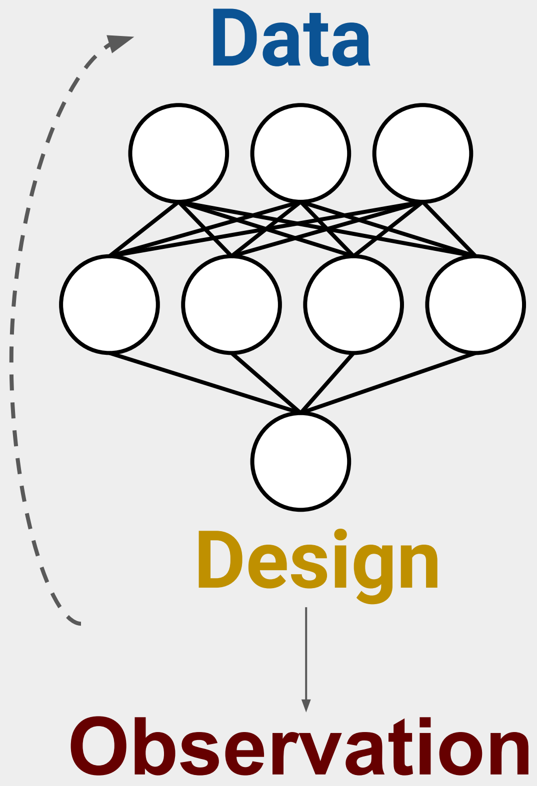
✓ Unbiased gradients w.r.t. ϕ

✓ Train π_ϕ end-to-end with SGA

DAD at work



DAD: key ingredients



Policy network π_ϕ

Unified objective $\mathcal{I}_T(\pi_\phi)$

Tractable lower bound $\mathcal{L}_T(\pi_\phi)$

How well does DAD do?

DAD	vs.	Traditional BOED approaches	
Policy-based		Variational myopic*	Fixed Strategy
✓ Adaptive		✓ Adaptive	✗ Adaptive
✓ Real-time		✗ Real-time	✓ Real-time

*SG-BOED of Foster et al (2020)

Location finding: Adaptive design

	$\mathcal{L}_{30}(\pi_\phi)$	$\mathcal{U}_{30}(\pi_\phi)$	Deployment time (sec.)
Variational	8.78	8.91	8963
DAD	10.93	12.38	0.005

>1mil times faster

Location finding: Adaptive design

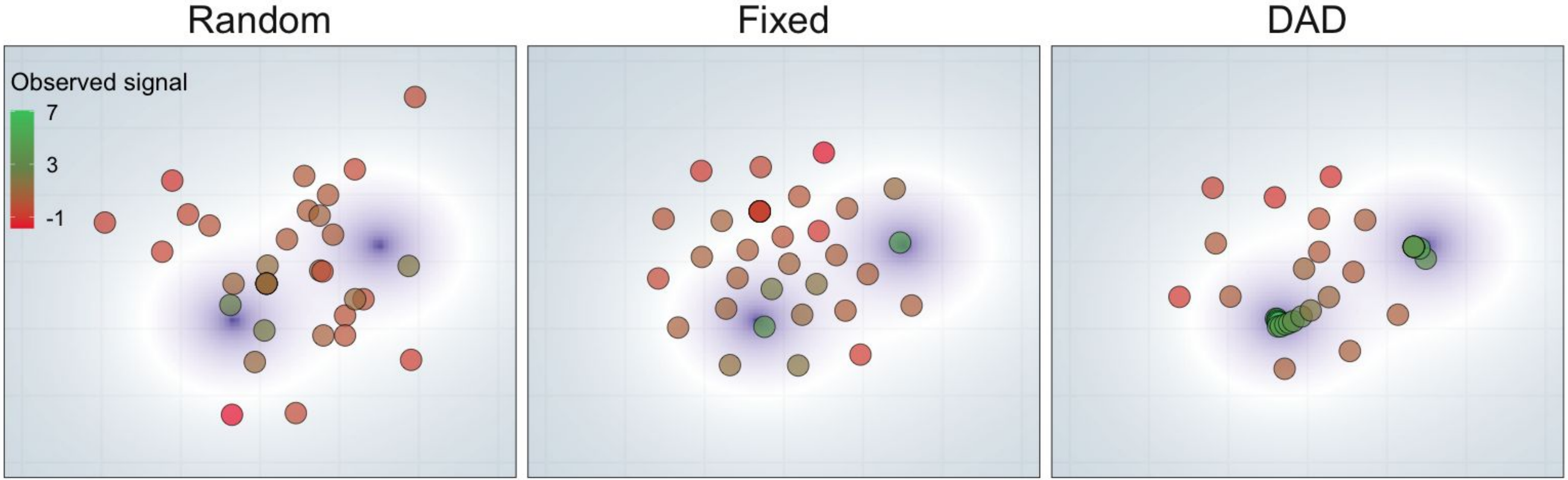
	$\mathcal{L}_{30}(\pi_\phi)$	$\mathcal{U}_{30}(\pi_\phi)$	Deployment time (sec.)
Variational	8.78	8.91	8963
DAD	10.93	12.38	0.005

