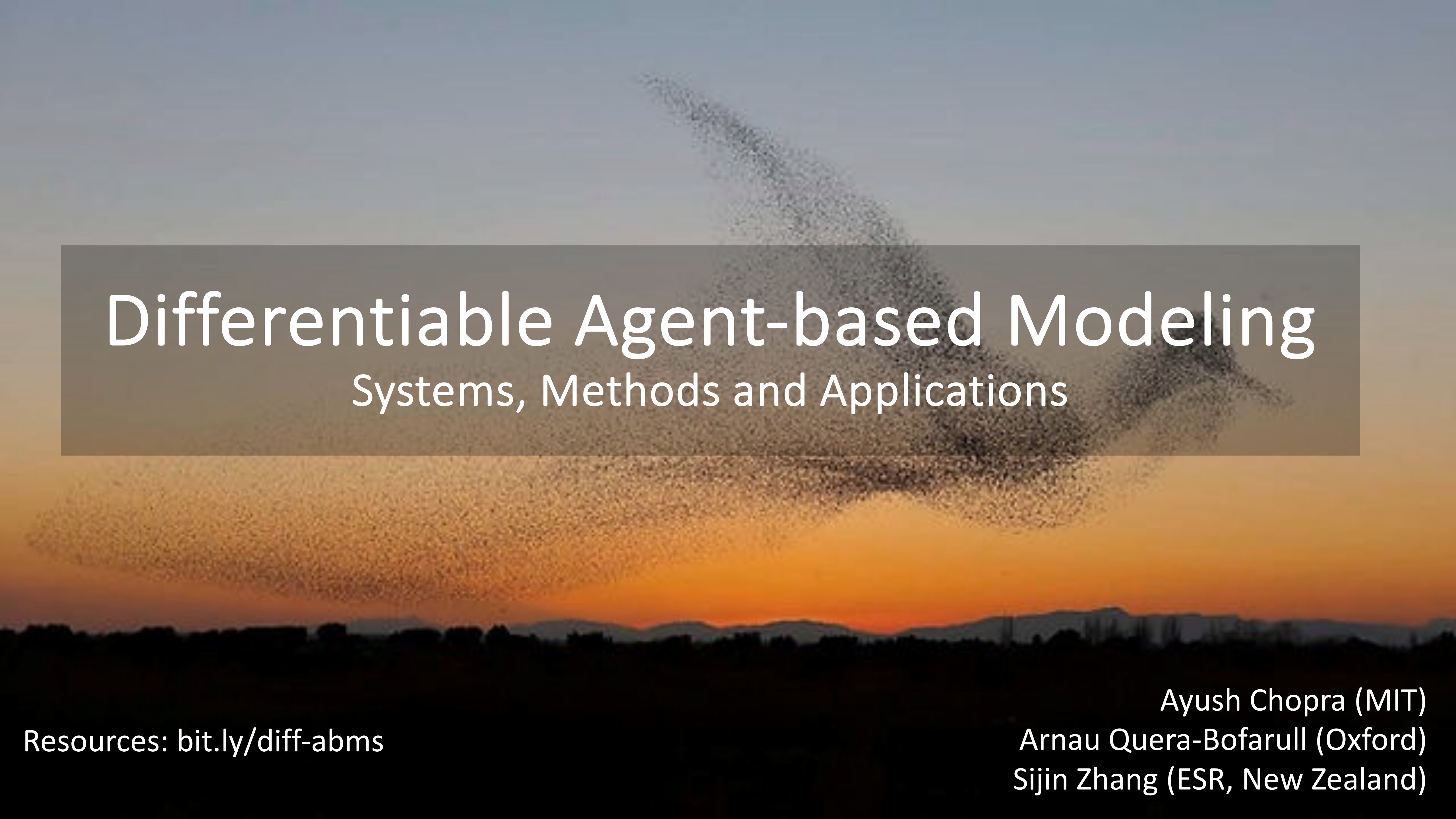


Differentiable Agent-based Modeling

Systems, Methods and Applications

Resources: bit.ly/diff-abms

Ayush Chopra (MIT)
Arnau Quera-Bofarull (Oxford)
Sijin Zhang (ESR, New Zealand)

Speakers

Ayush Chopra

PhD Candidate

MIT Media Lab

Arnau Quera-Bofarull

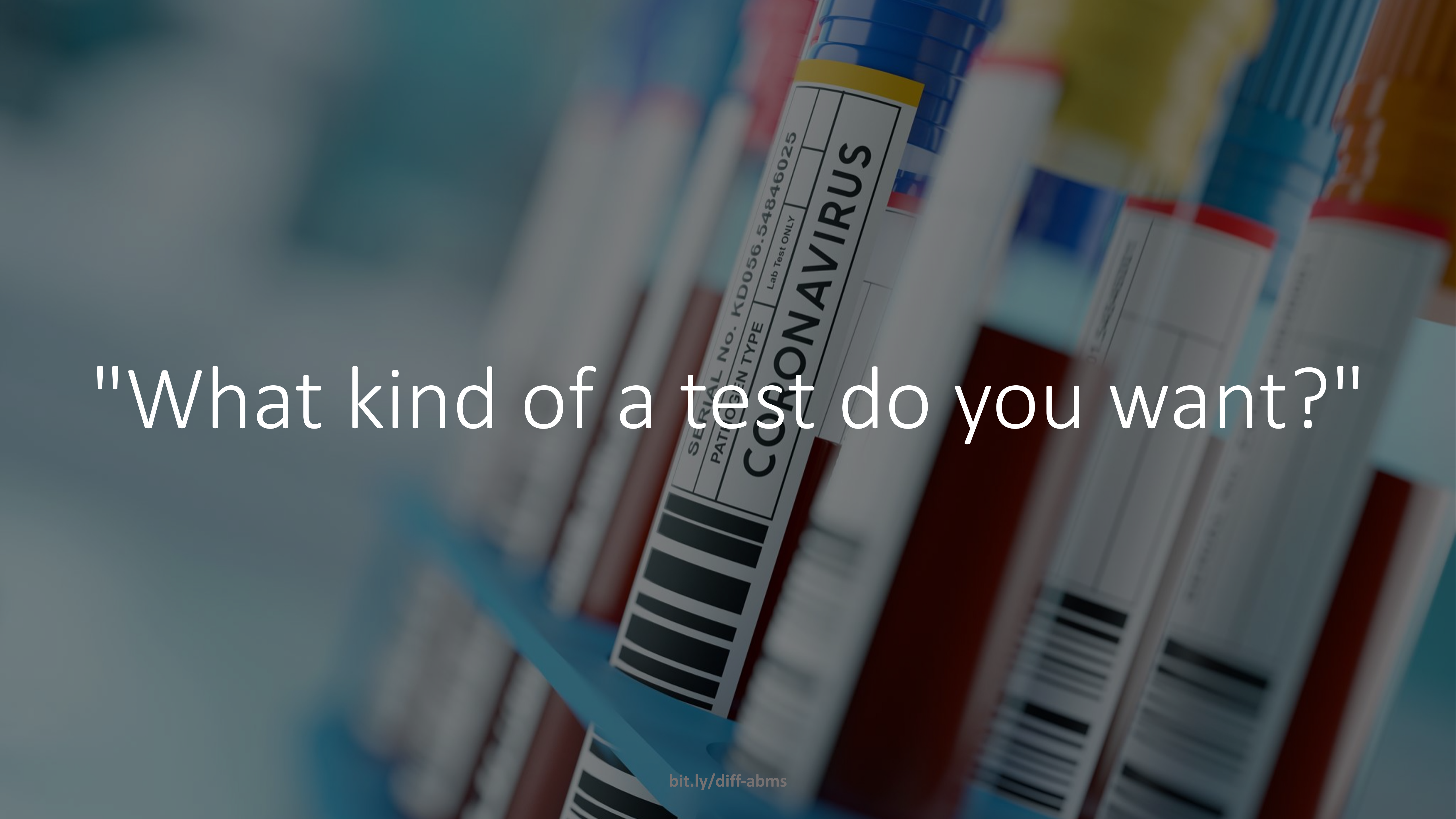
Postdoc Researcher

University of Oxford

Sijin Zhang

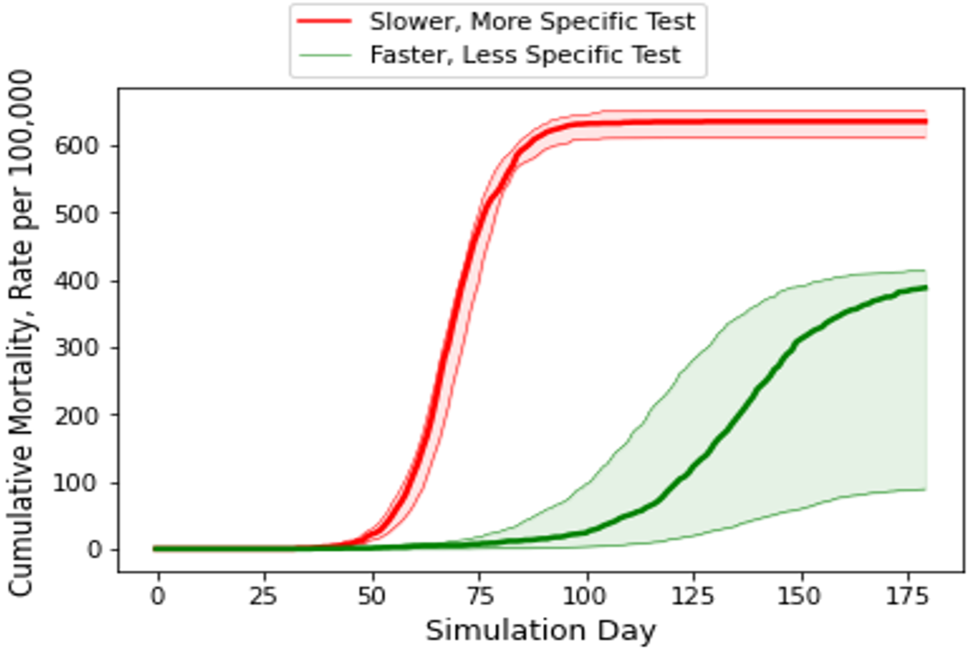
Senior Scientist

ESR, New Zealand



"What kind of a test do you want?"

collective prioritizes test speed over accuracy...



Collective outcomes can be ***very different*** from the sum of individual choices modeling collective behavior is critical


build bridges?

bit.ly/diff-abms

Collective behavior across scales and substrates

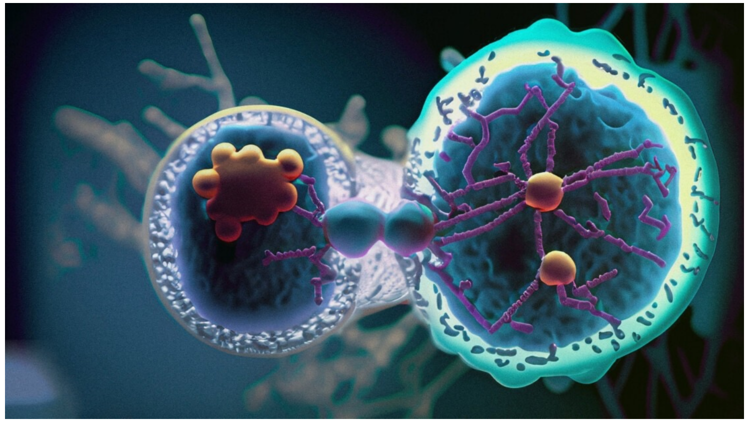
Cities

Supply Chains

Citizens

Pandemics

Cells



Morphogenesis

Collective behavior across scales and substrates

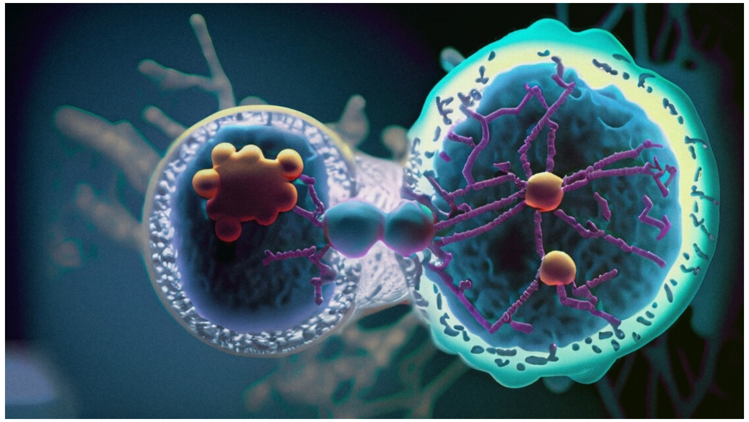
Cities

Supply Chains

Citizens

Pandemics

Cells

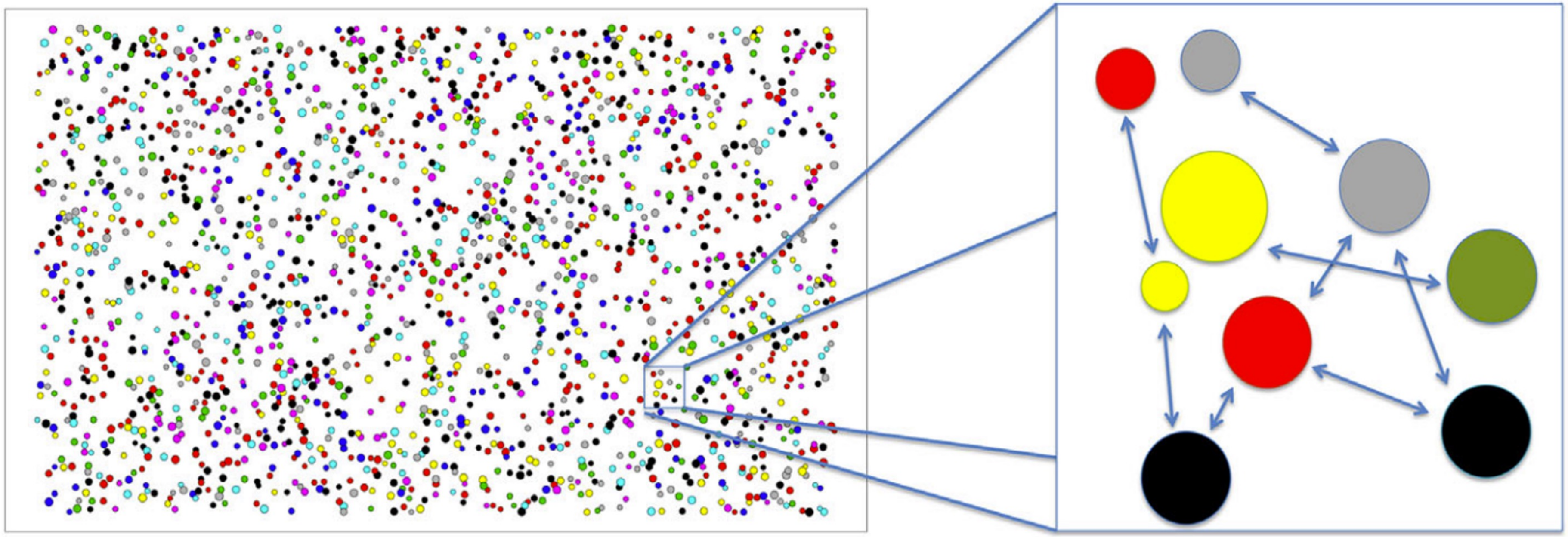


Morphogenesis

how to capture?

bit.ly/diff-abms

Agent-based Models

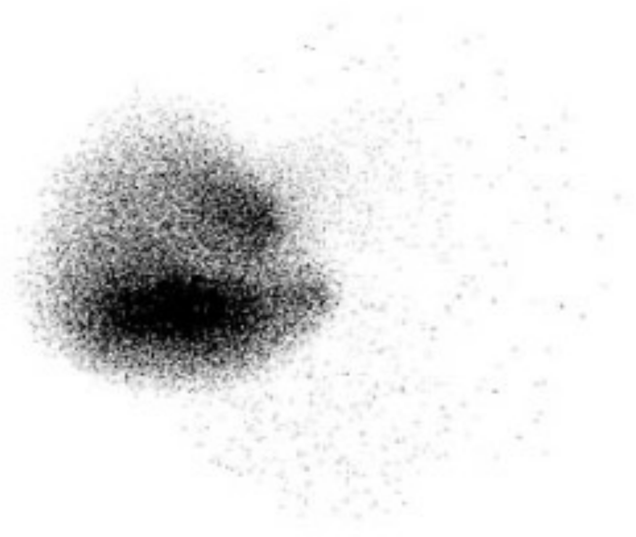


Simulate microscopic behavior and interactions in heterogeneous collectives

ABMs vs Multi-Agent Reinforcement Learning

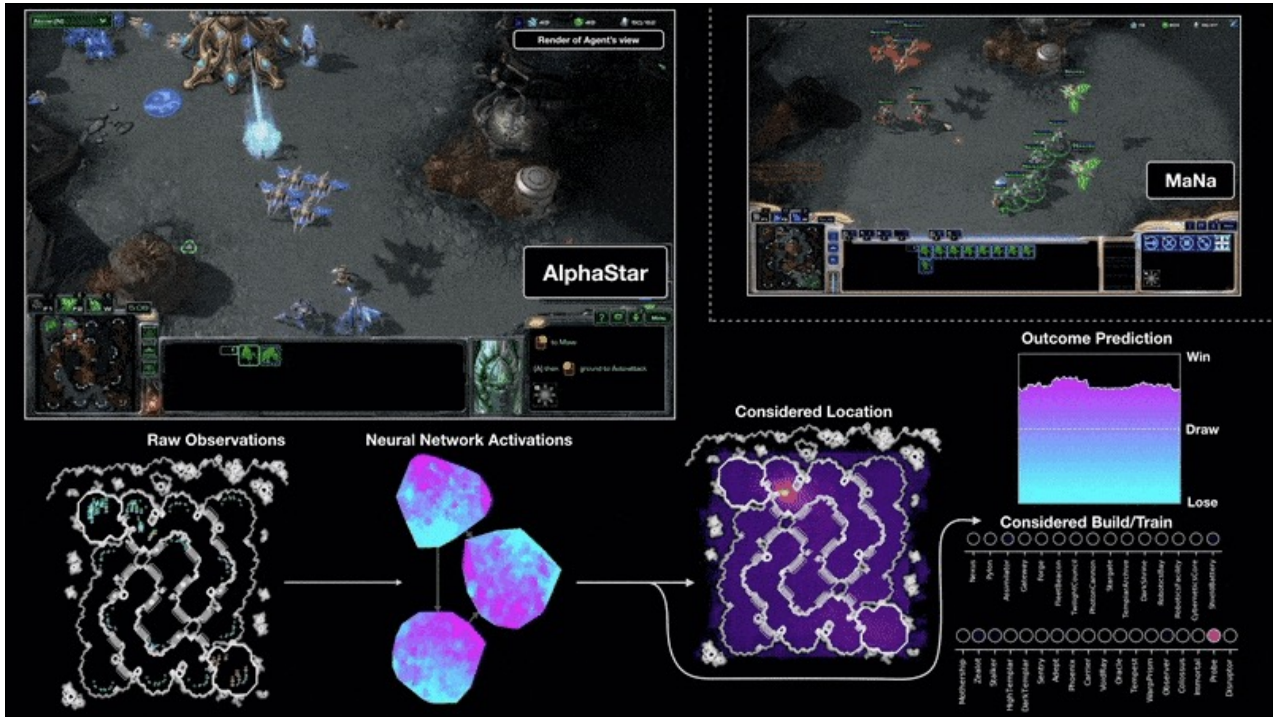
ABMs
- Many agents
- Simple behavior

MARL
- Few agents
- Complicated behavior



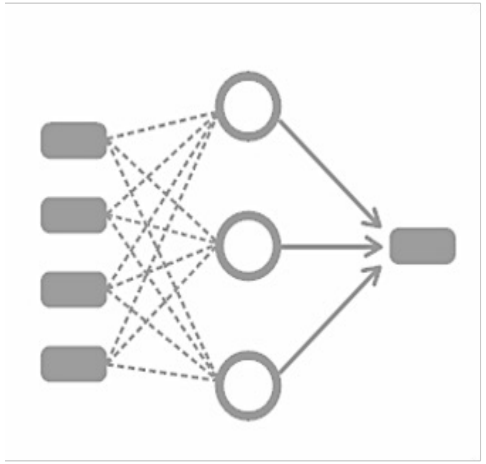
Flocking birds

bit.ly/diff-abms



Starcraft2 (AlphaStar)

long history of research and open challenges



Computation

Simulation?
Calibrate?
Analyze?

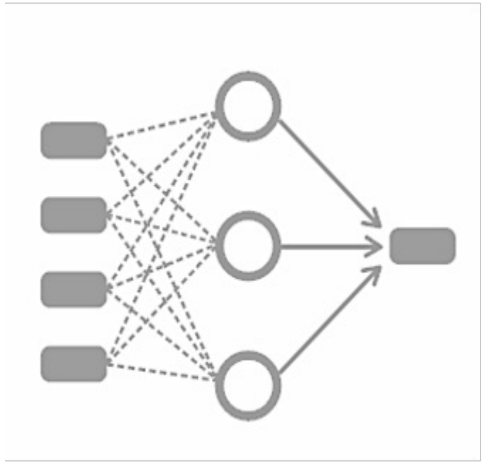
Data

Multi-modal?
Multi-scale?
Distributed?

