UNIVERSITY OF AMSTERDAM
Informatics Institute

DNA: Differential Privacy Neural Augmentation for Contact Tracing

Learning on Graphs - Amsterdam, 26 Nov

github.com/RobRomijnders/DNA

Rob Romijnders, Christos Louizos, Yuki M. Asano, Max Welling

ELSEVIER
Brain, Behavior, and Immunity
Volume 89, October 2020, Pages 531-542

Review Article

COVID-19 pandemic and mental health consequences: Systematic review of the current evidence

Nina Vindegaard, Michael Eriksen Benros

Show more

+ Add to Mendeley   Share   Cite

https://doi.org/10.1016/j.bbi.2020.05.048   Get rights and content

ELSEVIER
Best Practice & Research Clinical Anaesthesiology
Volume 35, Issue 3, October 2021, Pages 293-306

2

Economic impact of COVID-19 pandemic on healthcare facilities and systems: International perspectives

Alan D. Kaye MD, PhD (Provost & Vice Chancellor of Academic Affairs) [a],
Chikezie N. Okeagu MD (Assistant Professor) [b], Alex D. Pham MD (Resident Physician) [c],
Rayce A. Silva (Medical Student) [d], Joshua J. Hurley MD, PGY-1 (Resident Physician) [e],
Brett L. Arron MD (Associate Professor) [f], Noeen Sarfraz MD MPH (Resident Physician) [g],
Hong N. Lee MD (Assistant Professor) [h], G.E. Ghali DDS, MD, FACS, FRCS(Ed) (Chancellor) [i],
Jack W. Gamble (Professor and Chairman) [i], Henry Liu MD (Professor) [j],
Richard D. Urman MD (Associate Professor) [k],
Elyse M. Cornett PhD (Assistant Professor) [h]

Show more

Covid-19: Cities fear 'huge' economic impact of restrictions

29 September 2020

BBC, Sept 2020

Covid had negative impact on children's reading - Estyn

4 May

BBC, May 2023

This interactive tool tracks covid-19 travel restrictions by country

Skyscanner's detailed travel map is color-coded in stoplight-style green, yellow and red

Washington Post, December 2020

Privacy is important

"The top reasons against app use were as follows: mistrusting the government, concerns about data security and **privacy**, and doubts about efficacy." Jones et al. 2021

Privacy is important

"The top reasons against app use were as follows: mistrusting the government, concerns about data security and **privacy**, and doubts about efficacy." Jones et al. 2021

"The most cited reasons for not downloading were related to **data** (...) **concerns**" Gao et al. 2022

Privacy is important

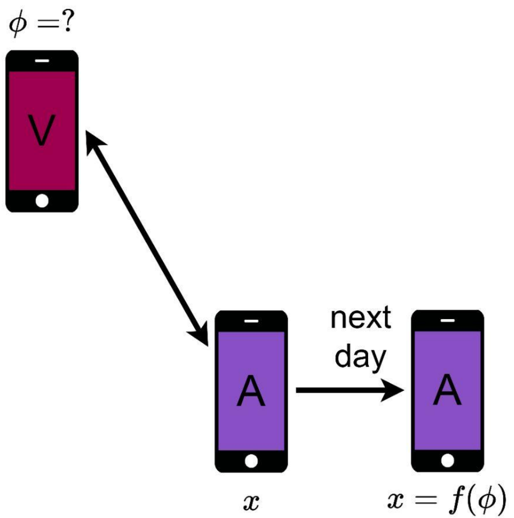
"The top reasons against app use were as follows: mistrusting the government, concerns about data security and **privacy**, and doubts about efficacy." Jones et al. 2021

"The most cited reasons for not downloading were related to **data** (...) **concerns**" Gao et al. 2022

"The main reasons for not downloading and using the app were (...) **worries about privacy**" Walrave et al. 2022