>1mil times faster

✓ No need for posterior estimation

✓ Non-myopic properties of DAD

Location finding: real-time design



$$\mathcal{L}_{30}(\pi_\phi) = 8.30 \qquad \mathcal{L}_{30}(\pi_\phi) = 8.84 \qquad \mathcal{L}_{30}(\pi_\phi) = 10.93$$

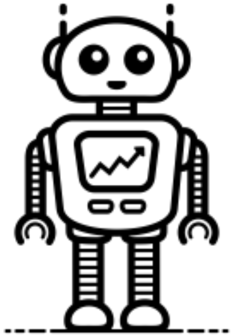
$$\mathcal{U}_{30}(\pi_\phi) = 8.32 \qquad \mathcal{U}_{30}(\pi_\phi) = 8.91 \qquad \mathcal{U}_{30}(\pi_\phi) = 12.38$$

Adaptive online surveys

AI powered by DAD

Would you prefer **$R** today, or $100 in **D days**?

$$\xi = (R, D)$$



Adaptive online surveys

AI powered by DAD

Would you prefer **$49** today, or $100 in **3 months**?

Human participant

$49 today

Adaptive online surveys

AI powered by DAD

Would you prefer **$49** today, or $100 in **3 months**?

Human participant

$49 today

AI powered by DAD

Would you prefer **$49** today, or $100 in **4 weeks**?

Human participant

$49 today

Adaptive online surveys

AI powered by DAD

Would you prefer **$49** today, or $100 in **3 months**?

Human participant

$49 today

AI powered by DAD

Would you prefer **$49** today, or $100 in **4 weeks**?

Human participant

$49 today

AI powered by DAD

Would you prefer **$49** today, or $100 in **16 days**?

Human participant

$100 in 16 days

Adaptive online surveys

AI powered by DAD

Would you prefer **$49** today, or $100 in **3 months**?

Human participant

$49 today

AI powered by DAD

Would you prefer **$49** today, or $100 in **4 weeks**?

Human participant

$49 today

AI powered by DAD

Would you prefer **$49** today, or $100 in **16 days**?

Human participant

$100 in 16 days

AI powered by DAD

Would you prefer **$49** today, or $100 in **23 days**?

Human participant

$49 today

Temporal discounting model

AI powered by DAD

Would you prefer **$49** today, or $100 in **3 months**?

Human participant

$49 today

AI powered by DAD

Would you prefer **$49** today, or $100 in **4 weeks**?

Human participant

$49 today

AI powered by DAD

Would you prefer **$49** today, or $100 in **16 days**?

Human participant

$100 in 16 days

AI powered by DAD

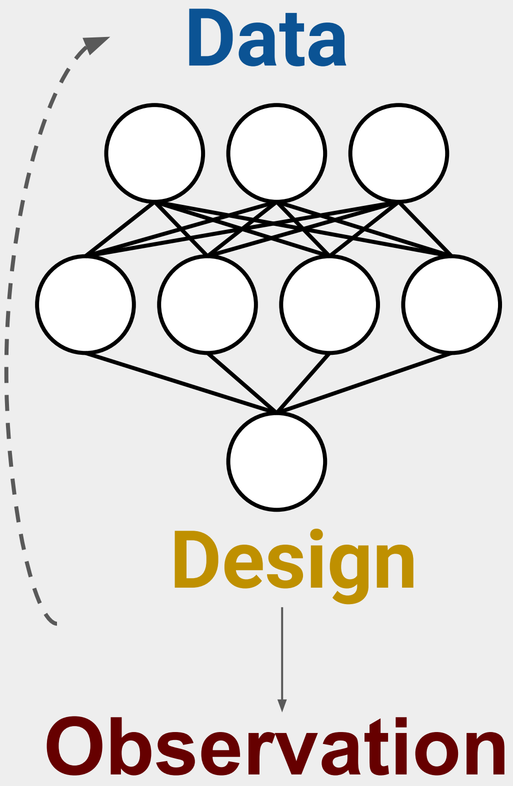
Would you prefer **$49** today, or $100 in **23 days**?

Human participant

$49 today

Hyperbolic temporal discounting

	$\mathcal{L}_{30}(\pi_\phi)$	$\mathcal{U}_{30}(\pi_\phi)$	Deployment time (sec.)
Heuristic (Frye et al. 2006)	3.50	3.51	0.09
Traditional (Vincent & Rainforth 2017)	4.45	4.54	25.27
DAD	**5.02**	**5.12**	**0.09**



Conclusion

Policy-based
Amortized
Non-myopic

Deep Adaptive Design

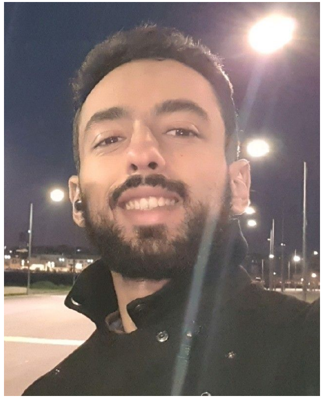
Design network

Unified objective

Tractable lower bound

Adaptive experiments in real-time

Thank you



UNIVERSITY OF OXFORD

Implementation in Pyro https://github.com/ae-foster/dad

Full paper https://arxiv.org/pdf/2103.02438.pdf