Expressiveness

Behaviour?
Mechanism?
Real-world feedback?

Proposal: Differentiable Agent-based Modeling



Computation	Data	Expressiveness
Simulation? Calibrate? Analyze?	Multi-modal? Multi-scale? Distributed?	Behaviour? Mechanism? Real-world feedback?
Vectorization	**Gradient-based learning**	**Neural Network composition**

Agent-based Model

$$\theta \longrightarrow \boxed{\text{Simulator}} \longrightarrow x$$

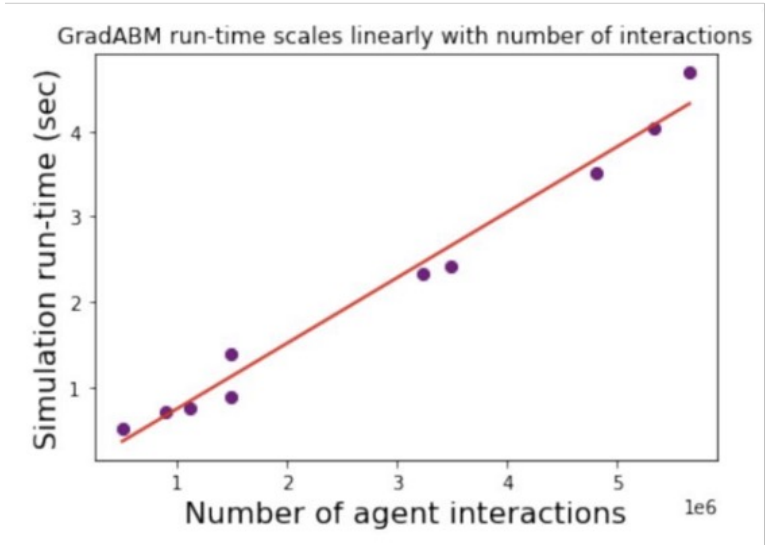
Differentiable if

$$\nabla_\theta \mathbb{E}[f(\theta)] \quad \text{exists}$$

Why do we care about the gradient?

Simulate **country-scale ecosystems** for few hundred dollars on commodity hardware

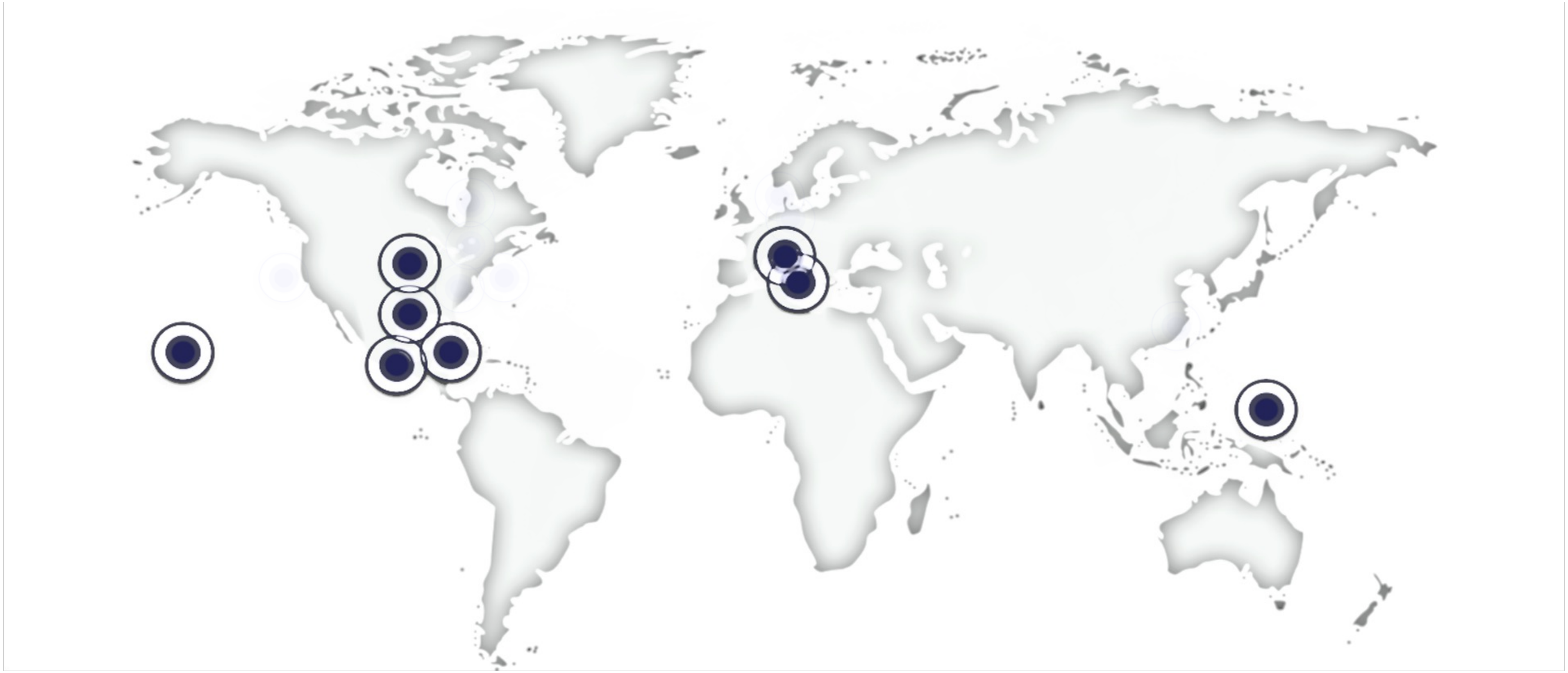
Method	Simulation	Calibration	Analysis
ABM	50 hours	100,000 hours	5,000 hours
Differentiable ABM	5 minutes	20 minutes	10 seconds



https://royalsocietypublishing.org/doi/10.1098/rsos.210506

8 million London agents

Differentiable ABMs are being deployed across domains



bit.ly/diff-abms

Scope of tutorial

- Preliminaries
  - Background to automatic differentiation
  - Implement a differentiable ABM
- Algorithms
  - Techniques to calibrate and analyze differentiable ABMs
- Applications
  - Real-world case study in New Zealand
- Systems
  - Tooling to build and calibrate differentiable ABMs at scale

Scope of tutorial

- Preliminaries
  - Background to automatic differentiation
  - Implement a differentiable ABM
- Algorithms
  - Techniques to calibrate and analyze differentiable ABMs
- Applications
  - Real-world case study in New Zealand
- Systems
  - Tooling to build and calibrate differentiable ABMs at scale

Automatic Differentiation

Stochastic Automatic Differentiation

Implement a Differentiable ABM

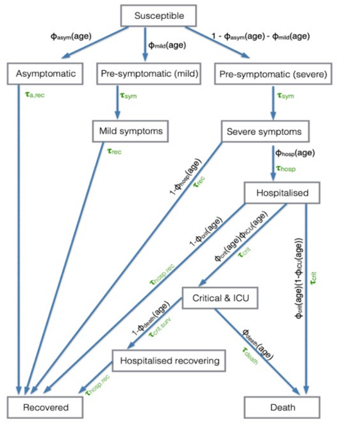
Scope of tutorial

- Preliminaries
  - Background to automatic differentiation
  - Implement a differentiable ABM
- Algorithms
  - Techniques to calibrate and analyze differentiable ABMs
- Applications
  - Real-world case study in New Zealand
- Systems
  - Tooling to build and calibrate differentiable ABMs at scale

Critical Challenges
Rapid Action
Effective Policies

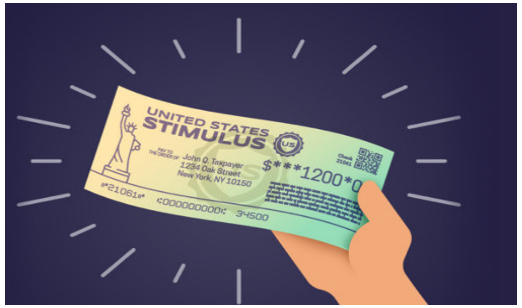
Dynamics and Interventions

New Transmission



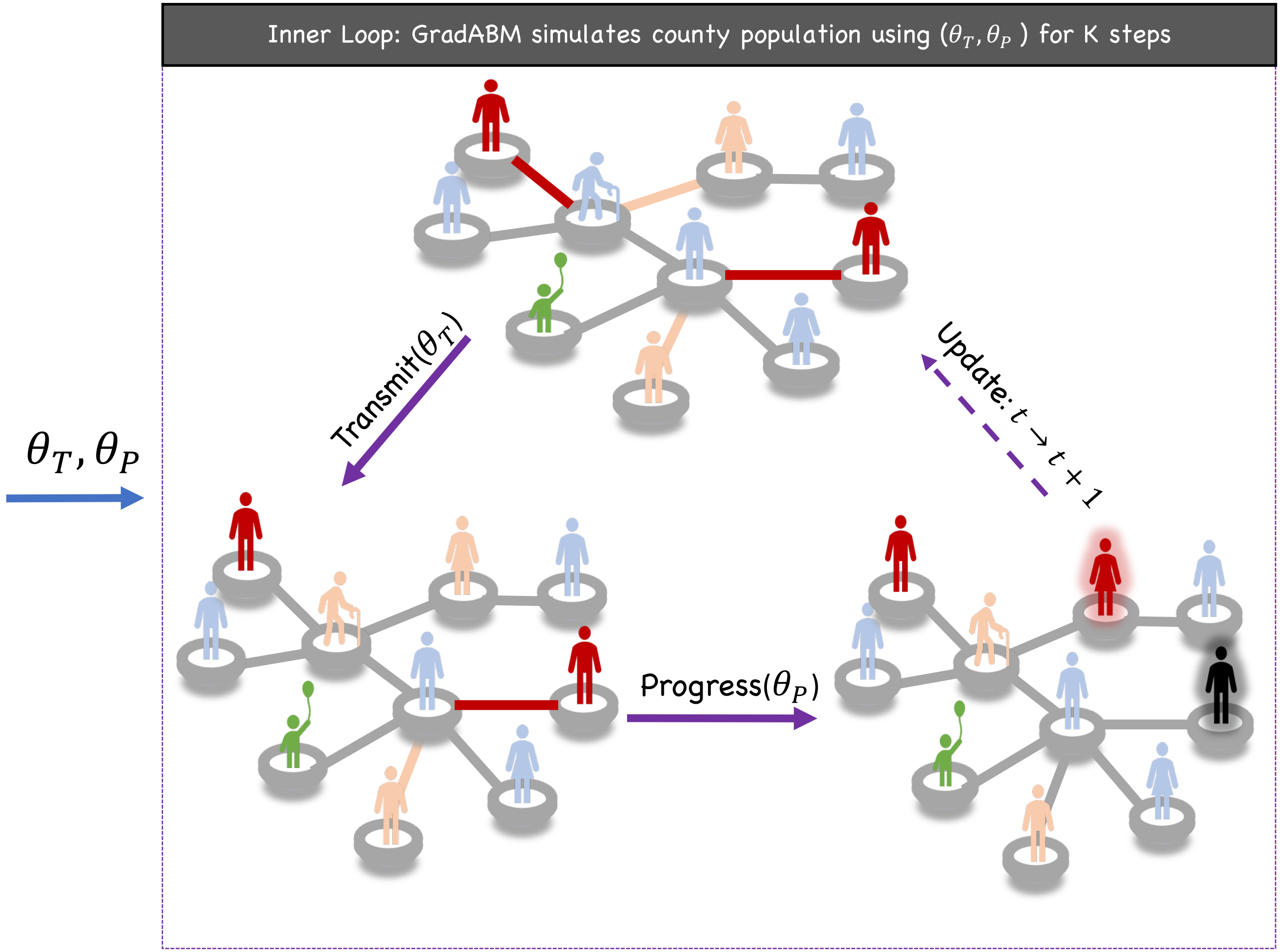
Disease Progression

Health Interventions
(Testing, Vaccination, Lockdowns)

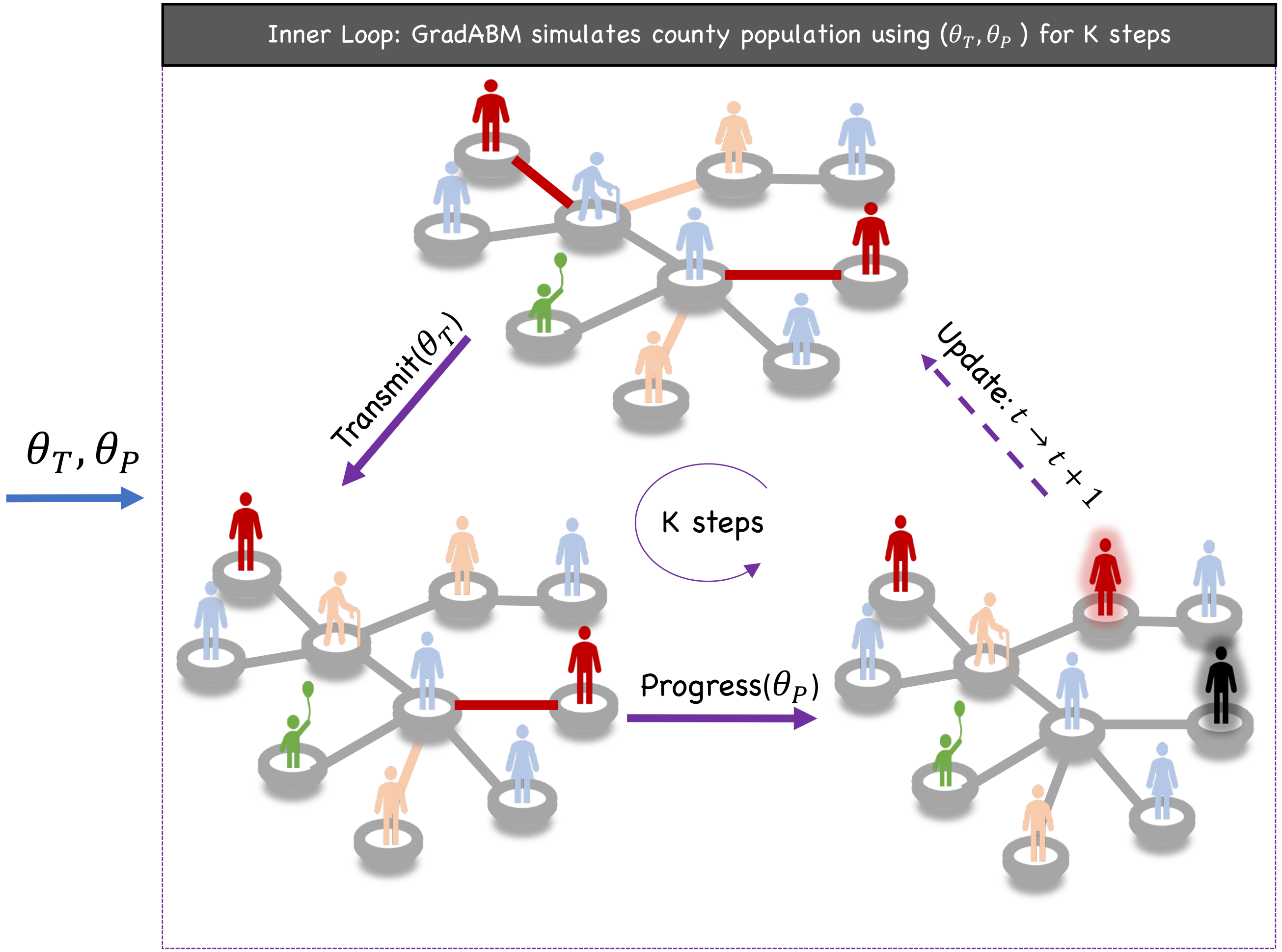


Financial Interventions
(Stimulus, PUA, PPP, FPUC)

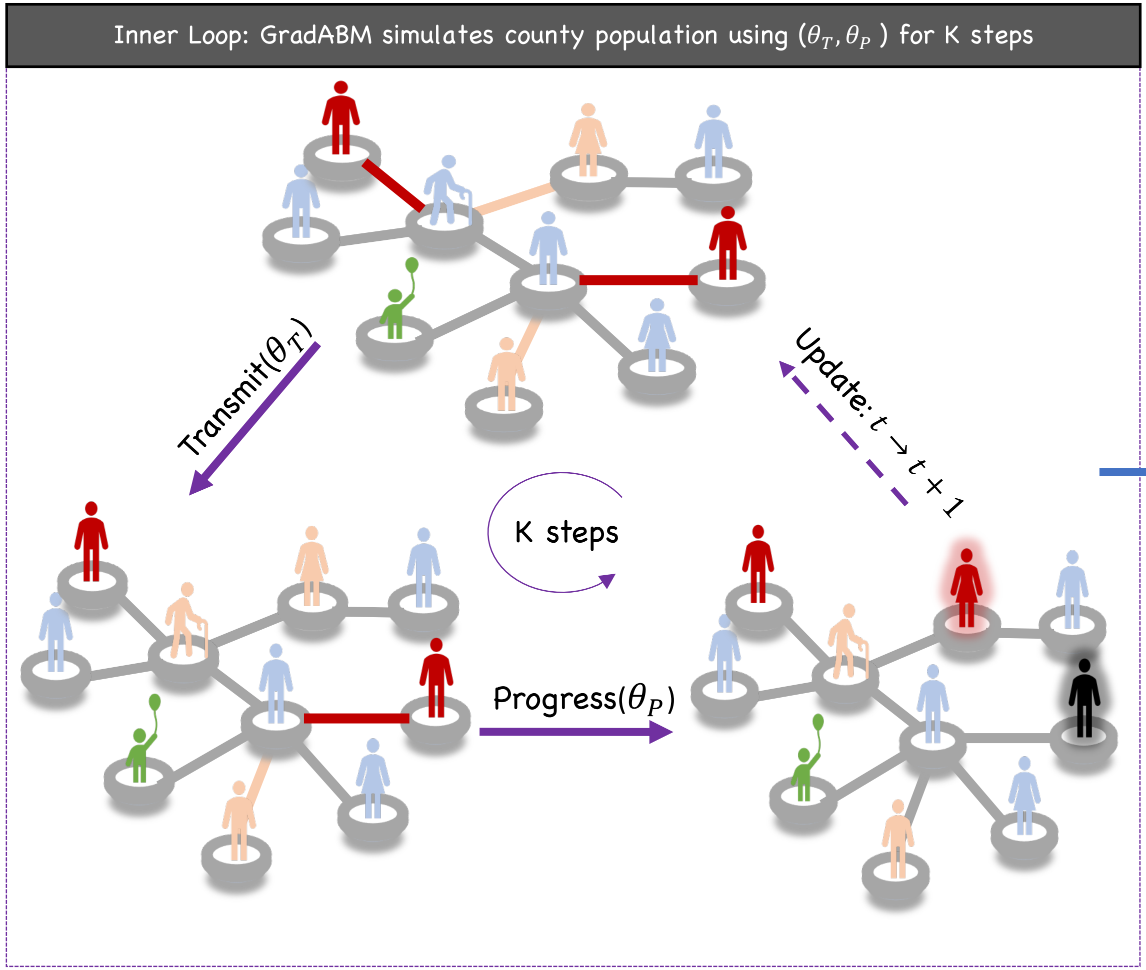
Gradient-assisted calibration

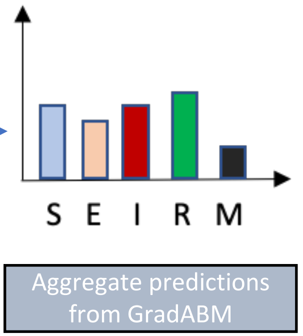


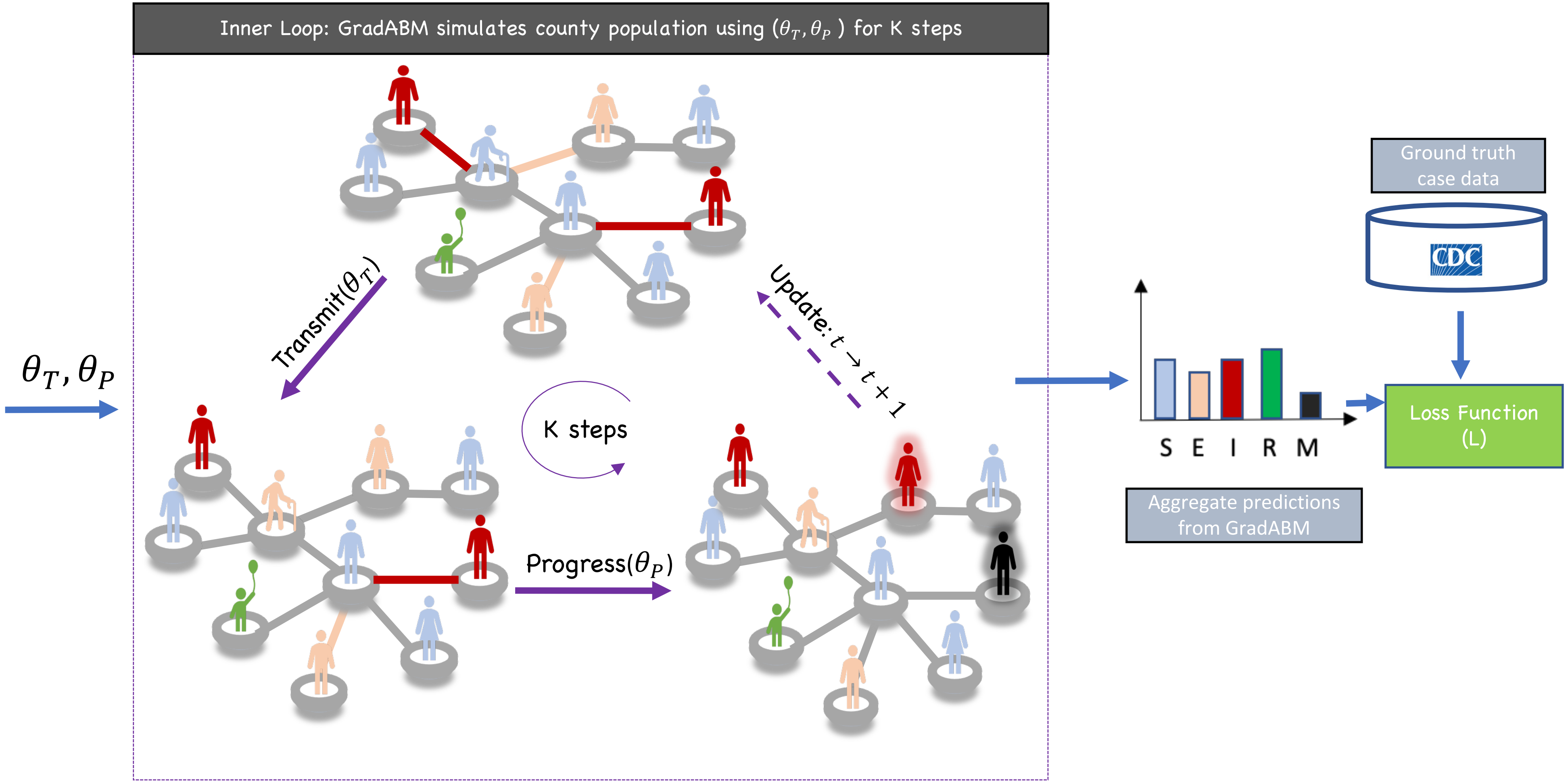
bit.ly/diff-abms


Inner Loop: GradABM simulates county population using (θ_T, θ_P) for K steps
θ_T, θ_P
Transmit(θ_T)
K steps
Update: $t \to t+1$
Progress(θ_P)