Attack Scenario on contact tracing apps

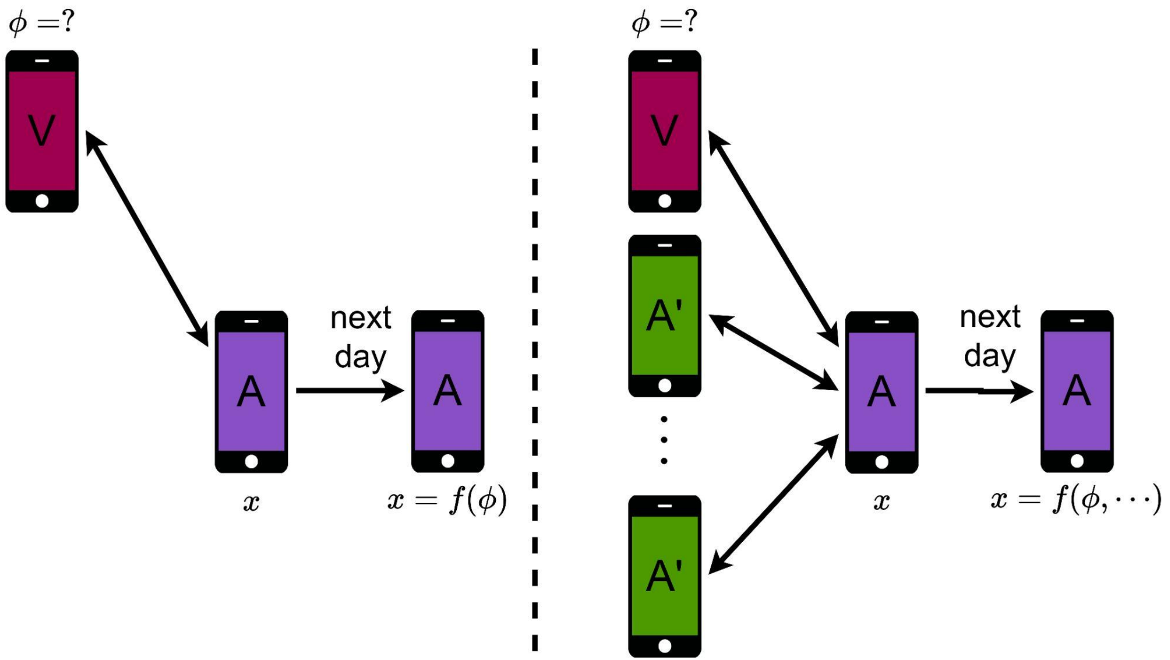
Privacy with respect to released risk score



V is victim, A is attacker

Attack Scenario

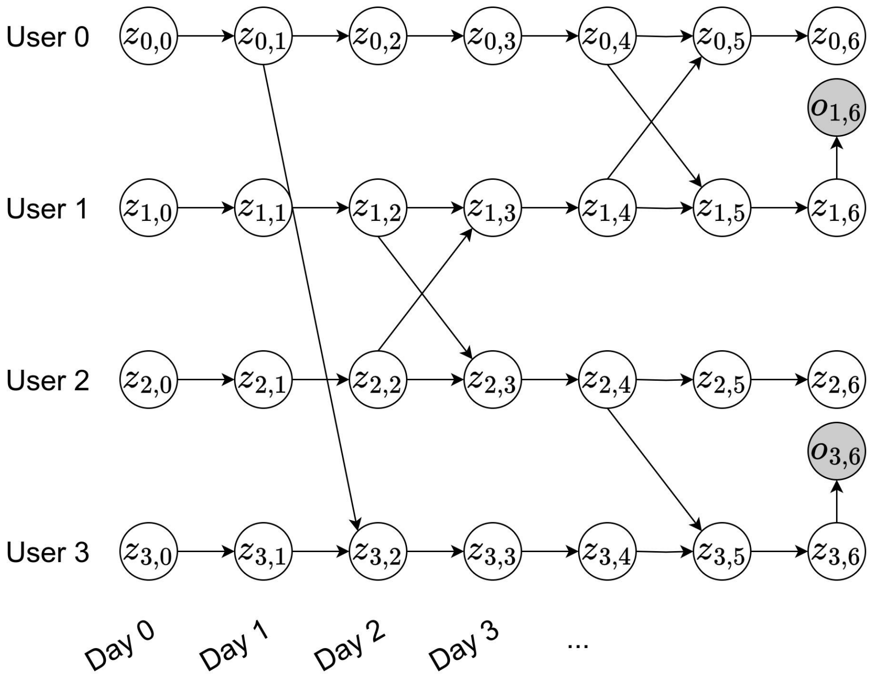
Privacy with respect to released covidscore