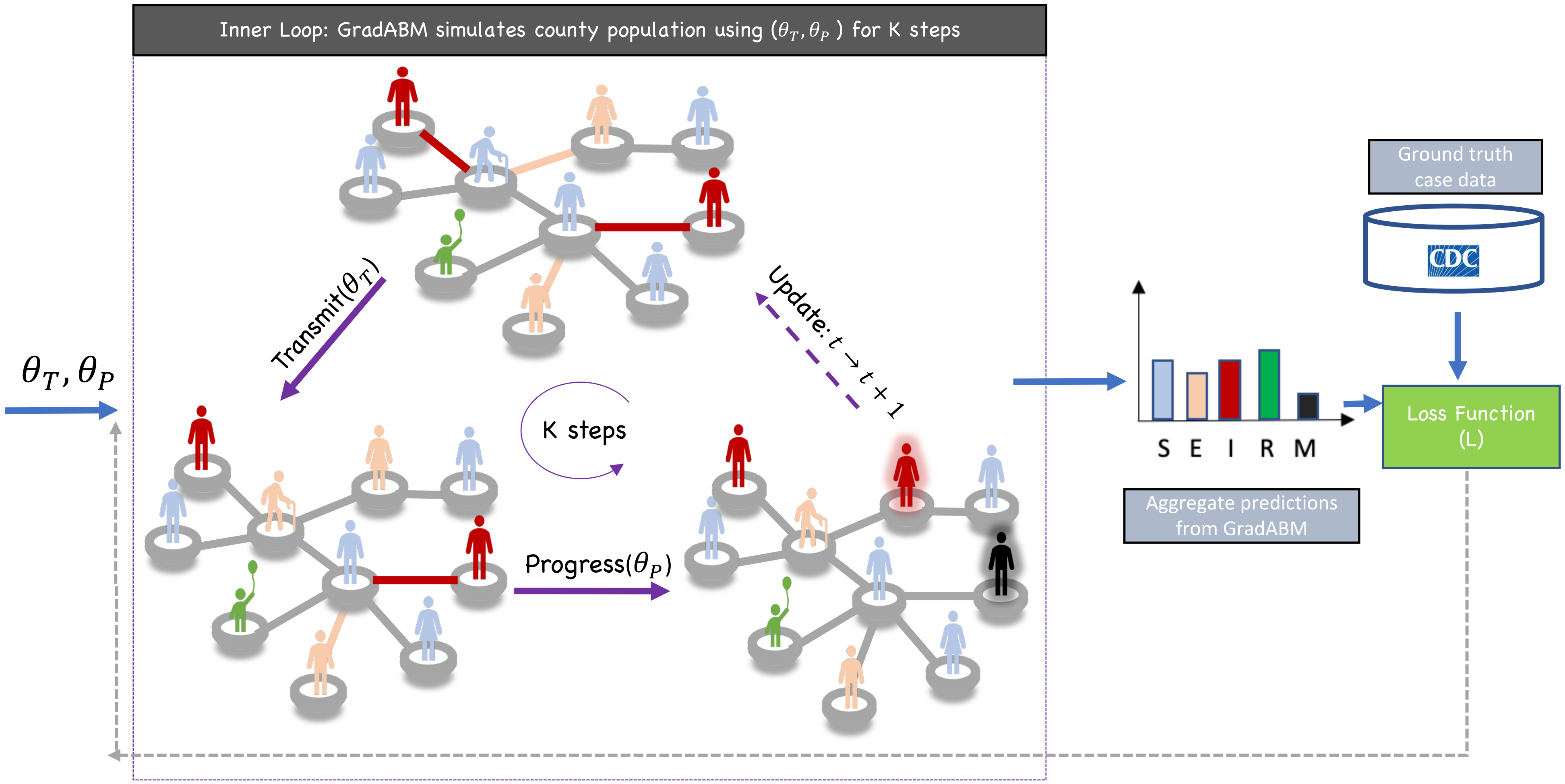
θ_T, θ_P





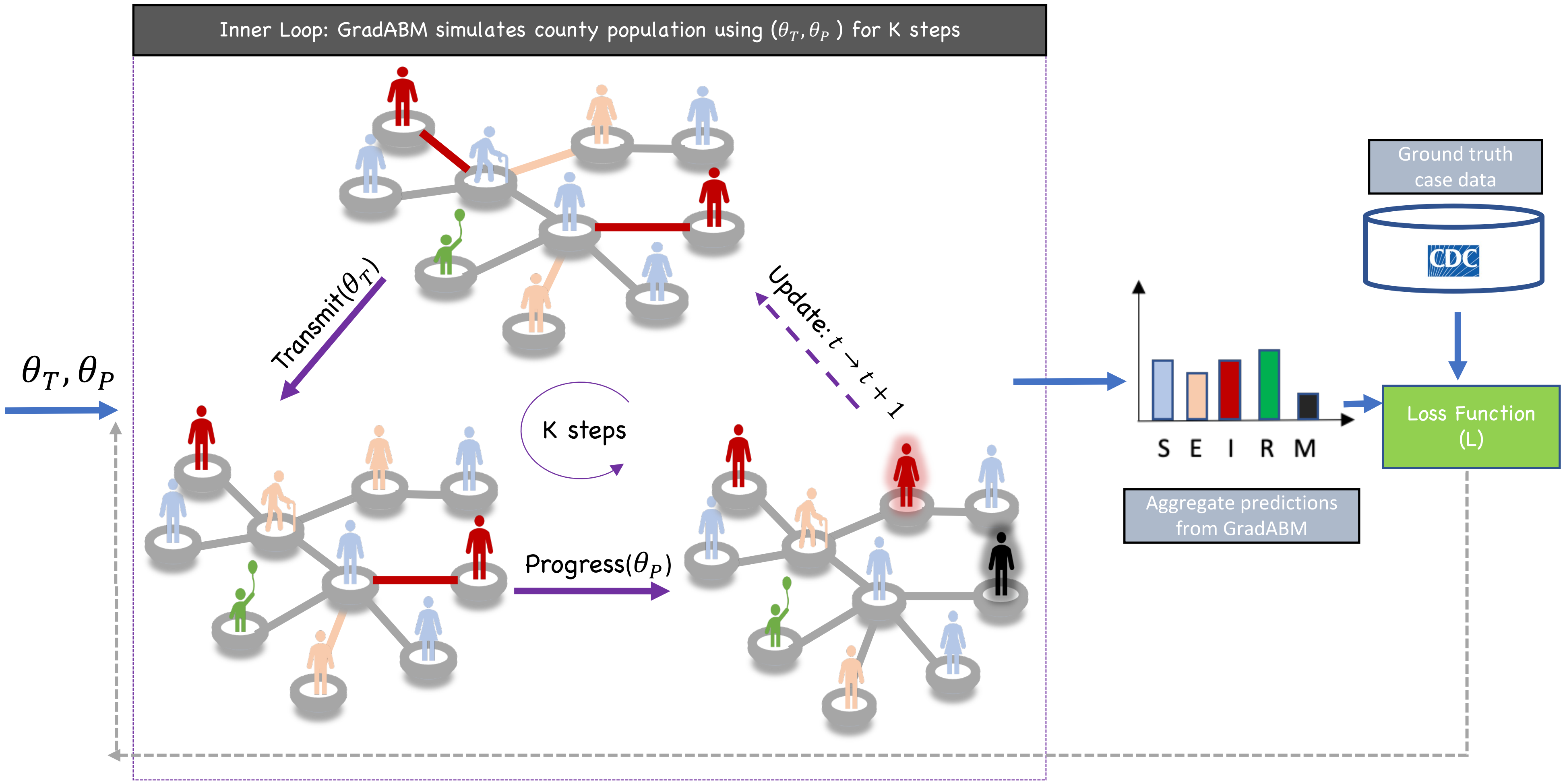


bit.ly/diff-abms



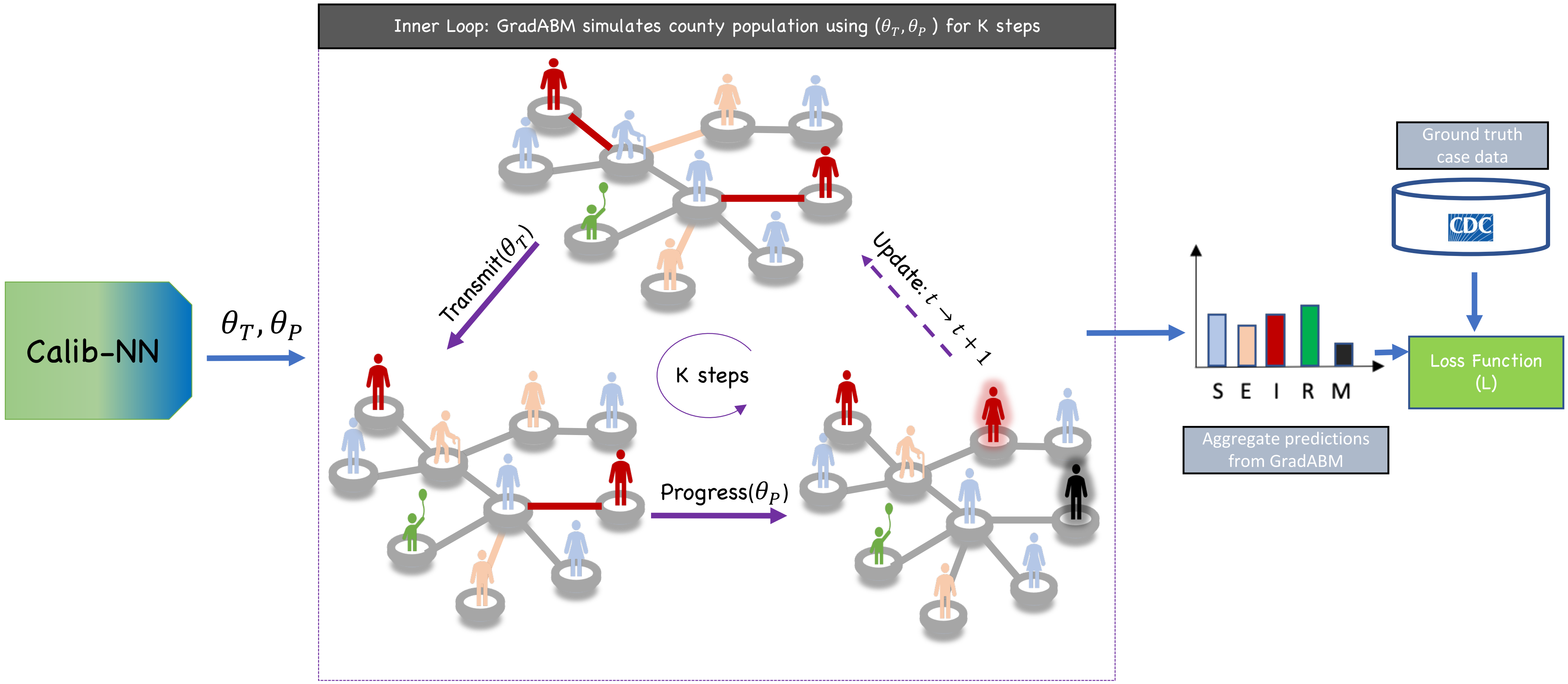
$$\theta_T = \theta_T - \alpha \frac{\partial \mathcal{L}(\hat{y}, y)}{\partial \theta_T} \qquad \theta_P = \theta_P - \alpha \frac{\partial \mathcal{L}(\hat{y}, y)}{\partial \theta_P}$$

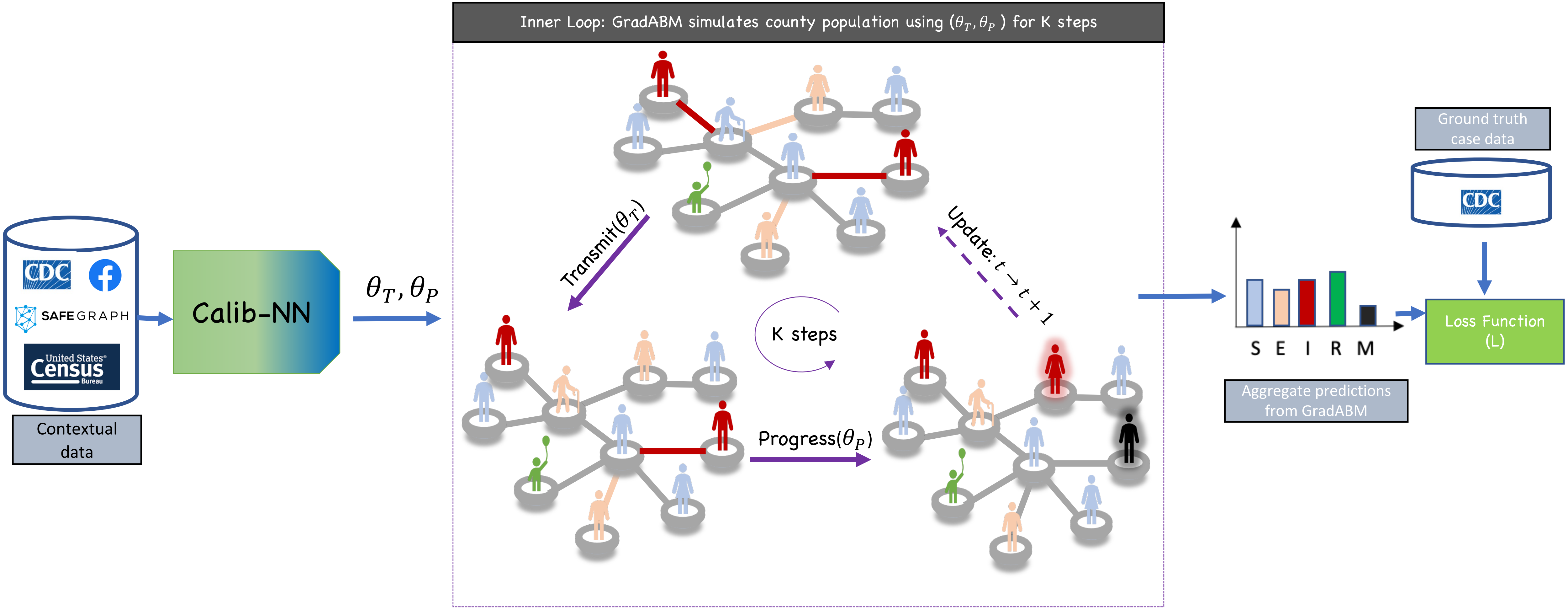
bit.ly/diff-abms



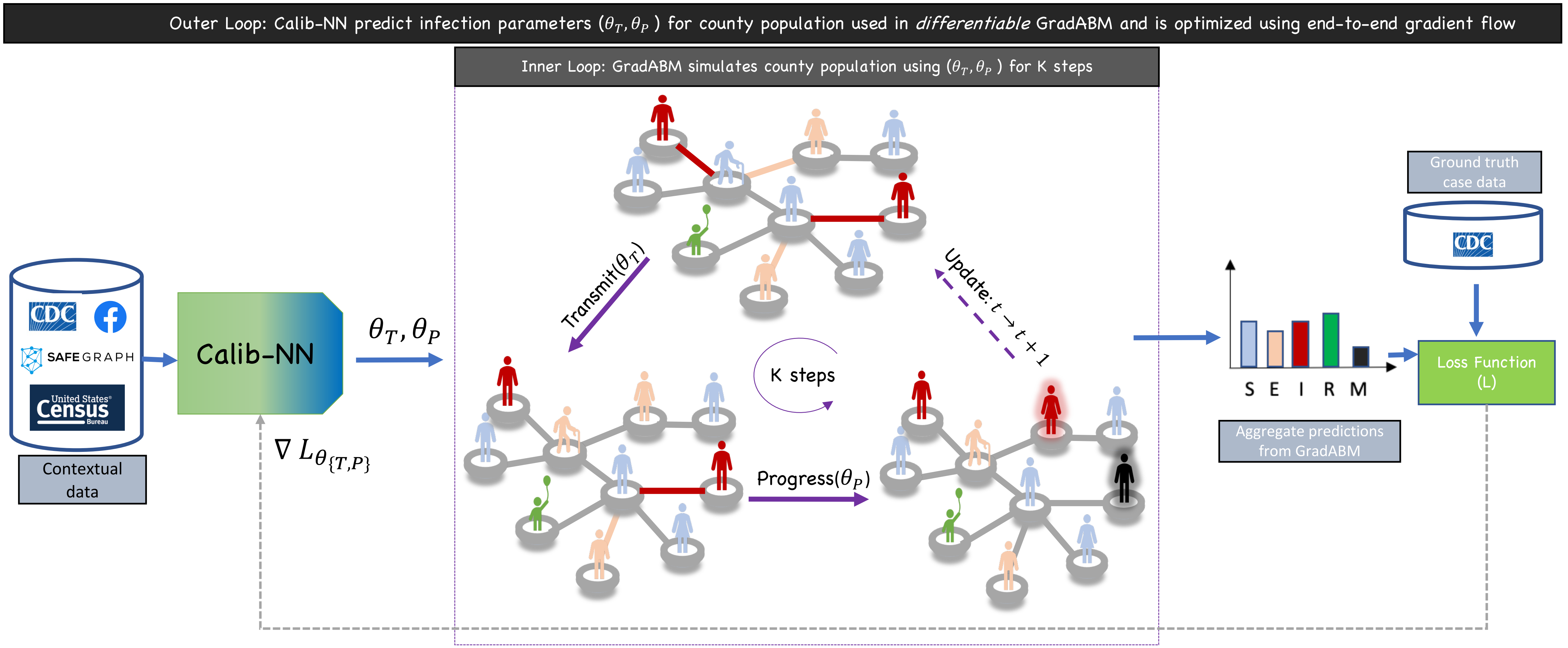
$$\theta_T = \theta_T - \alpha \frac{\partial \mathcal{L}(\hat{y}, y)}{\partial \theta_T} \qquad \theta_P = \theta_P - \alpha \frac{\partial \mathcal{L}(\hat{y}, y)}{\partial \theta_P}$$

Mode 1: Calibrate **parameters** with gradient descent (c-GRADABM)


Calib-NN
θ_T, θ_P
Inner Loop: GradABM simulates county population using (θ_T, θ_P) for K steps
Transmit(θ_T)
K steps
Progress(θ_P)
Update: $t \rightarrow t+1$
S E I R M
Aggregate predictions from GradABM
Ground truth case data
CDC
Loss Function (L)