V is victim, A is attacker

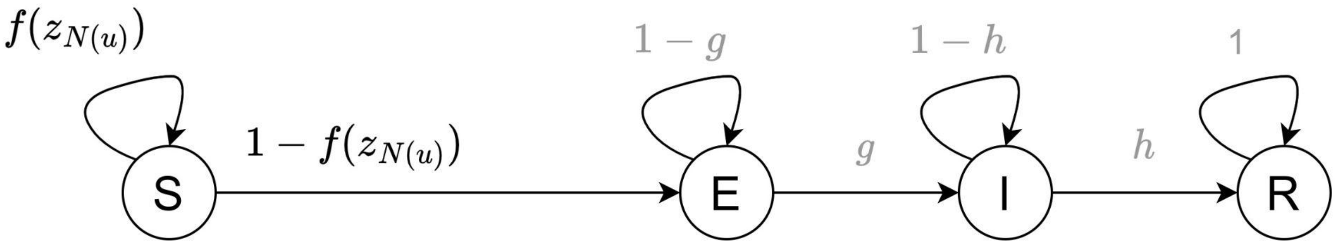
Green phones, A' are co-attackers

Example of graphical model of 4 users in 6 days



Statistical model for contact tracing

Susceptible - Exposed - Infected - Recovered

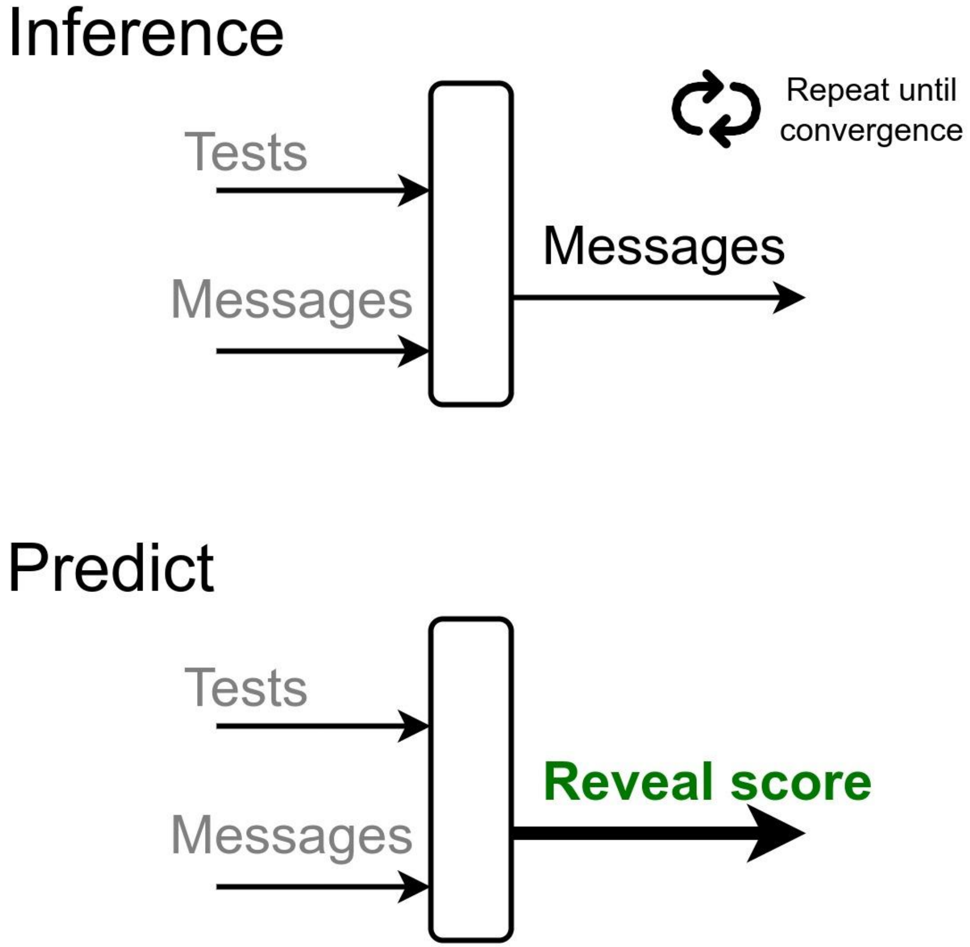


$$f(z_{N(u)}) = (1-p_0)(1-p_1)^{|\{z \in z_{N(u)}:\ z=I\}|}$$

Modular view

Do approximate inference by either Belief Propagation or Factorized Neighbors

$$
\begin{aligned}
b_u(z_u) &= \sum_{z_{N(u)}} P(z_u | z_{N(u)}, \mathcal{O}) B_{N(u)}(z_{N(u)}) \\
&= E_{B_{N(u)}(z_{N(u)})}[P(z_u | z_{N(u)}, \mathcal{O})].
\end{aligned}
$$