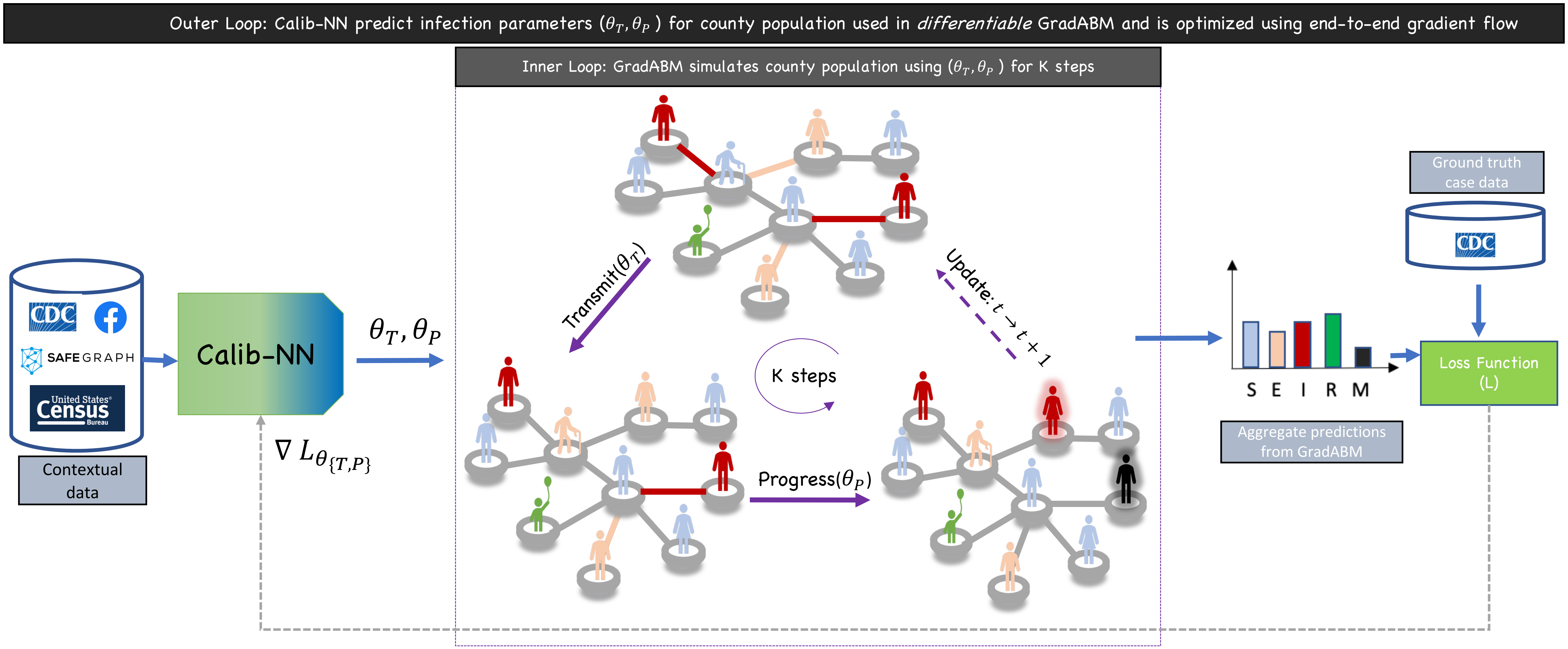


bit.ly/diff-abms



$$\phi = \phi - \alpha \frac{\partial \mathcal{L}(\hat{y}, y; (\theta_T^t, \theta_P^t))}{\partial \phi},$$

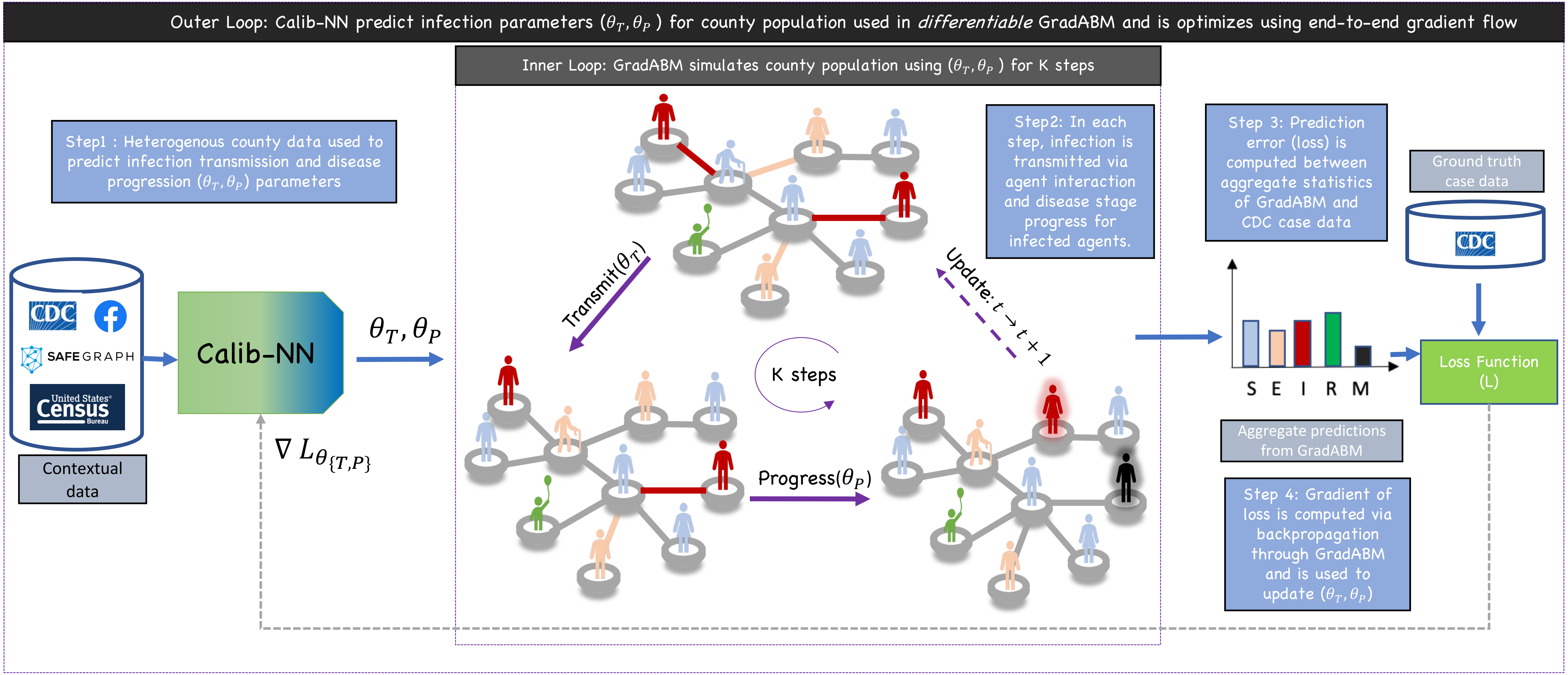
bit.ly/diff-abms



<u>Mode 2</u>: Calibrate ge**nerator function** with gradient descent (dc-GRADABM)

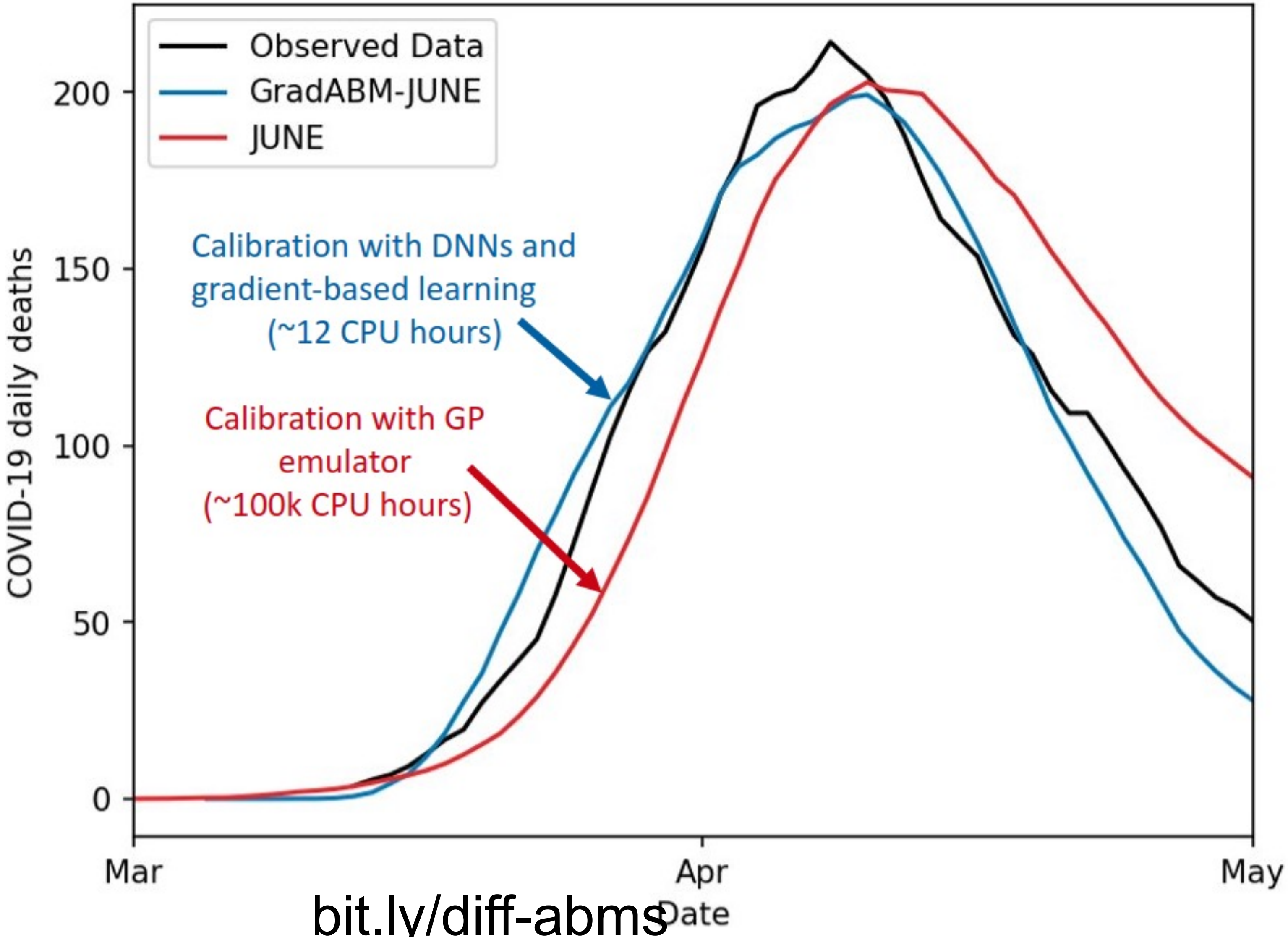
$$\phi = \phi - \alpha \frac{\partial \mathcal{L}(\hat{y}, y; (\theta_T^t, \theta_P^t))}{\partial \phi},$$

bit.ly/diff-abms



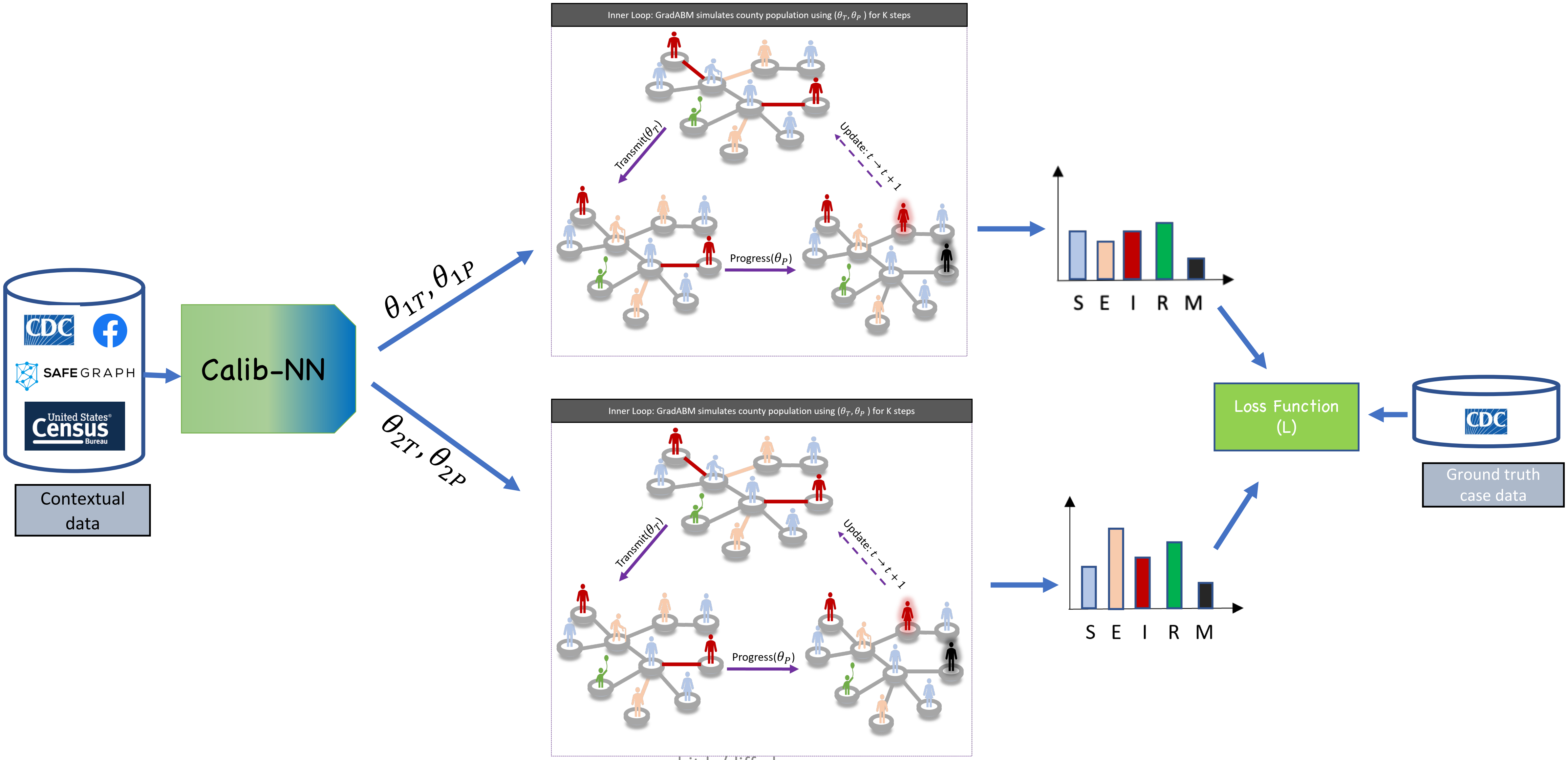
bit.ly/diff-abms

Gradients enable fast calibration over emulators: 100k to 12 CPU hours



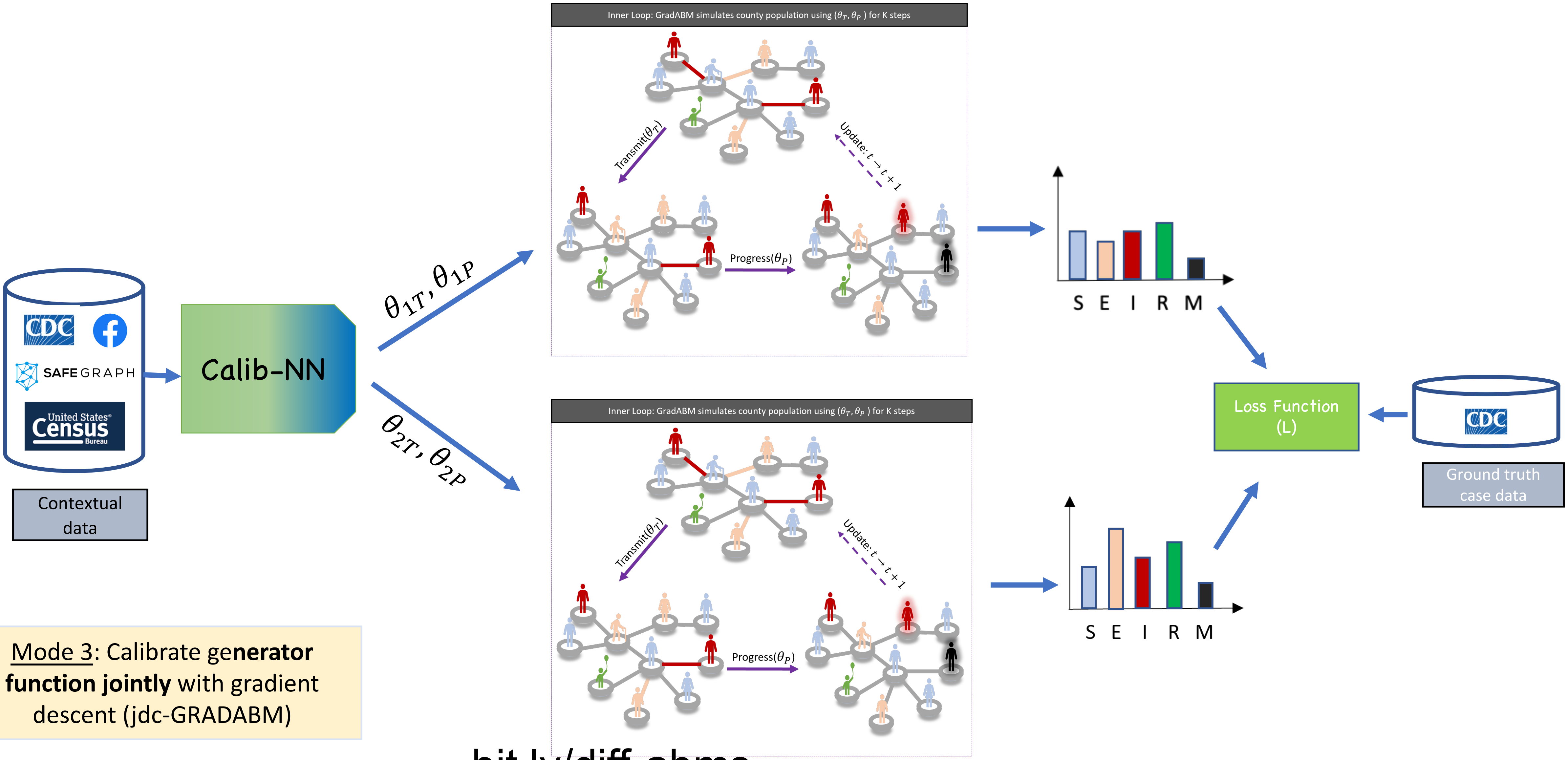
bit.ly/diff-abms

Calibrate with ensemble learning to reduce overfitting



bit.ly/diff-abms

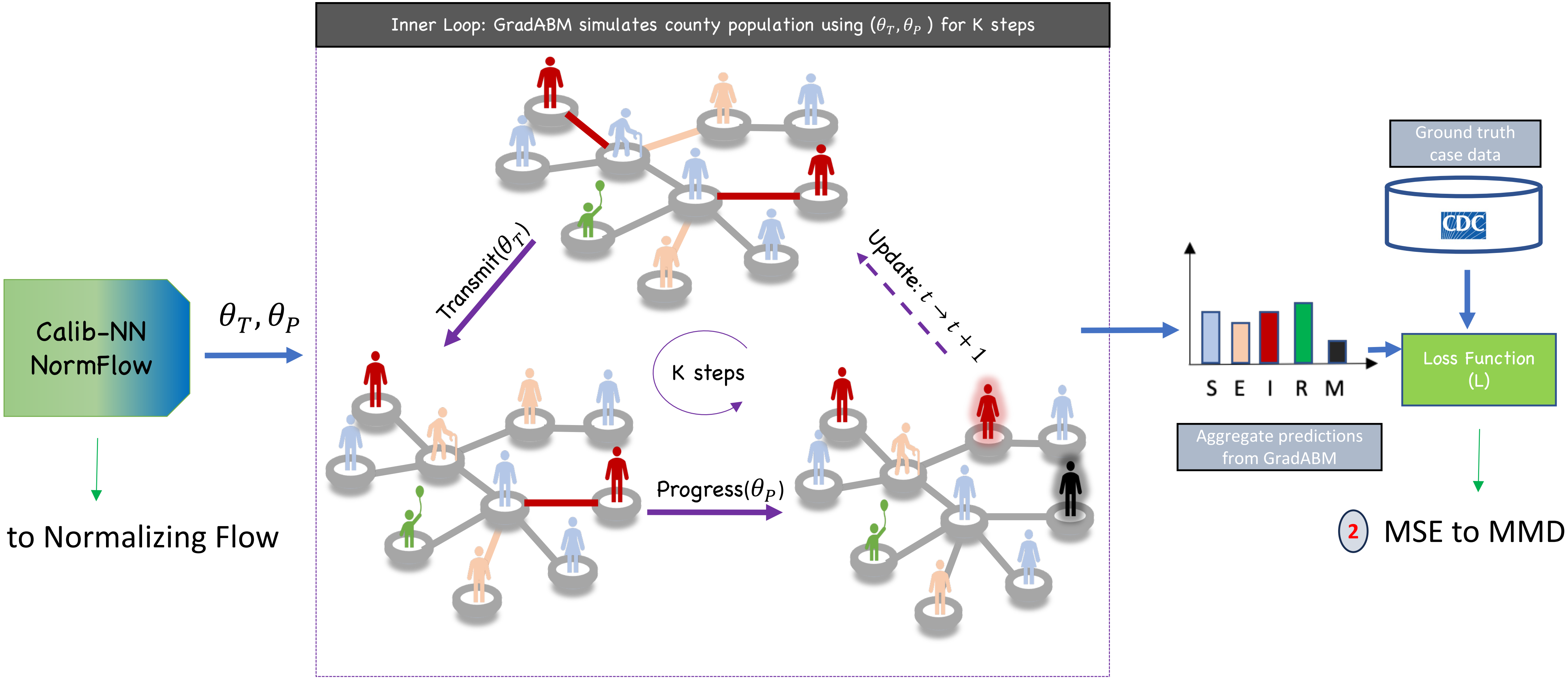
Calibrate with ensemble learning to reduce overfitting



Mode 3: Calibrate **generator function jointly** with gradient descent (jdc-GRADABM)

bit.ly/diff-abms

Calibrate posteriors with variational inference



1 RNN to Normalizing Flow

2 MSE to MMD

Mode 4: Calibrate with uncertainty quantification (dc-GRADABM)

bit.ly/diff-abms

Gradient-assisted sensitivity analysis

Sensitivity Analysis is critical for validation

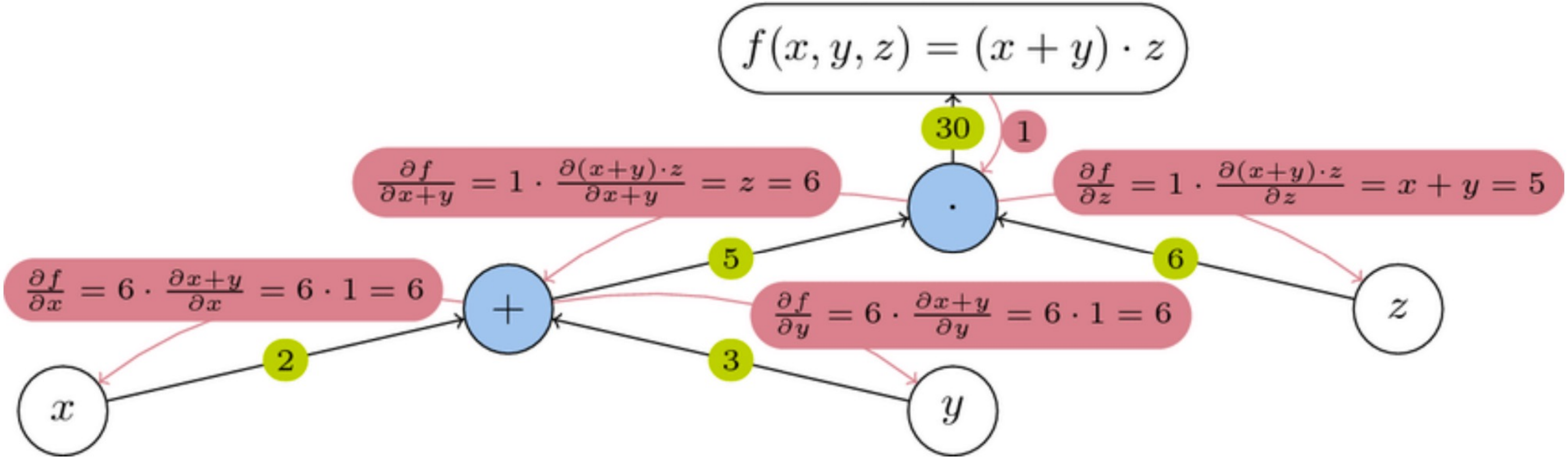
The impact of uncertainty on predictions of the CovidSim epidemiological code

Wouter Edeling[1], Hamid Arabnejad[2], Robbie Sinclair[3], Diana Suleimenova[2], Krishnakumar Gopalakrishnan[3], Bartosz Bosak[4], Derek Groen[2], Imran Mahmood[2], Daan Crommelin[1,5] and Peter V. Coveney[3,6]

Epidemiological modelling has assisted in identifying interventions that reduce the impact of COVID-19. The UK government relied, in part, on the CovidSim model to guide its policy to contain the rapid spread of the COVID-19 pandemic during March and April 2020; however, CovidSim contains several sources of uncertainty that affect the quality of its predictions: parametric uncertainty, model structure uncertainty and scenario uncertainty. Here we report on parametric sensitivity analysis and uncertainty quantification of the code. From the 940 parameters used as input into CovidSim, we find a subset of 19 to which the code output is most sensitive—imperfect knowledge of these inputs is magnified in the outputs by up to 300%. The model displays substantial bias with respect to observed data, failing to describe validation data well. Quantifying parametric input uncertainty is therefore not sufficient: the effect of model structure and scenario uncertainty must also be properly understood.

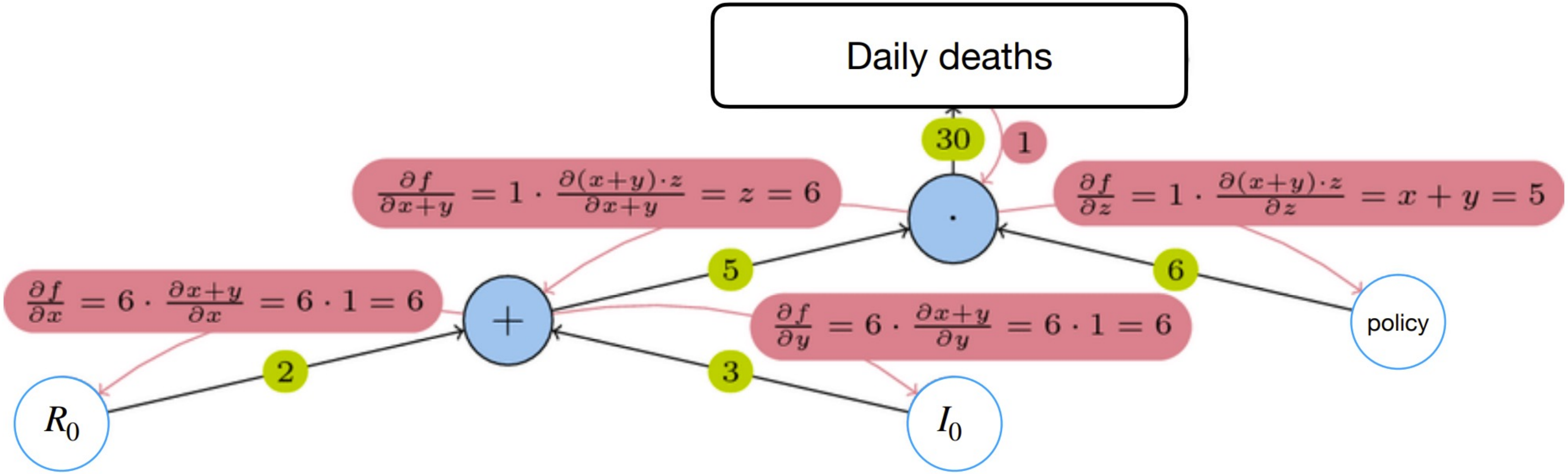
Ensemble execution. Consequently, through the use of adaptive methods we make the uncertainty analysis of CovidSim tractable, but our analysis nevertheless required us to perform thousands of runs, each with its own unique set of input parameters. Specifically, we used the Eagle supercomputer at the Posnan

Recap: Reverse-mode automatic differentiation



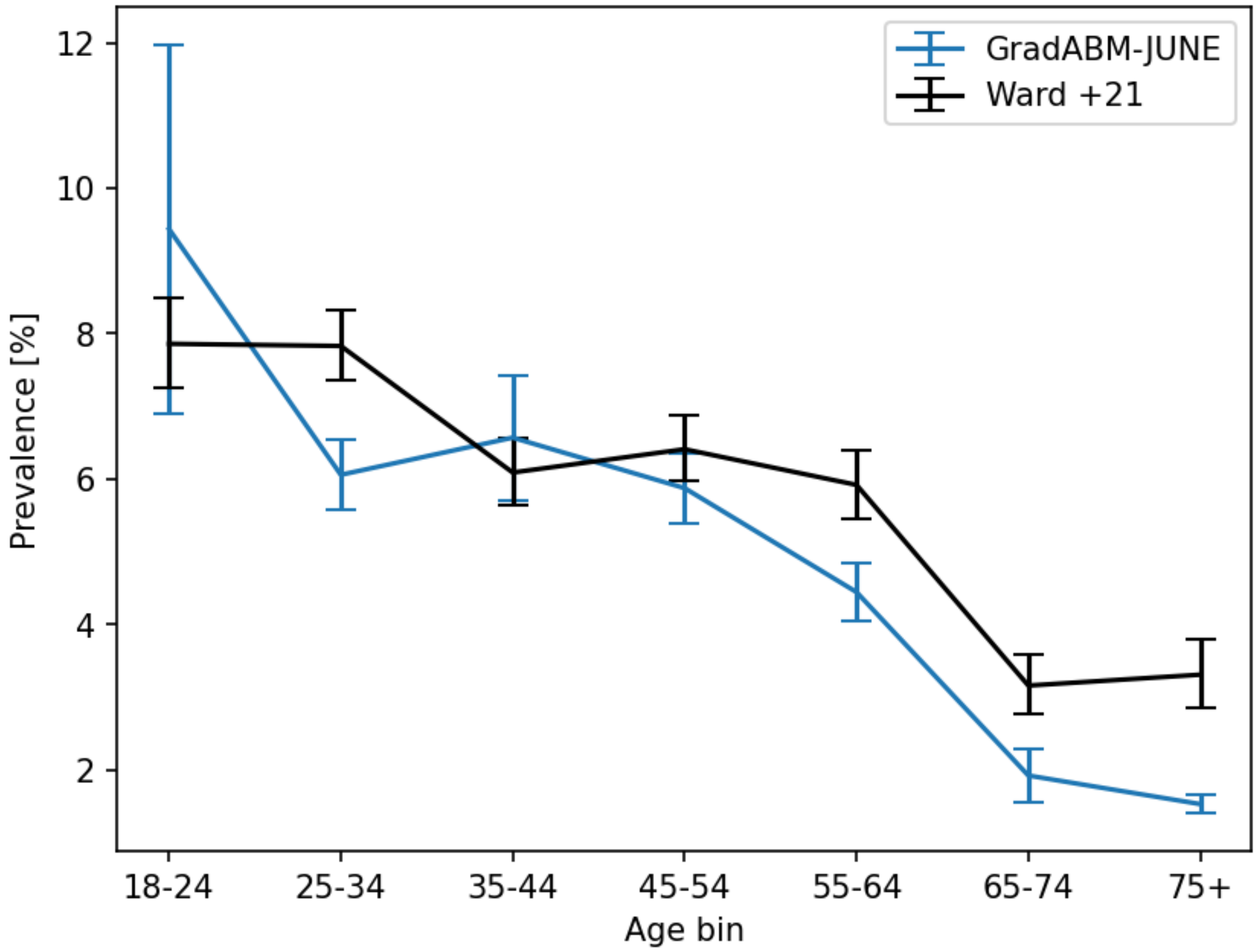
Sensitivity analysis via reverse-mode automatic differentiation

Reverse-mode automatic differentiation is independent of the number of parameters!!



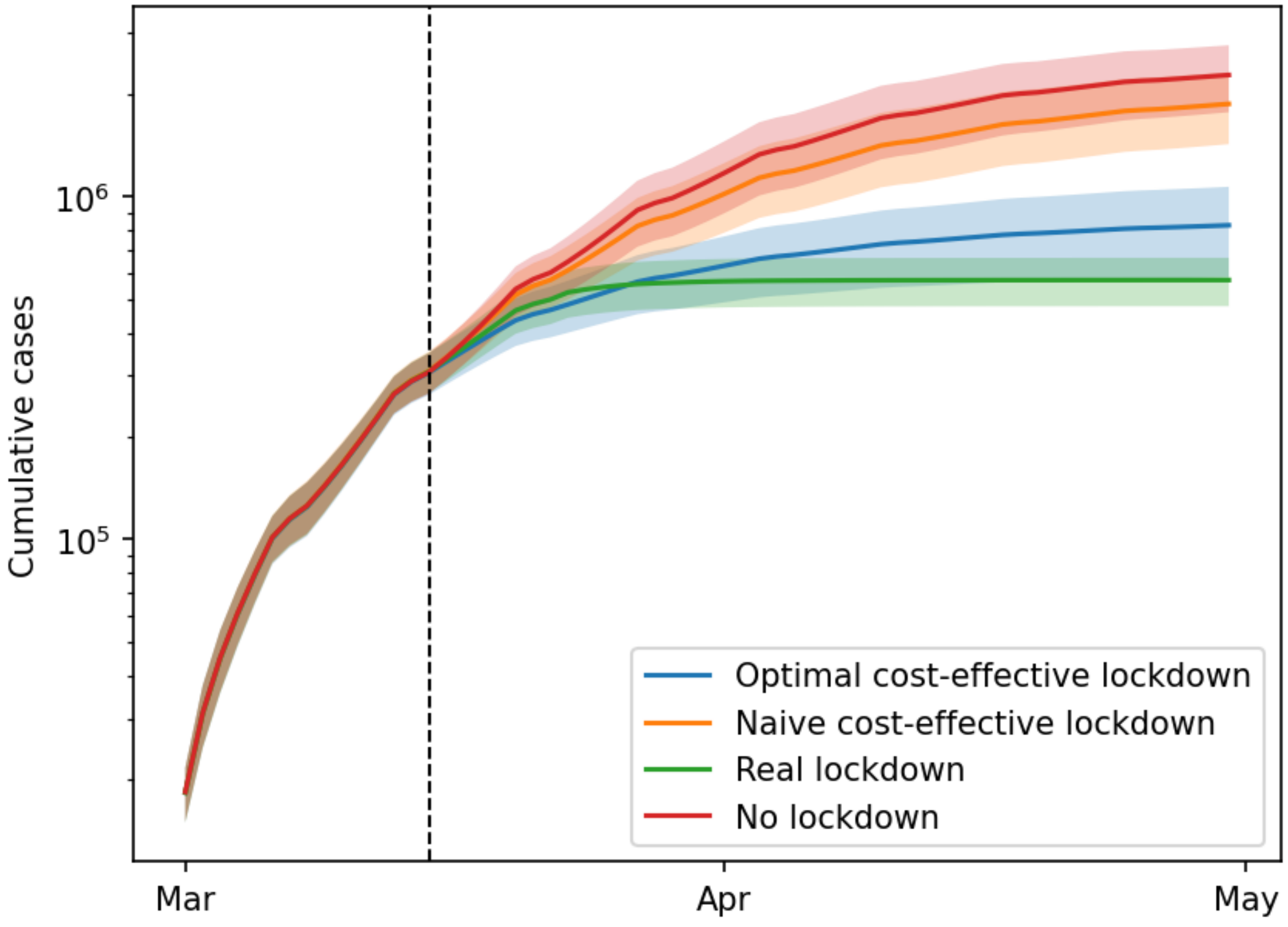
How effective *really* were lockdown policies?

Analyze retrospective decisions by reproducing seroprevalence studies in-silico



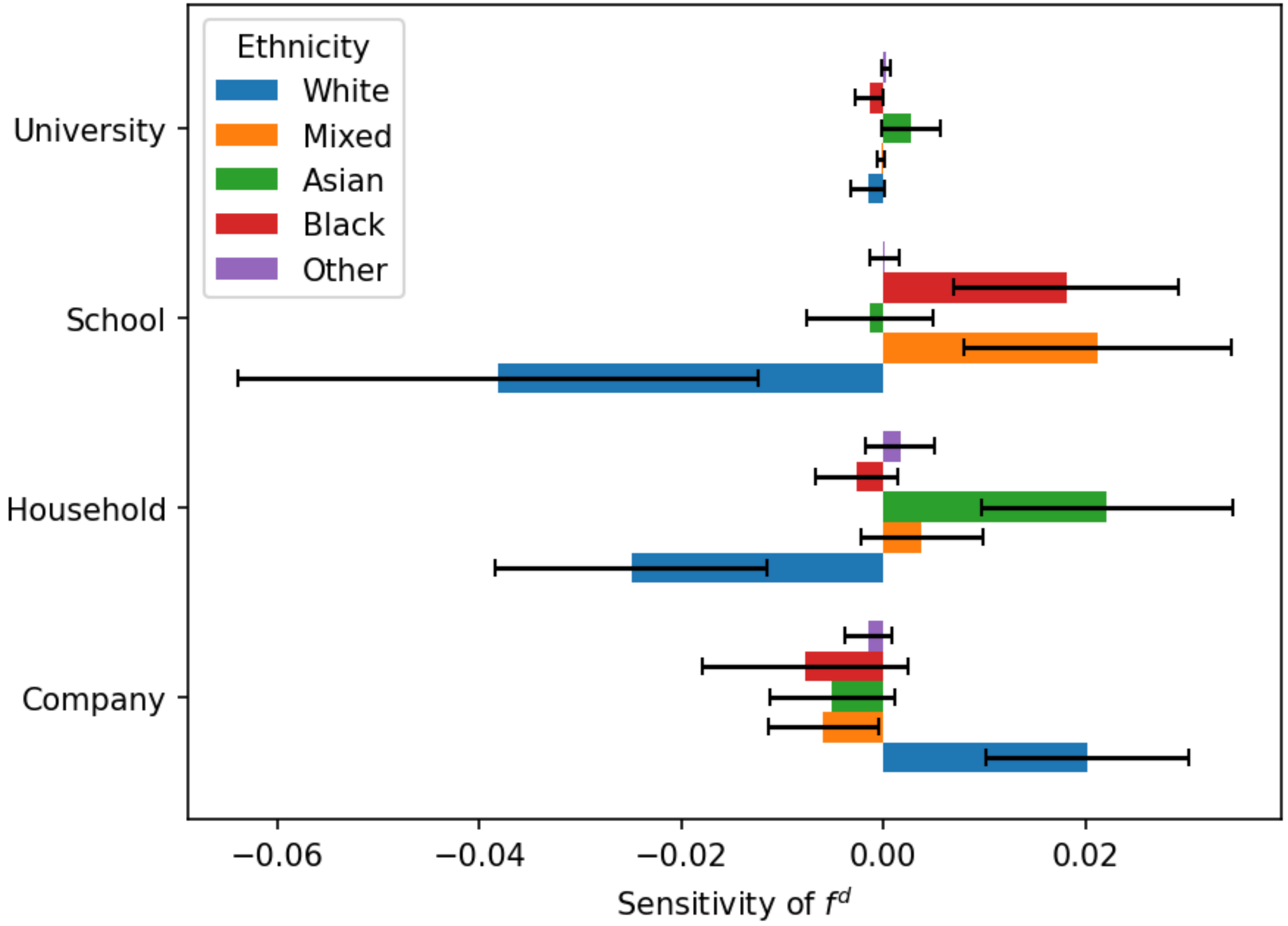
What could we have done differently?

Design counterfactual lockdown policies with multiple constraints in-silico!



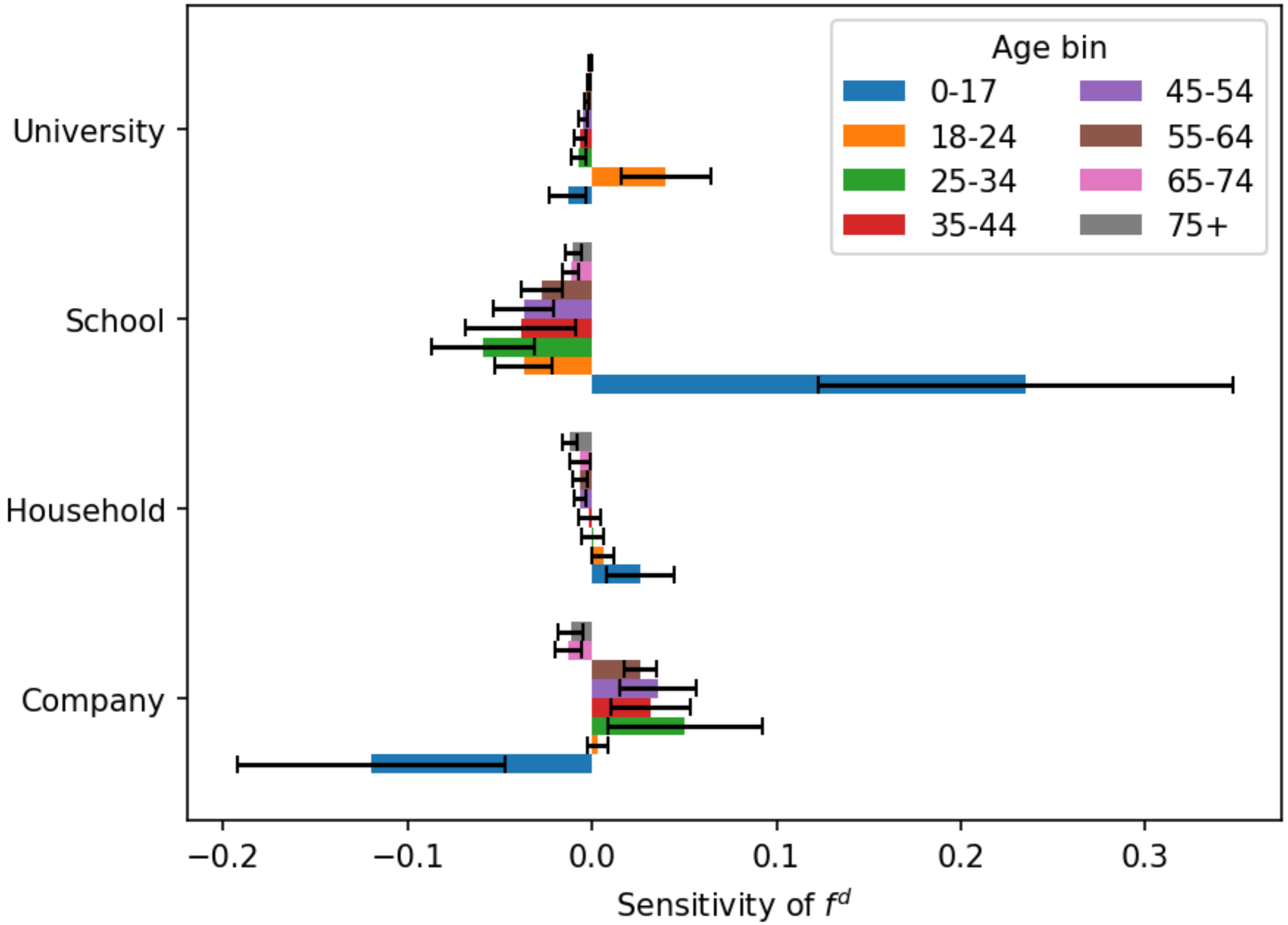
How sensitive was infection to ethnicity?