Practical use of Differential Privacy

- Emoji suggestions at **Apple**
- QuickType suggestions at **Apple**
- **US Census** releases data under DP
- Executive order US gov. mentions Differential Privacy multiple times
- **Governments** releasing birth rate data
- **Facebook** releases mobility data of users during covid pandemic
- **Google GBoard** language next word prediction
- **LinkedIn** user analytics
- Telemetry on **Windows**

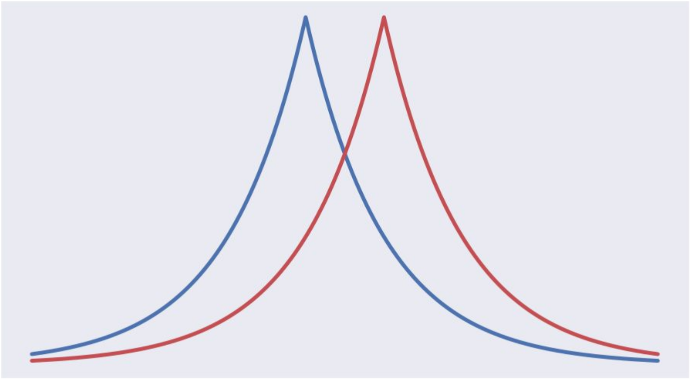
Ponomareva et al. JAIR 2023
Apple blog on DP

Executive order US white house
Facebook Mobility, 2020

LinkedIn (slides, post)
Windows telemetry (NeurIPS paper)

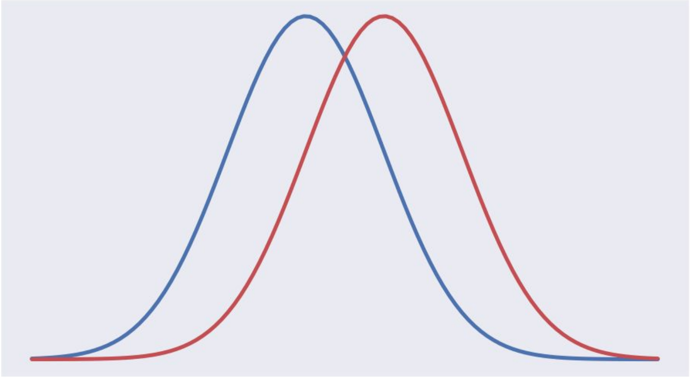
Israeli Birthrate release
Other great overview

Differential privacy



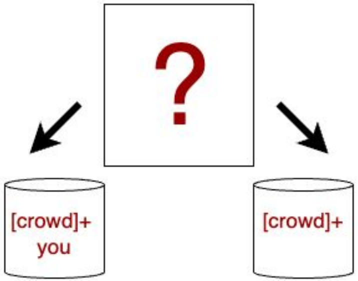
Definition of (ε, δ) differential privacy (Dwork and Roth 2014): for every $\varepsilon > 0$, $\delta \in [0, 1)$, a mechanism $f(\cdot)$, for any outcome Φ in the range of $f(\cdot)$, and any two adjacent data sets D, D' that differ in at most one element, satisfies the constraint:

$$p(f(D) \in \Phi) \leq e^{\varepsilon} p(f(D') \in \Phi) + \delta$$

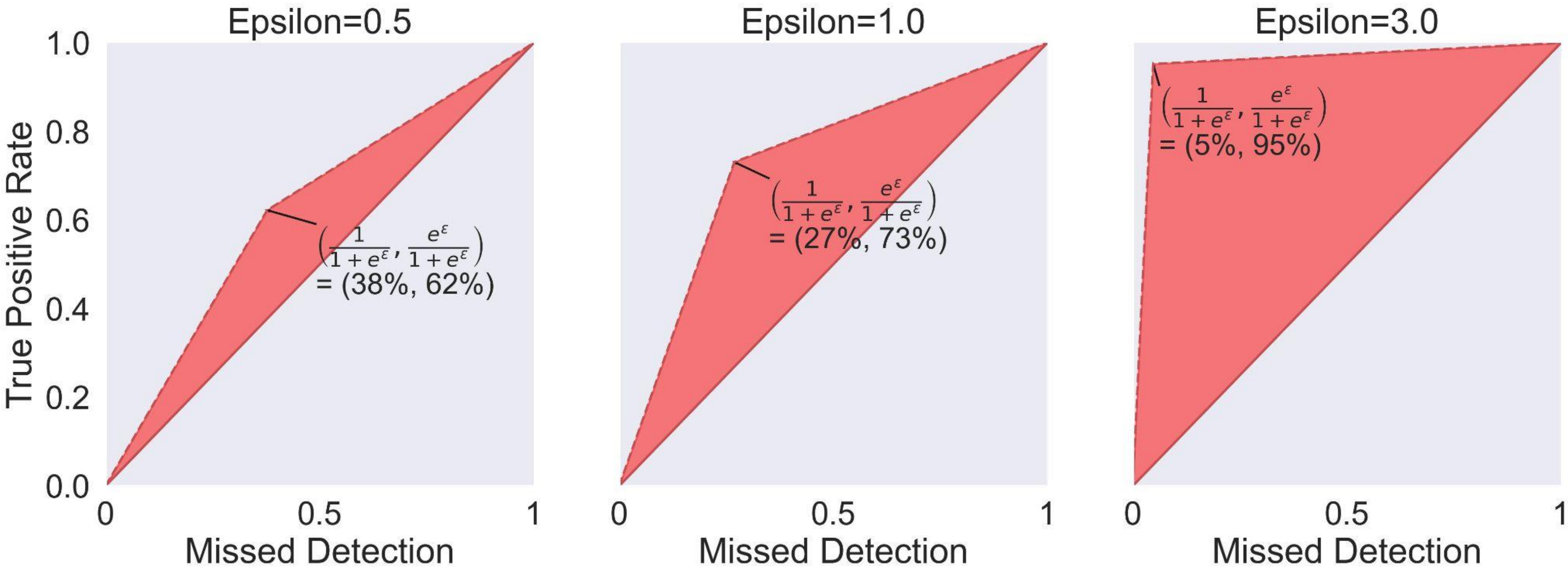


Gaussian Mechanism:

$$\sigma > \frac{\Delta}{\varepsilon}\Big(2\log(\frac{1.25}{\delta})\Big)^{\frac{1}{2}}$$



What does DP mean?



https://en.wikipedia.org/wiki/Differential_privacy#Hypothesis_testing_interpretation

Privacy bound, definition of adjacent datasets

Dataset:

$$D = \{(\mu_i, t_i)\}_{i=1}^{C}$$

Sensitivity:

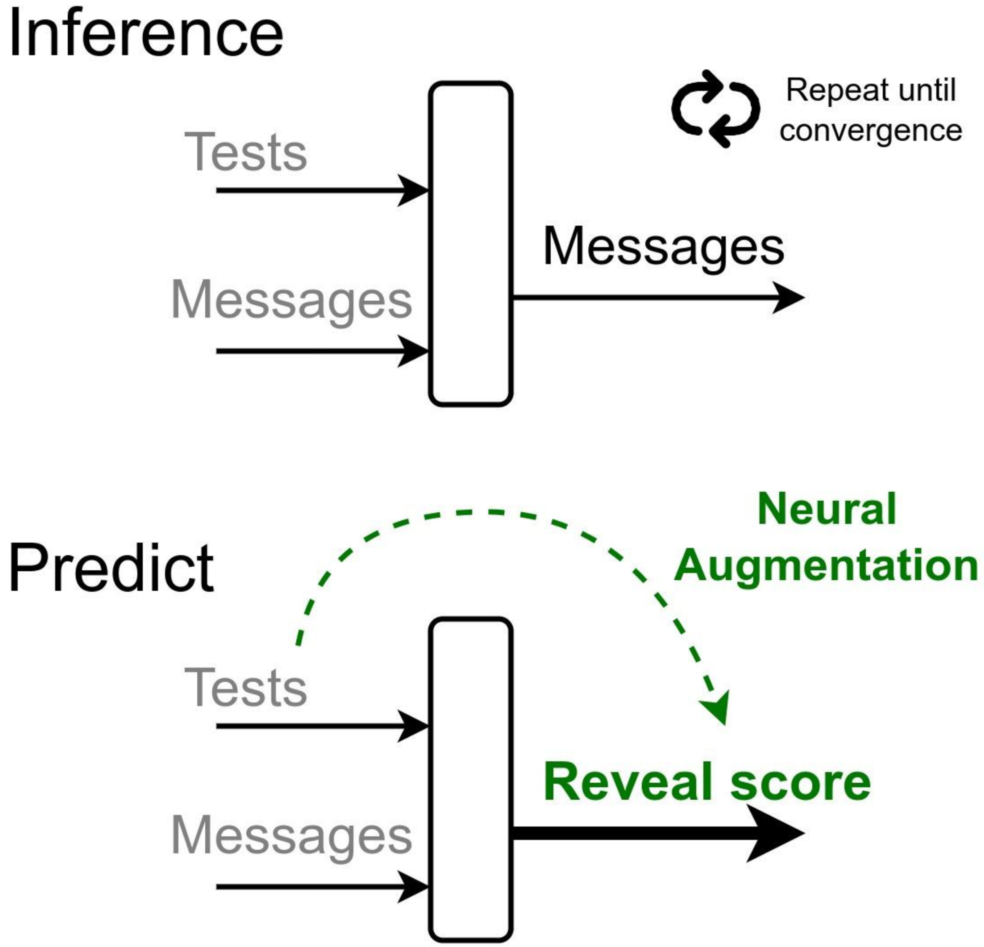
$$\Delta = \max_{\mu_1, \mu_1' \in [0, \gamma_u]} |F(\ (\mu_1, t_1) \cup D\ ) - F(\ (\mu_1', t_1) \cup D\ )| \leq p_1 \gamma_u \qquad \forall\ D.$$

p1 around 0.05, and gamma around 0.7

Neural Augmentation

Neural augmentation known from:

- MRI reconstruction
  (Lønning et al. Medical image analysis, 2019)
- Enhanced belief propagation
  (Satorras et al., AISTATS 2021)
- Fast sparse coding
  (Gregor et al. ICML 2010)



Lipschitz-bounded Neural Network

$$\phi = G_\theta(\{(\mu_i, t_i)\}_{i=1}^{C_T}) = g_\theta^{(2)}\left(\frac{1}{C} \sum_i g_\theta^{(1)}([\mu_i, t_i]^T) \right)$$

During training: estimate Lipschitz constant with power iterations O(p^2)
During testing: calculate Spectral norm exactly once O(p^3)

“Deep sets” Zaheer et al. NeurIPS 2017
“Spectral Normalization..” Miyato et al. ICLR 2018

Make Lipschitz function DP with Gaussian noise

Algorithm 1 DNA: Differentially private Neural Augmentation

Require: Dataset $D = \{(\mu_i, t_i)\}_{i=1}^{C_T}$, constants $p_1, \gamma_u \in (0, 1)$;