More infection among South Asians through households in contrast to white British people



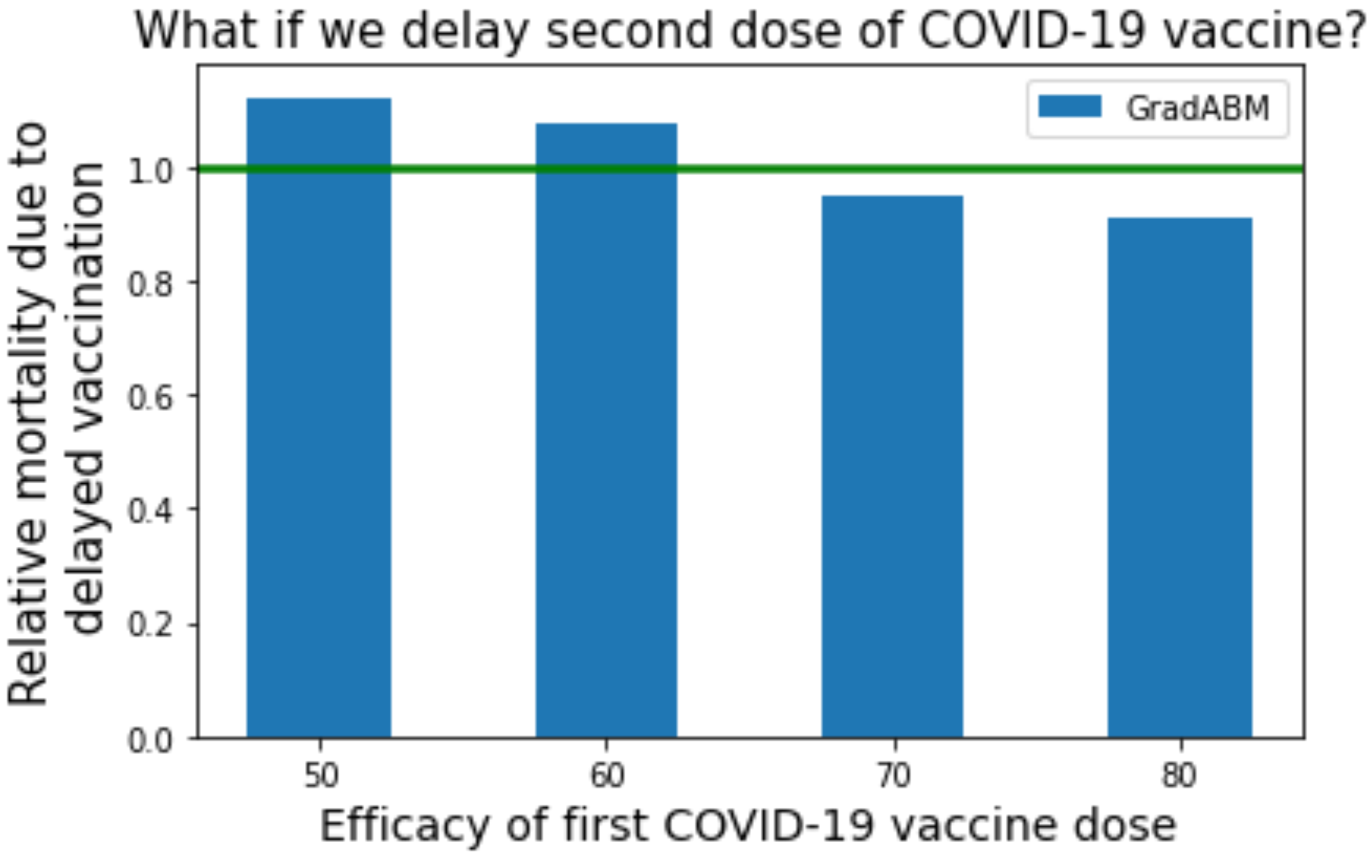
How sensitive was infection to age?

Dominant infection spread through schools for 0-17 and university for 18-24



What if we delay second dose of the COVID-19 vaccine?

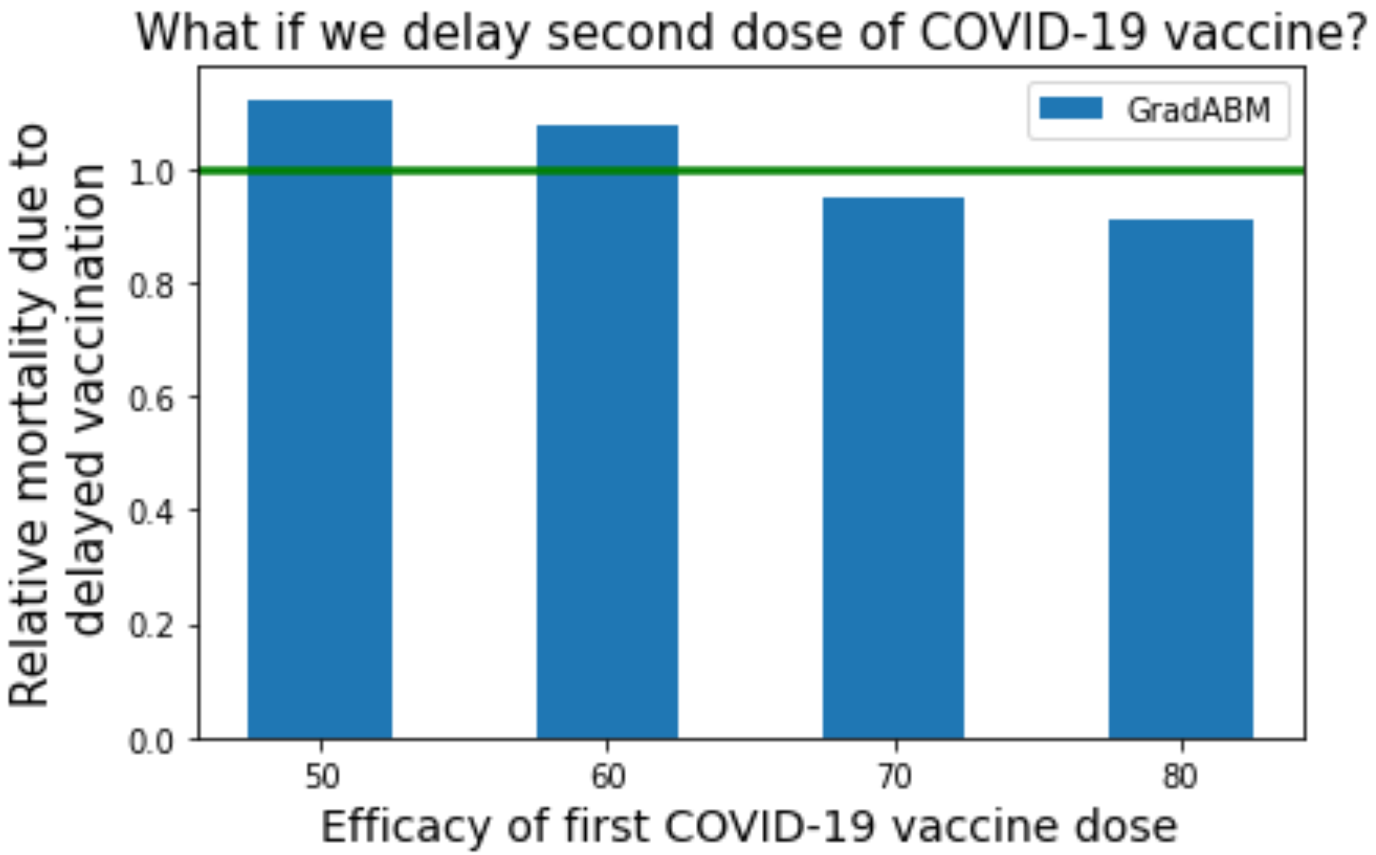
Supply chain limitations and population behavior to design immunization policies



bit.ly/diff-abms

What if we delay second dose of the COVID-19 vaccine?

Consider supply chain limitations and population behavior to design immunization policies



was higher under the 12-week strategy than the 3-week strategy. For this period, we estimated that delaying the interval between the first and second COVID-19 vaccine doses from 3 to 12 weeks averted a median (calculated as the median of the posterior sample) of 58 000 COVID-19 hospital admissions (291 000 cumulative hospitalisations [95% credible interval 275 000–319 000] under the 3-week strategy *vs* 233 000 [229 000–238 000] under the 12-week strategy) and 10 100 deaths (64 800 deaths [60 200–68 900] *vs* 54 700 [52 800–55 600]). Similarly, we estimated that the

Retrospective impact of delaying 2nd covid-19 dose in England (Lancet '23)

bit.ly/diff-abms

More details: Jade Room 3 on Friday at 10 am

- Composing with neural networks
- Using LLM as agents for million-scale simulations
- Modeling with private and distributed data
- Generating diverse simulation scenarios

Scope of tutorial

- Preliminaries
  - Background to automatic differentiation
  - Implement a differentiable ABM
- Algorithms
  - Techniques to calibrate and analyze differentiable ABMs
- Applications
  - Real-world case study in New Zealand
- Systems
  - Tooling to build and calibrate differentiable ABMs at scale

Differentiable ABMs in action: Case Study of New Zealand

Scope of tutorial

- Preliminaries
  - Background to automatic differentiation
  - Implement a differentiable ABM
- Algorithms
  - Techniques to calibrate and analyze differentiable ABMs
- Applications
  - Real-world case study in New Zealand
- Systems
  - Tooling to build and calibrate differentiable ABMs at scale

Variational Inference with Blackbirds

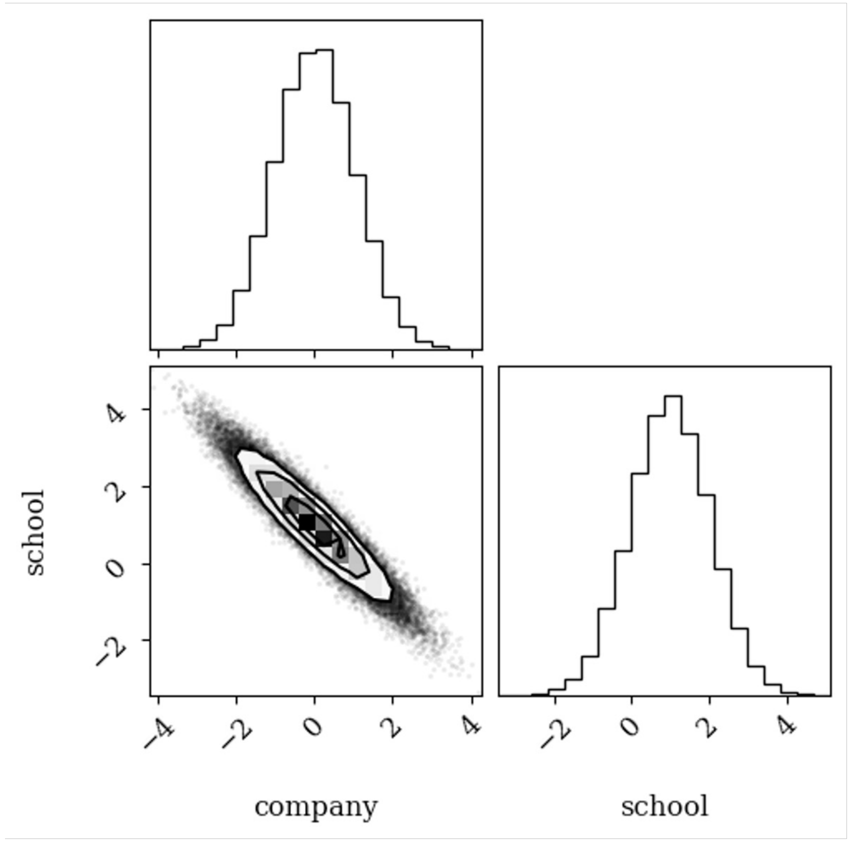
github.com/arnauqb/blackbirds

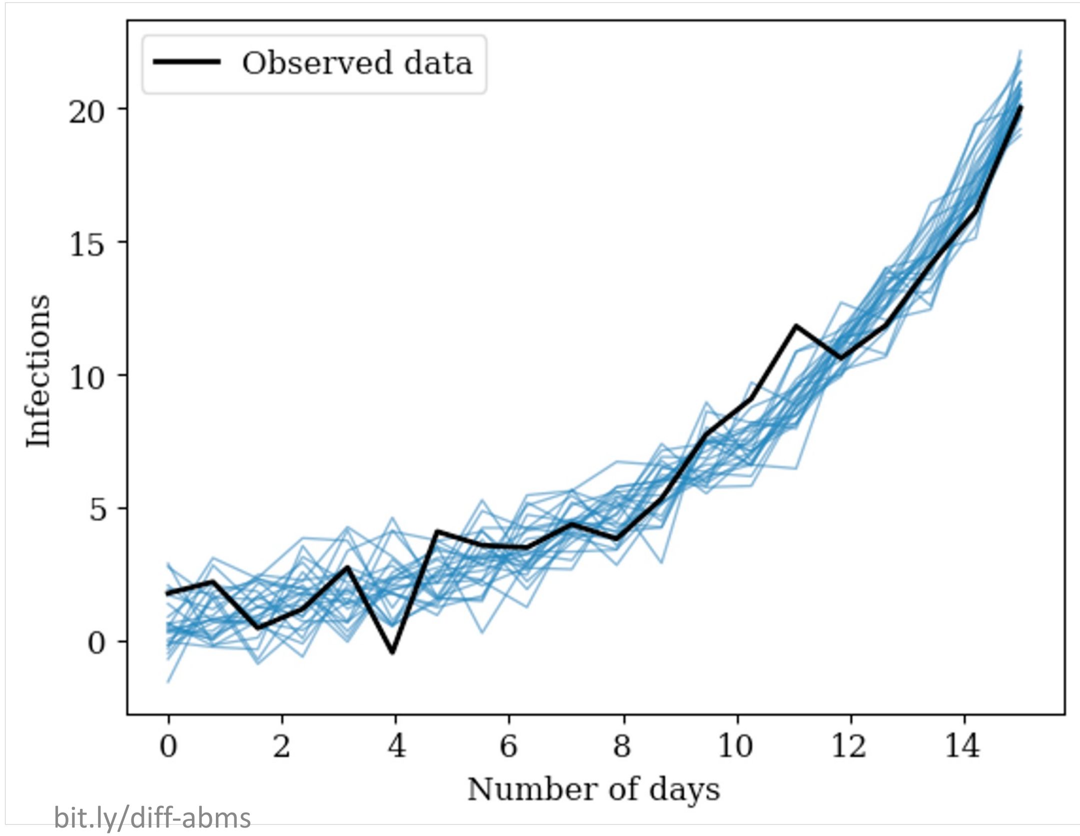
pip install blackbirds

Bayesian inference

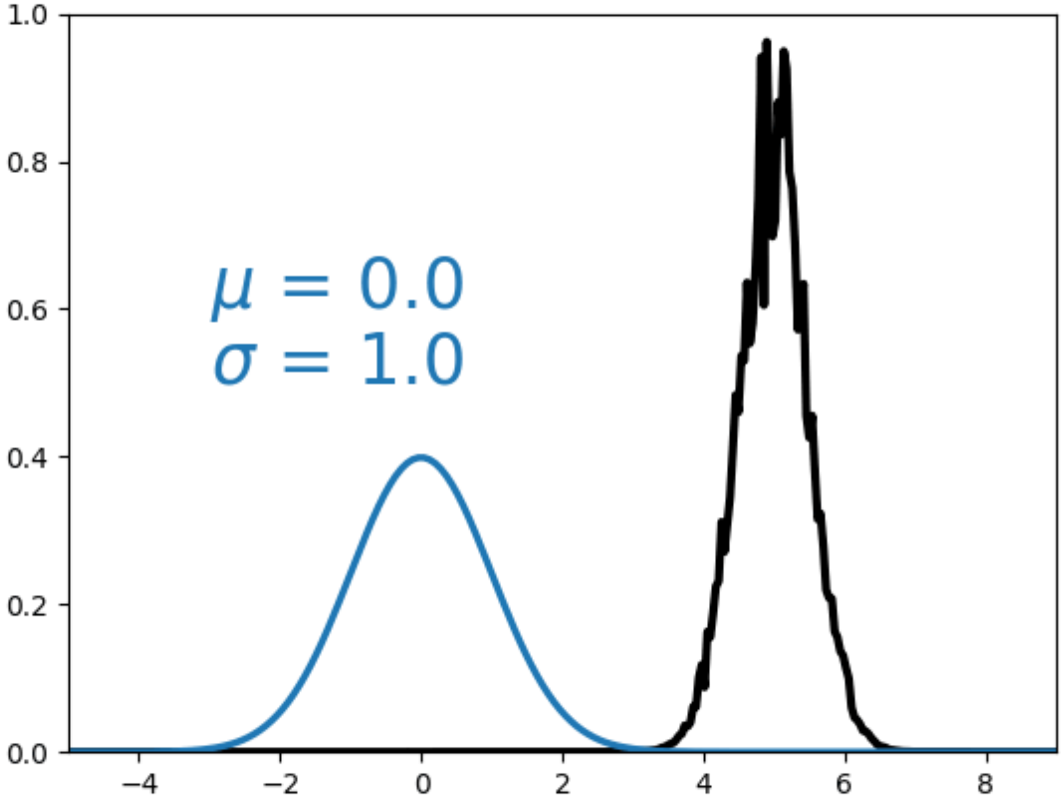
$$p(\theta|x) \propto p(x|\theta)p(\theta)$$

posterior likelihood prior



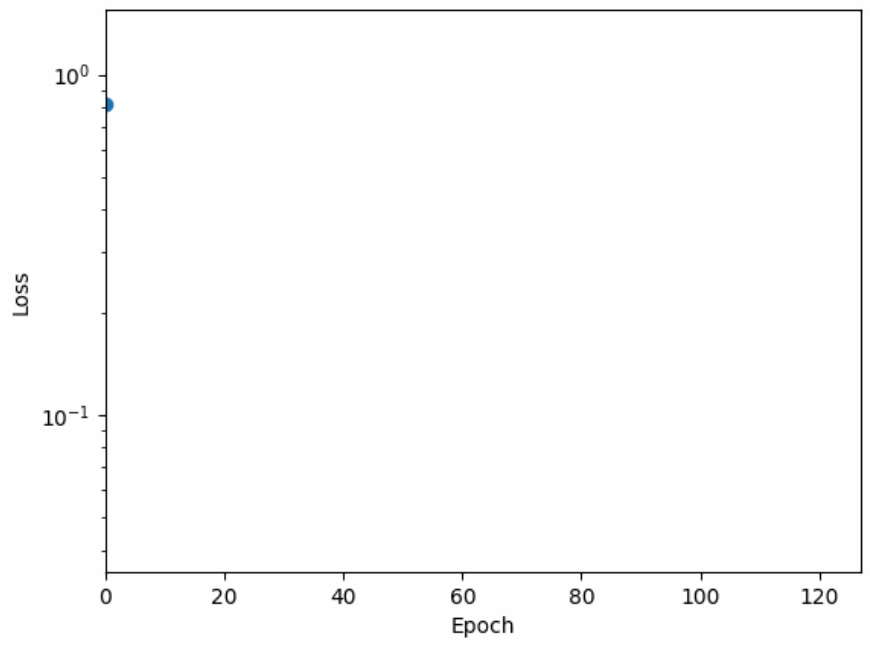


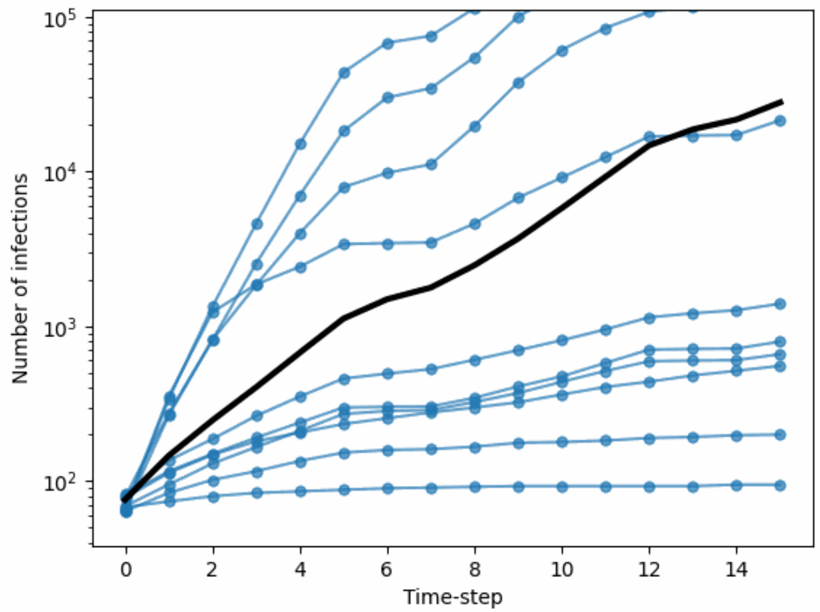
Variational Inference: Bayesian Inference as an optimization problem

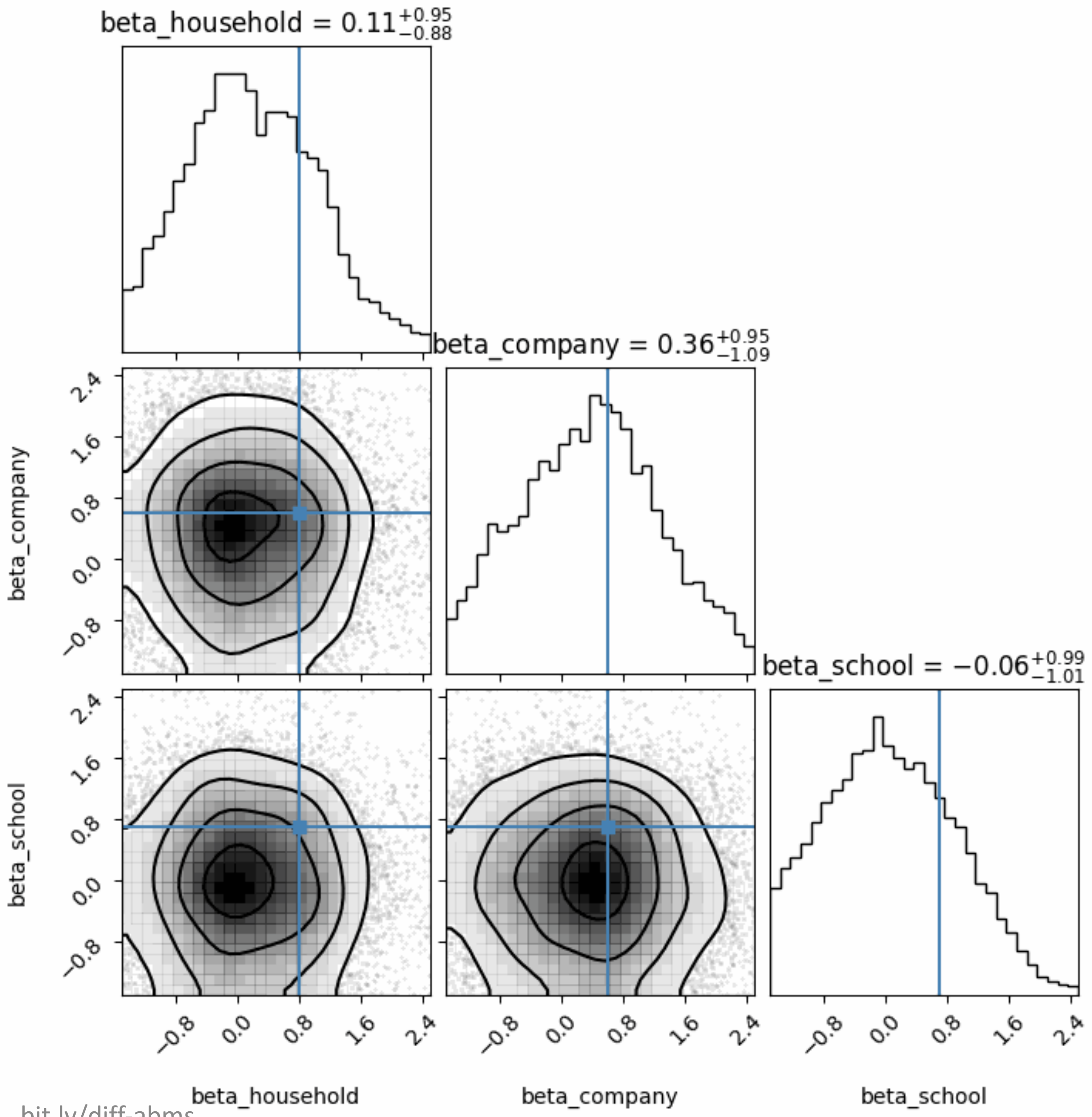


1. Assume posterior can be approximated by a family of distributions

2. Optimise for optimal parameters







bit.ly/diff-abms

Build your own Differentiable ABMs with AgentTorch

github.com/AgentTorch/AgentTorch

pip install agent-torch

Using the AgentTorch API

execute simulation
talk to your simulation
customize agents (eg: LLM as agent)
customize population (eg: NZ -> NYC)

Execute a simulation with AgentTorch

Simple Python API. Get started in 3 lines of code. Massive Acceleration. "AI Compatible"

```python
from AgentTorch.models import disease
from AgentTorch.populations import new_zealand

from AgentTorch.execute import Executor

simulation = Executor(disease, new_zealand)
simulation.execute()
```

Gradient-based learning with AgentTorch

Pytorch compatible. Optimize parameters. Compose with Neural Networks

```
from torch.optim import SGD

optimizer = SGD(simulation.parameters())
for i in range(episodes):
  optimizer.zero_grad()
  simulation.execute()
  optimizer.step()
  simulation.reset()
```

Visualize your simulation with AgentTorch

Interactive geo-plots and natural language interface

```
from AgentTorch.visualize import Visualizer
from AgentTorch.LLM.qa import load_state_trace

state_trace = load_state_trace(simulation)
visualizer = Visualizer(state_trace)

visualizer.plot('agent_behavior')
```

Talk to your AgentTorch simulation

Understand the past. "brainstorm" for the future. Verify the data. Speculate *reliably.*

```
from AgentTorch.LLM.qa import SimulationAnalysisAgent

analyzer = SimulationAnalysisAgent(simulation, state_trace)

analyzer.query("Which age group has lowest median income, how
much is it?")

analyser.query("how are stimulus payments affecting disease?")
```

Customize agents in AgentTorch

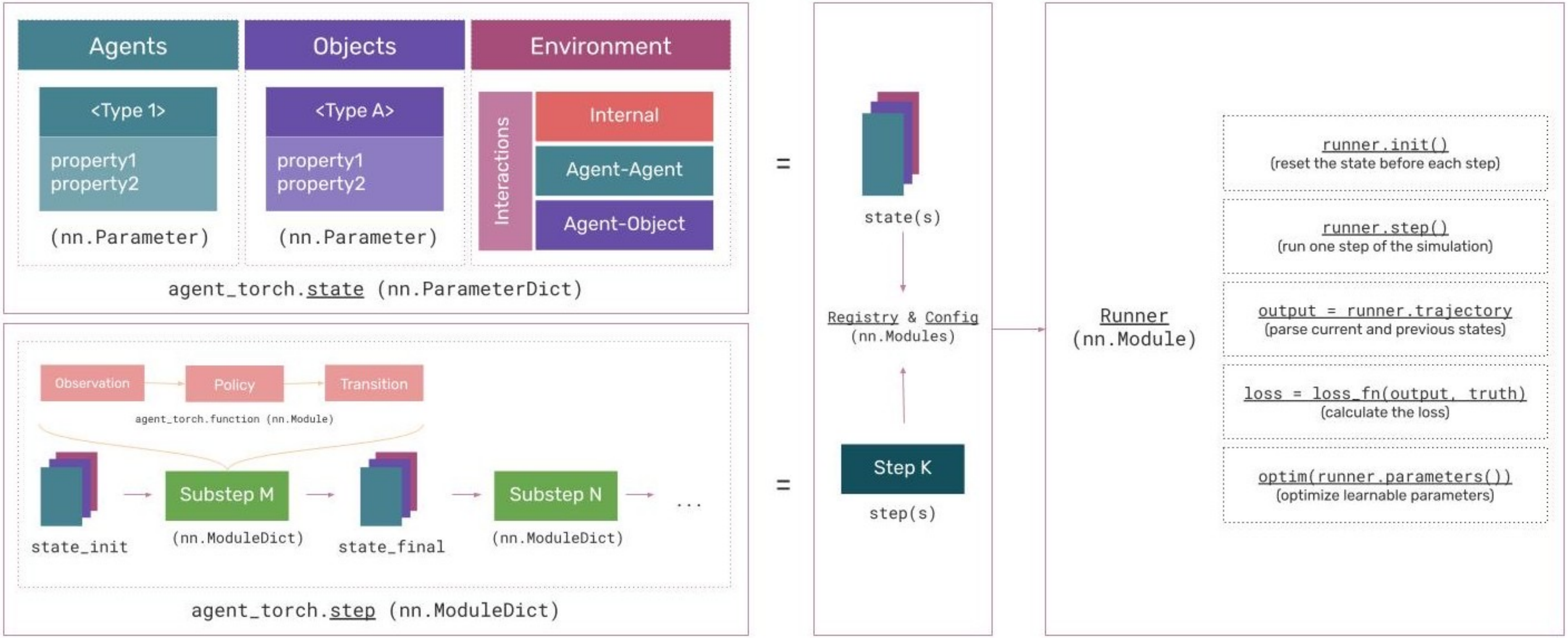
Agents can be heuristic, LLMs or neural networks

```
from AgentTorch.dataloader import DataLoader

dataloader = DataLoader(new_zealand)
dataloader._set_config_attribute('use_llm_agent', True)
dataloader._set_config_attribute('prompt', AGENT_PROMPT)

llm_simulation = Executor(disease, dataloader)
llm_simulation.execute()
```

Build a new simulator: Predator prey model


Agents
<Type 1>
property1
property2
(nn.Parameter)
Objects
<Type A>
property1
property2
(nn.Parameter)
Environment
Interactions
Internal
Agent-Agent
Agent-Object
agent_torch.state (nn.ParameterDict)
Observation
Policy
Transition
agent_torch.function (nn.Module)
state_init
Substep M
(nn.ModuleDict)
state_final
Substep N
(nn.ModuleDict)
...
agent_torch.step (nn.ModuleDict)
=
=
state(s)
Registry & Config
(nn.Modules)
Step K
step(s)
Runner
(nn.Module)
runner.init()
(reset the state before each step)
runner.step()
(run one step of the simulation)
output = runner.trajectory
(parse current and previous states)
loss = loss_fn(output, truth)
(calculate the loss)
optim(runner.parameters())
(optimize learnable parameters)

Scope of tutorial

- Preliminaries
  - Background to automatic differentiation
  - Implement a differentiable ABM
- Algorithms
  - Techniques to calibrate and analyze differentiable ABMs
- Applications
  - Real-world case study in New Zealand
- Systems
  - Tooling to build and calibrate differentiable ABMs at scale

Questions and Discussion

References

Systems

- A framework for learning in Agent-based Models (AAMAS 2024)
- BlackBIRDS: Black-Box Inference foR Differentiable Simulators (JOSS 2023)

Methods

- Differentiable Agent-based Epidemiology (AAMAS 2023)
- Don't Simulate Twice: One-Shot Sensitivity Analyzes via Automatic Differentiation (AAMAS 2023)
- Private Agent-based Modeling (AAMAS 2024)
- Population synthesis as scenario generation for simulation-based planning under uncertainty (AAMAS 2024)

Applications

- Public health impact of delaying second dose of BNT162b2 or mRNA-1273 covid-19 vaccine (BMJ 2021)
- Composing and evaluating interventions with ABM (AAMAS 2024, Best Student Paper Award Finalist!)

Differentiable Agent-based Modeling

Systems, Methods and Applications

Ayush Chopra (MIT)
Arnau Quera-Bofarull (Oxford)
Sijin Zhang (ESR, New Zealand)

Resources: bit.ly/diff-abms

Private Agent-based Modeling

What about the data?

ABM still rely on stale, coarse-grained data

limited granularity due to privacy concerns not scarcity of data.

Thousands of New Zealanders being contacted after personal details leaked in Covid-19 data breach

MOTHERBOARD
TECH BY VICE

T-Mobile 'Put My Life in Danger' Says Woman Stalked With Black Market Location Data

Telecom giants are giving up customers' real-time location data to stalkers and bounty hunters. Now, Motherboard speaks to a victim.

Illinois Bought Invasive Phone Location Data From Safegraph

Report: Indonesian Government's Covid-19 App Accidentally Exposes Over 1 Million People in Massive Data Leak

MOTHERBOARD
TECH BY VICE

Thousands of New Zealanders being contacted after personal details leaked in Covid-19 data breach

T-Mobile 'Put My Life in Danger' Says Woman Stalked With Black Market Location Data

Telecom giants are giving up customers' real-time location data to stalkers and bounty hunters. Now, Motherboard speaks to a victim.

Illinois Bought Invasive Phone Location Data From Safegraph

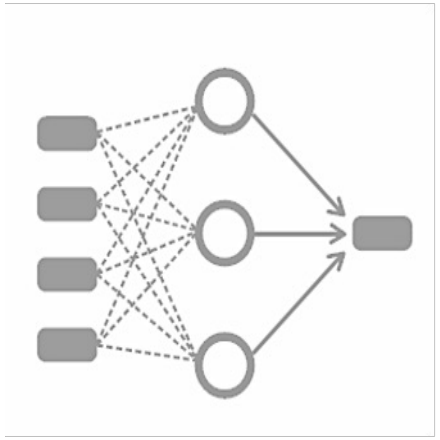
Report: Indonesian Government's Covid-19 App Accidentally Exposes Over 1 Million People in Massive Data Leak

GDPR

HIPAA

how can we access high-resolution data, *securely*?

Conventional focus on "de-sensitizing the data" for simulation



Synthetic generation

Learn distribution and re-sample

high sim2real gap

Differential Privacy

Add noise and release

bad privacy-utility trade-off

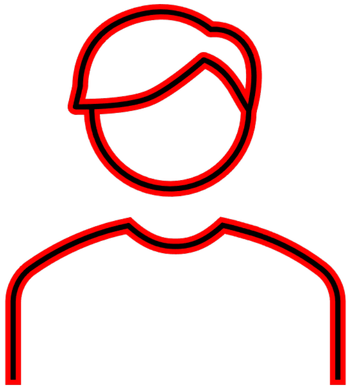
need low sim2real gap + perfect privacy!

decentralize simulation >> centralize data

rethinking the paradigm of agent-based modeling!

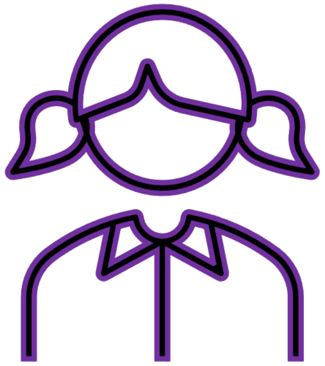
secure multi-party computation

Alice



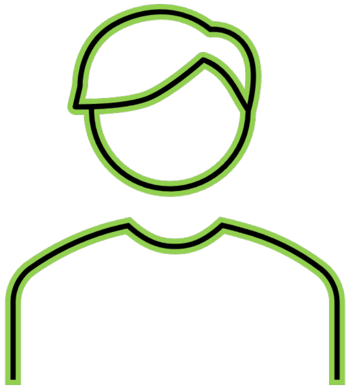
2

Bob



3

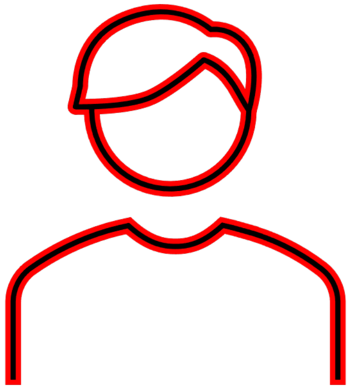
Charlie



5

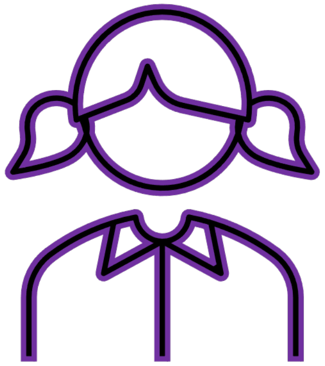
n=11

Alice



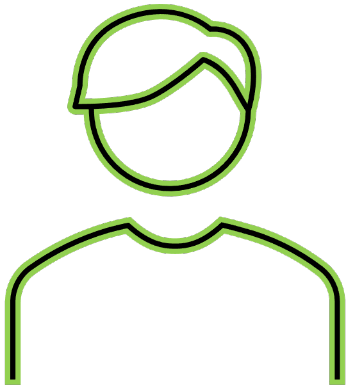
2

Bob



3

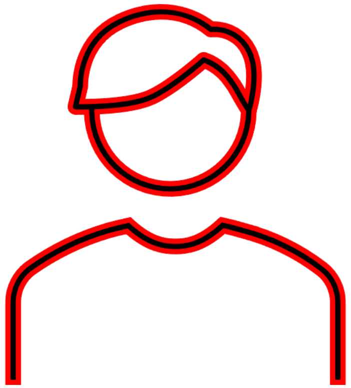
Charlie



5

n=11

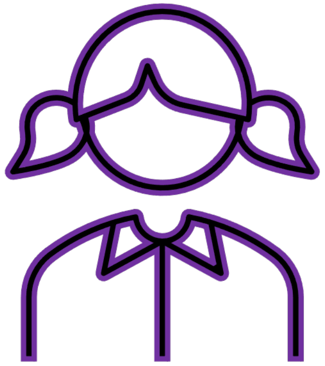
Alice



2

7 + 5 + 1

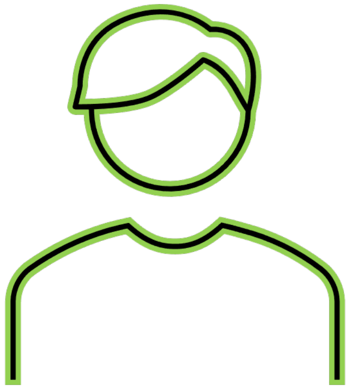
Bob



3

2 + 0 + 1

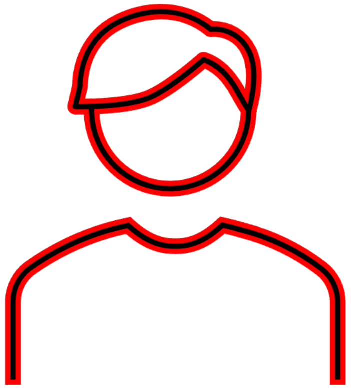
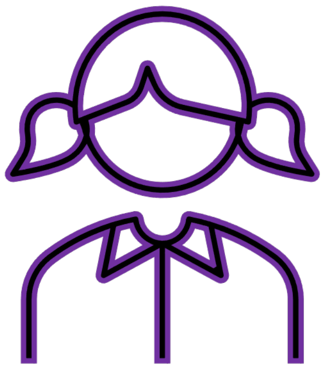
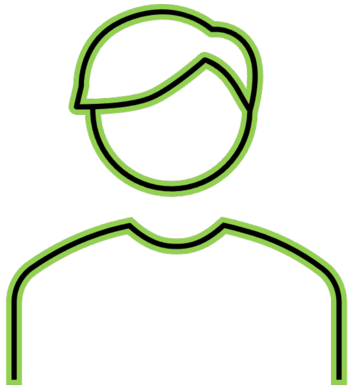
Charlie



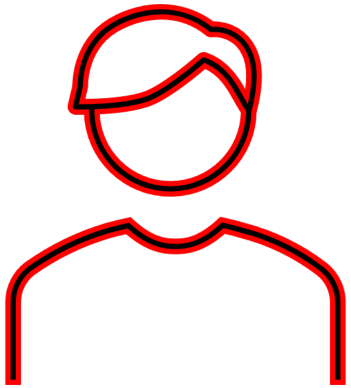
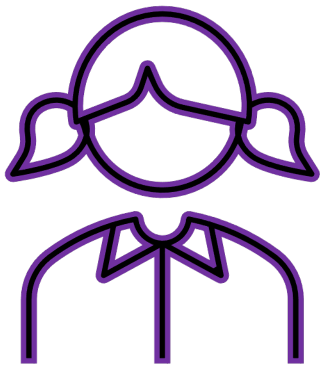
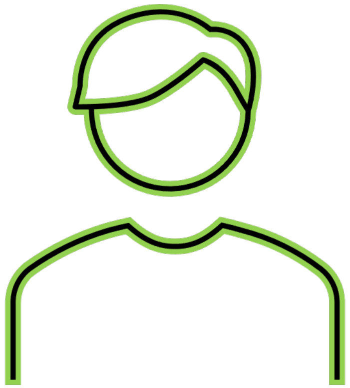
5

3 + 1 + 1

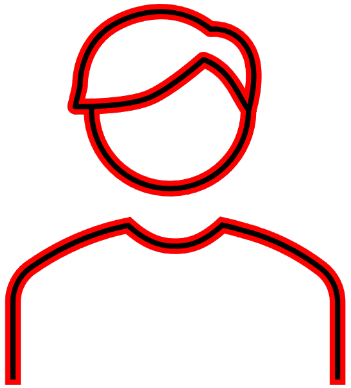
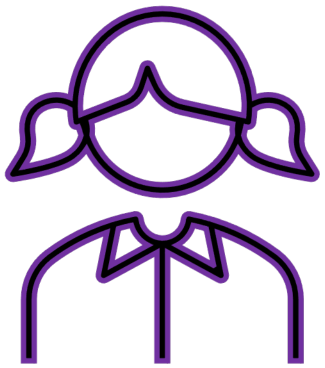
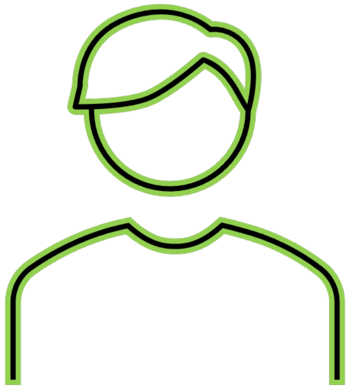
n=11