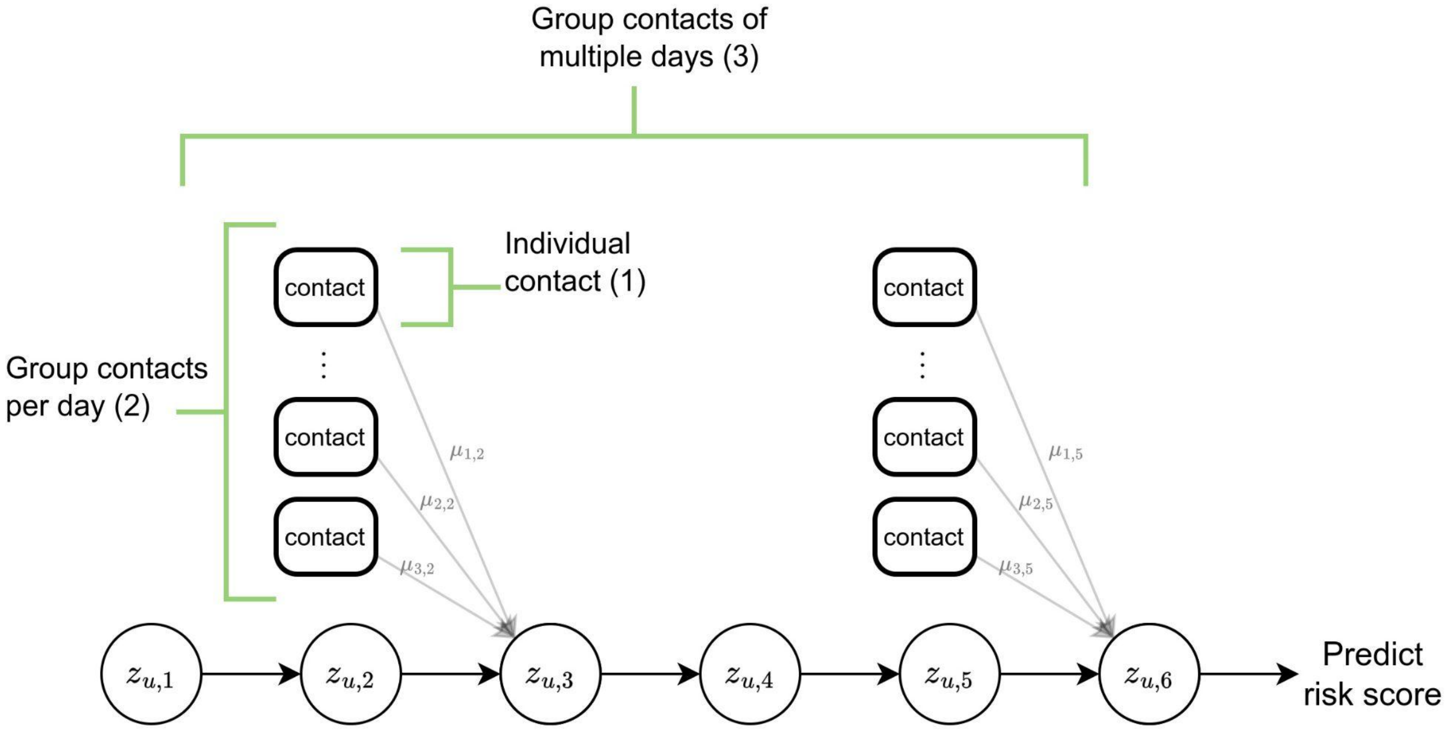
$\mu_i \leftarrow \min(\mu_i, \gamma_u)$

$\bar{\phi} \leftarrow F(\{(\mu_i, t_i)\}_{i=1}^{C_T}) + p_1 \times G_\theta(\{(\mu_i, t_i)\}_{i=1}^{C_T})$

$\phi \leftarrow \bar{\phi} + \mathcal{N}(0, \frac{2}{\varepsilon^2}(\gamma_u p_1 (1 + \frac{1}{C_T}))^2 \log(\frac{5}{4\delta}))$

Jha et al. SIAM Journal on computing 2013

Privacy hierarchy



Different algorithms to compare on simulator

- Traditional contact tracing (Baker et al. 2021)
- Per-message, level 1 (Romijnders et al. 2023)
- Per-day, DPFN, level 2 (Romijnders et al. 2024)
- Per-window, DPFN-S, level 3 (Ours)
- Per-window, DNA, level 3+ (Ours)