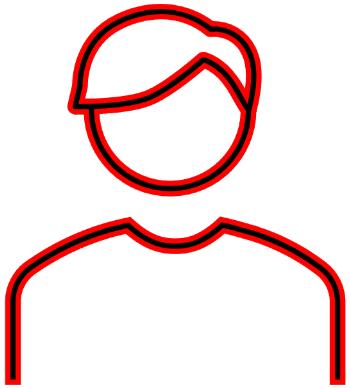
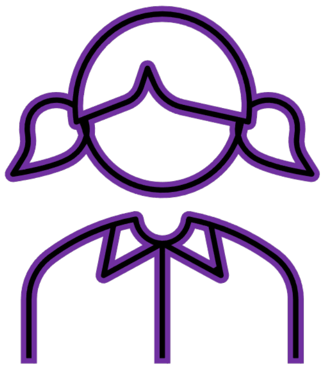
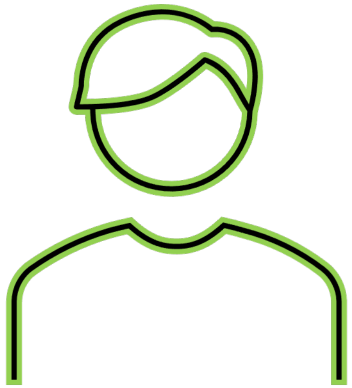
Alice	Bob	Charlie
		
2	3	5
7 + 5 + 1	2 + 0 + 1	3 + 1 + 1
7 + 2 + 3	5 + 0 + 1	1 + 1 + 1

n=11

Alice	Bob	Charlie
		
2	3	5
7 + 5 + 1	2 + 0 + 1	3 + 1 + 1
7 + 2 + 3	5 + 0 + 1	1 + 1 + 1
1	6	3

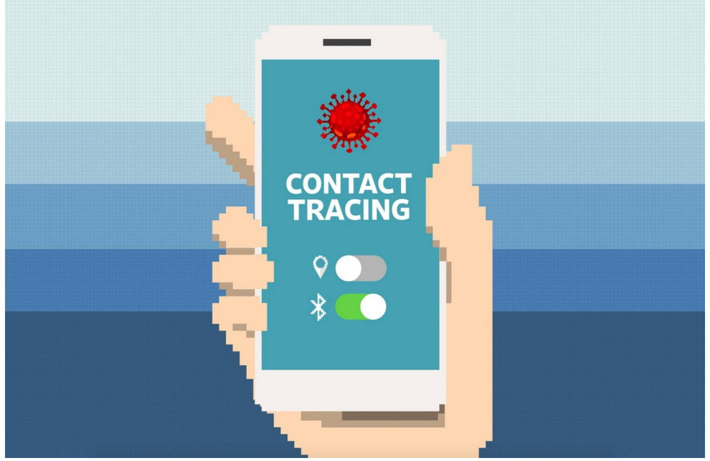
n=11

Alice	Bob	Charlie
		
2	3	5
7 + 5 + 1	2 + 0 + 1	3 + 1 + 1
7 + 2 + 3	5 + 0 + 1	1 + 1 + 1
1	6	3
1 + 6 + 3	1 + 6 + 3	1 + 6 + 3

n=11	Alice	Bob	Charlie
			
	2	3	5
	7 + 5 + 1	2 + 0 + 1	3 + 1 + 1
	7 + 2 + 3	5 + 0 + 1	1 + 1 + 1
	1	6	3
answer	10	10	10

SPMC in action

Settle transactions



Trace infections

ticketmaster®

Pick auction winners

Critical Challenges
Rapid Action
Effective Policies

$$p_{\mathrm{inf}}^{(i)}(t) = 1 - \exp\left(-\frac{\beta\, S_i \Delta t}{n_i} \sum_{j \in \mathcal{N}(i)} I_j(t) \right)$$

disease parameter

susceptibility of agent **i**

transmissibility of neighbor j

$$p_{\text{inf}}^{(i)}(t) = 1 - \exp\left(-\frac{\beta\, S_i \Delta t}{n_i} \sum_{j \in \mathcal{N}(i)} I_j(t)\right)$$

Infection probability for agent **i**

neighborhood of agent **i**

Privacy definition

Secure agent disease (S_i), demographics (I_j) and mobility trace (N_i) data

$$p_{\text{inf}}^{(i)}(t) = 1 - \exp\left(-\frac{\beta \boxed{S_i} \Delta t}{n_i} \sum_{\boxed{j \in \mathcal{N}(i)}} \boxed{I_j(t)}\right)$$

Aggregate *total* transmissibility over all neighbors using additive secret sharing

$$p_{\text{inf}}^{(i)}(t) = 1 - \exp\left(-\frac{\beta \boxed{S_i} \Delta t}{n_i} \underbrace{\sum_{\boxed{j \in \mathcal{N}(i)}} \boxed{I_j(t)}}_{}\right)$$

Additive secret sharing

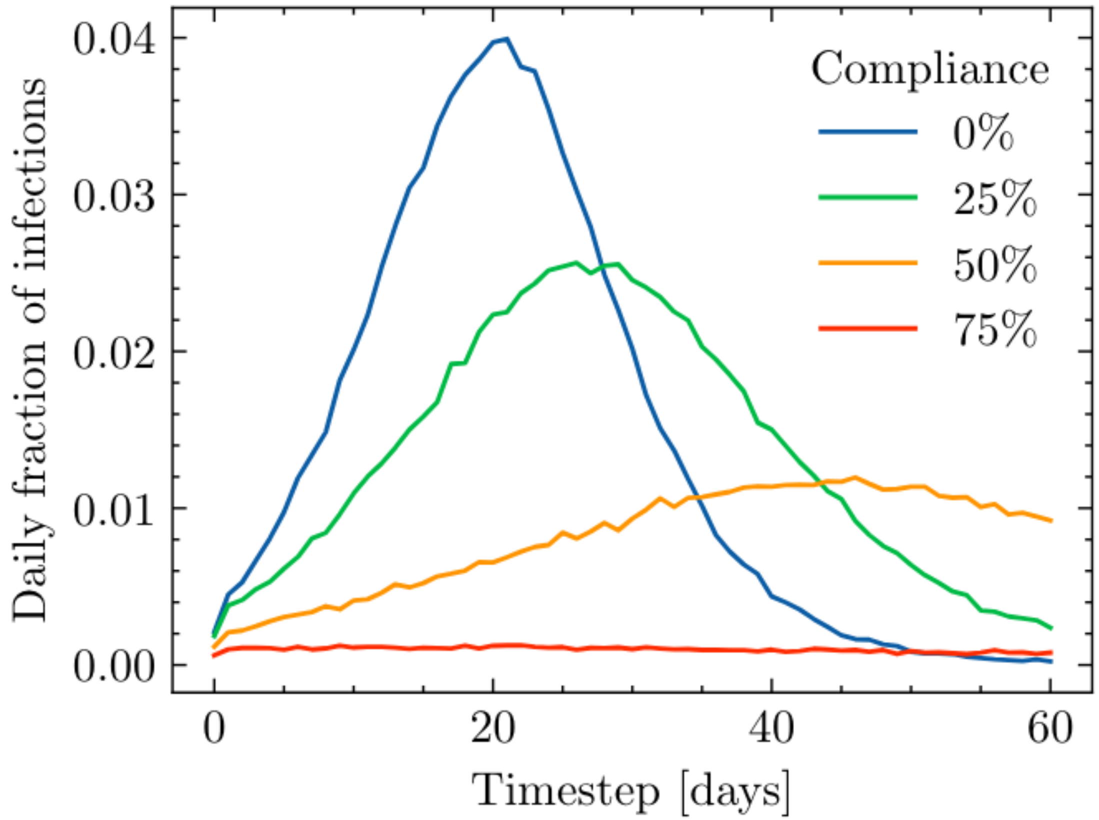
Secure Simulation

Compliance probability
of neighbor j

$$p_{\text{inf}}^{(i)}(t) = 1 - \exp\left(-\frac{\beta\, S_i \Delta_t}{n_i} \sum_{j \in \mathcal{N}(i)} I_j(t)(1 - c_j) \right)$$

How effective will lockdown policies be?

Evaluate interventions without leaking individual disease status or compliance preference



Secure Calibration

Calibrate the disease parameter to total infections at each time step

$$x = \text{number of infections}$$

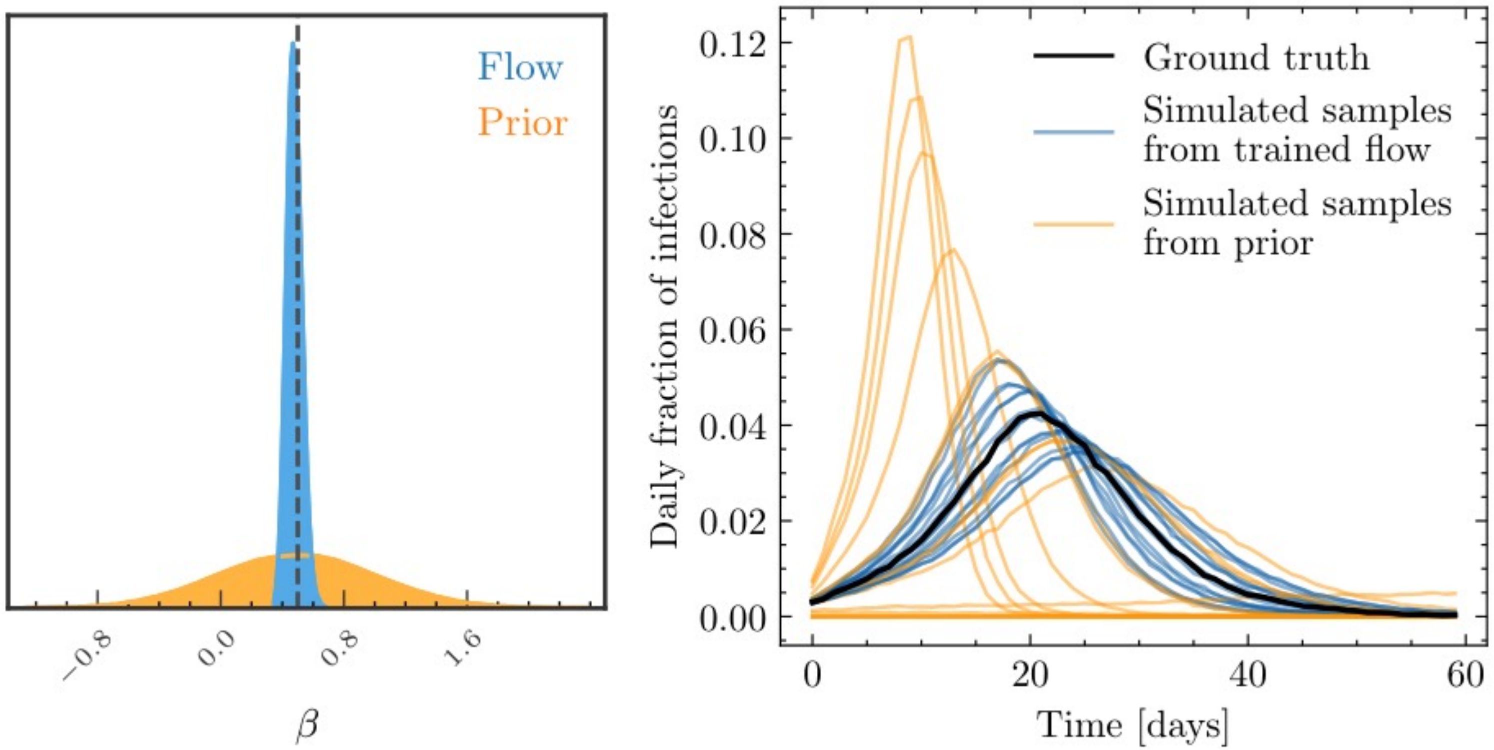
$$p_i = \text{prob agent } i \text{ is infected}$$

$$\frac{\partial x}{\partial \beta} = \sum_i \frac{\partial}{\partial \beta} \mathrm{Bernoulli}(\mathrm{p_i}) \approx \sum_i \frac{\partial p_i}{\partial \beta}$$

We can approximate the total gradient by summing the individual infection gradients (local and private).

Calibrate simulation parameter \beta

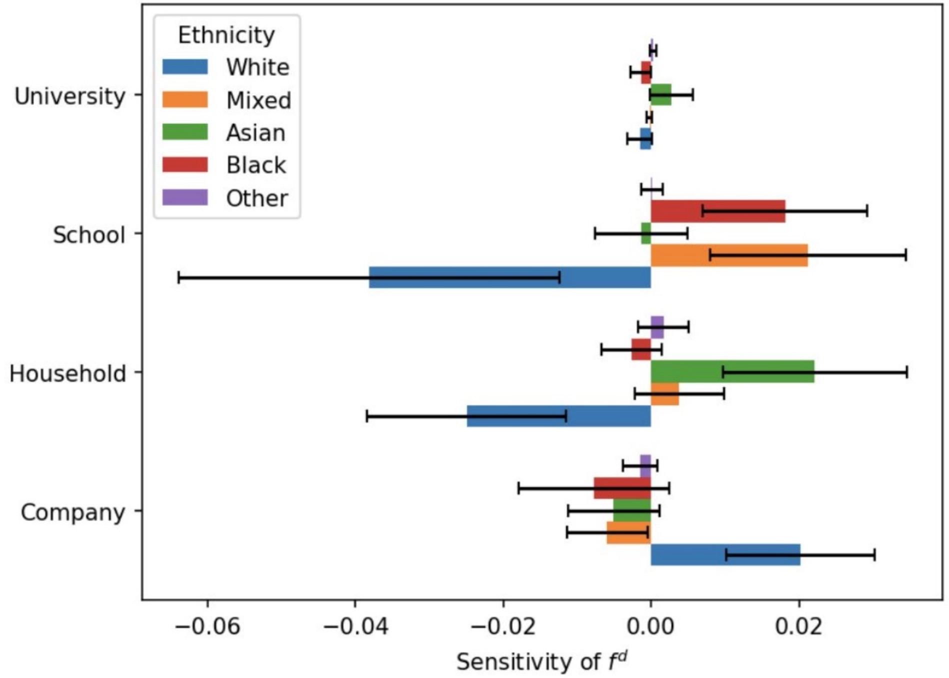
Calibrate disease parameters without leaking an agent's state or interaction trace

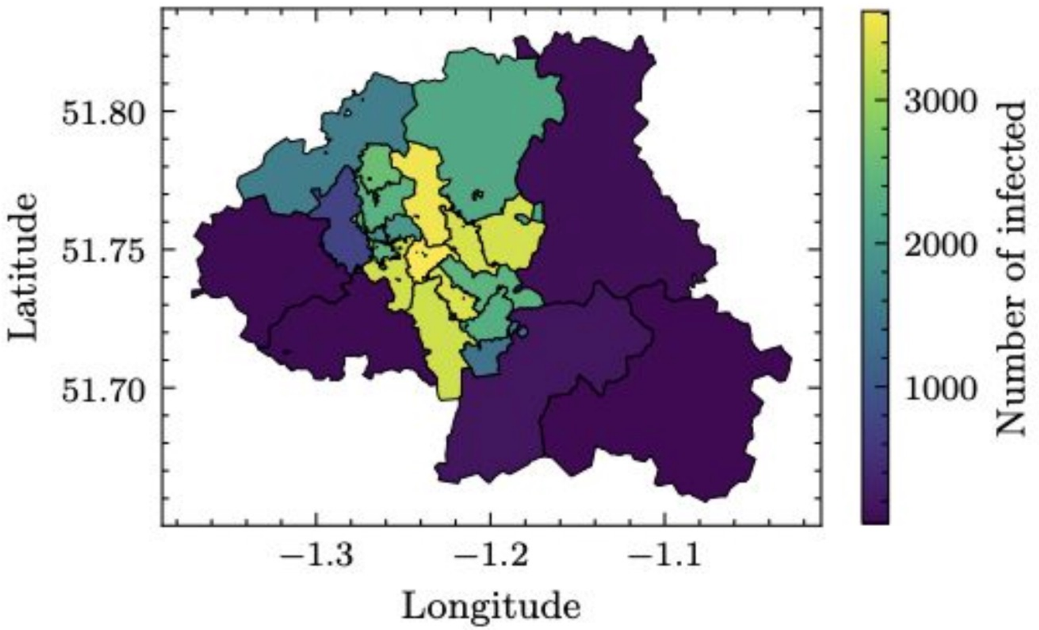


Secure Analysis

How does infection spread across age group and geography?

Analyze dynamics without leaking individual disease, demographic or geo-location



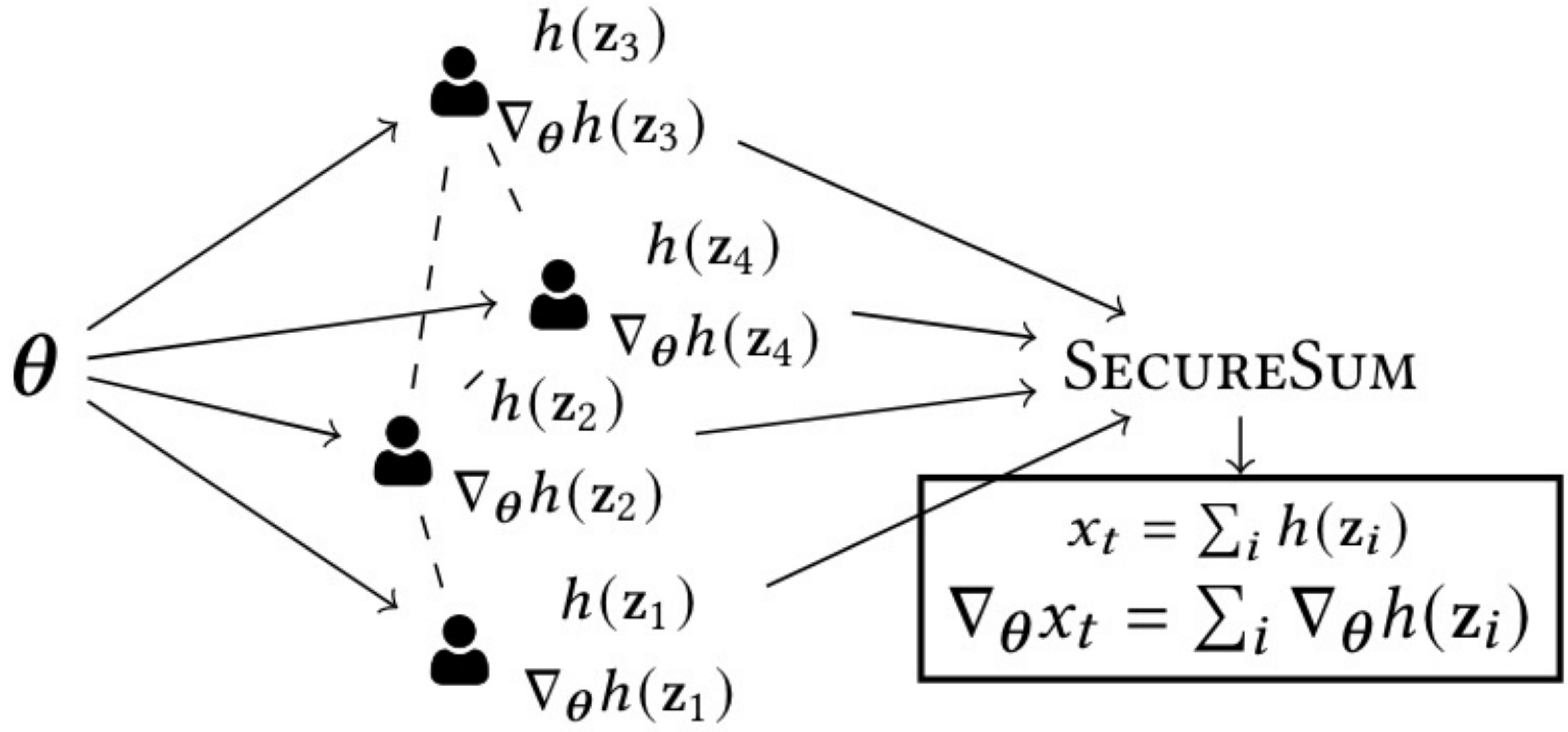


Protocol generalizes to any ABM with "permutation-invariant" message aggregation

$$\mathbf{z}_i(t+1) = f\left(\mathbf{z}_i(t),\ \bigoplus_{j \in \mathcal{N}_i(t)} M_{ij}(t),\ \boldsymbol{\theta}\right)$$

See Section 2 in the paper (https://arxiv.org/pdf/2404.12983)

SMPC to Aggregate Message and Calibration Gradient



Growing trend of decentralized protocols across the world!

UPI
UNIFIED PAYMENTS INTERFACE

Financial networks

ONDC
Open Network for Digital Commerce

Supply chain networks

namma yatri

mobility networks

Differentiable and Private Agent-based Models

github.com/AgentTorch/AgentTorch

Collaborate

ayushc@mit.edu