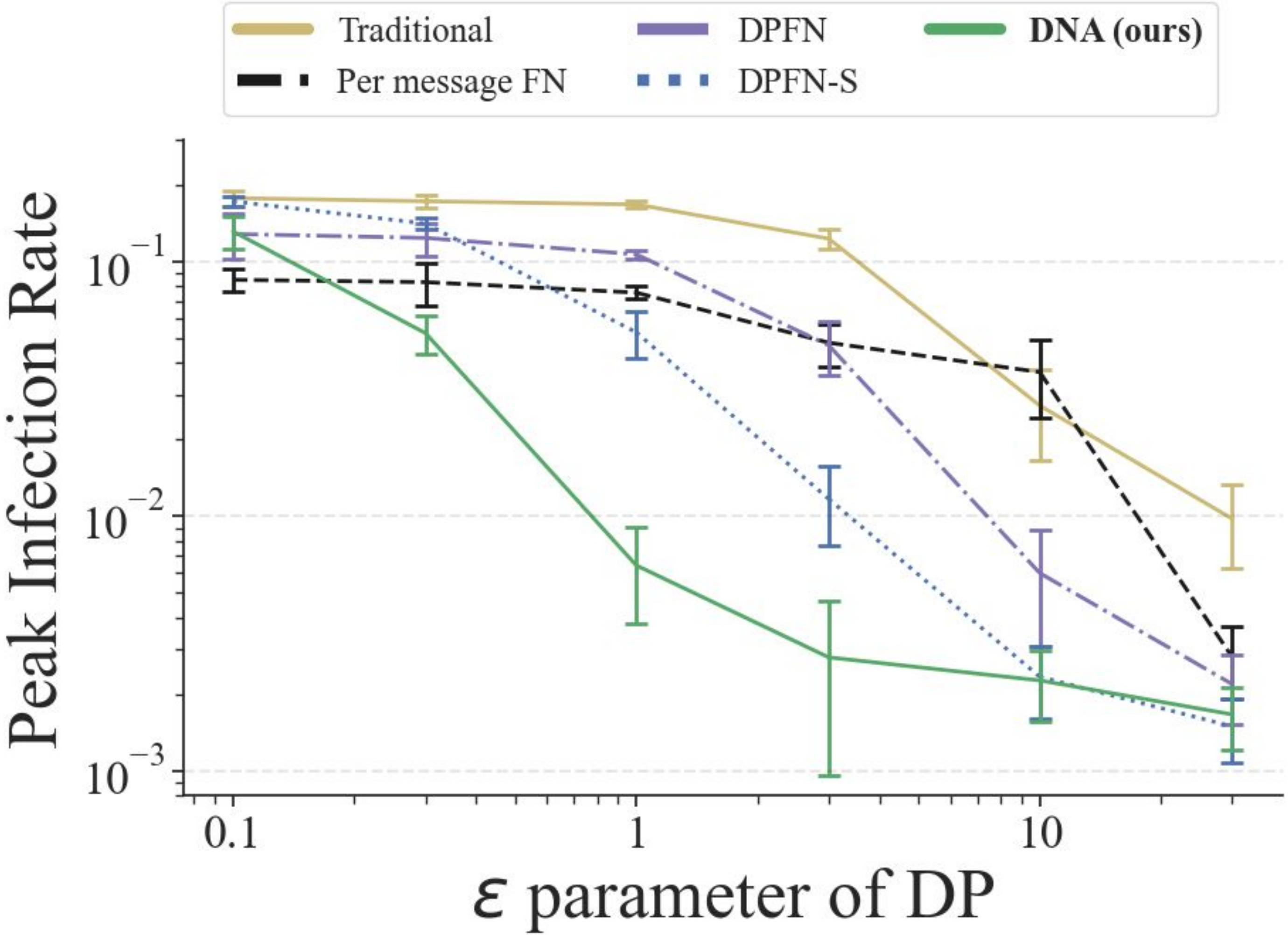
Simulator for experiments

Need simulator as better predictions interact with agents

OpenABM (Hinch et al. 2021)

- Stratifying for
  - 9 age categories
  - 3 occupations
  - 6 household types
- In total 150 parameters calibrated against a typical city in the UK

Experimental results



"DPFN" from Romijnders et al. AAAI 2024

DNA has better utility under various noise scenarios

Even when up to 50% of the agents don't follow the protocol, or when the tests become more noisy, the DNA method **achieves lower infection rate**, compared to the same method without neural augmentation

	DPFN-S (‰)	**DNA** (‰)
Follow protocol		
100%	$52.7_{\pm 10.9}$	$6.4_{\pm 2.6}$
80%	$60.4_{\pm 9.6}$	$6.4_{\pm 2.2}$
50%	$100.1_{\pm 4.4}$	$27.2_{\pm 8.6}$
Noisy tests		
FPR 1%, FNR .1%	$52.7_{\pm 10.9}$	$6.4_{\pm 2.6}$
FPR 10%, FNR 1%	$81.3_{\pm 2.6}$	$19.5_{\pm 2.5}$
FPR 25%, FNR 3%	$130.4_{\pm 1.5}$	$81.3_{\pm 1.8}$

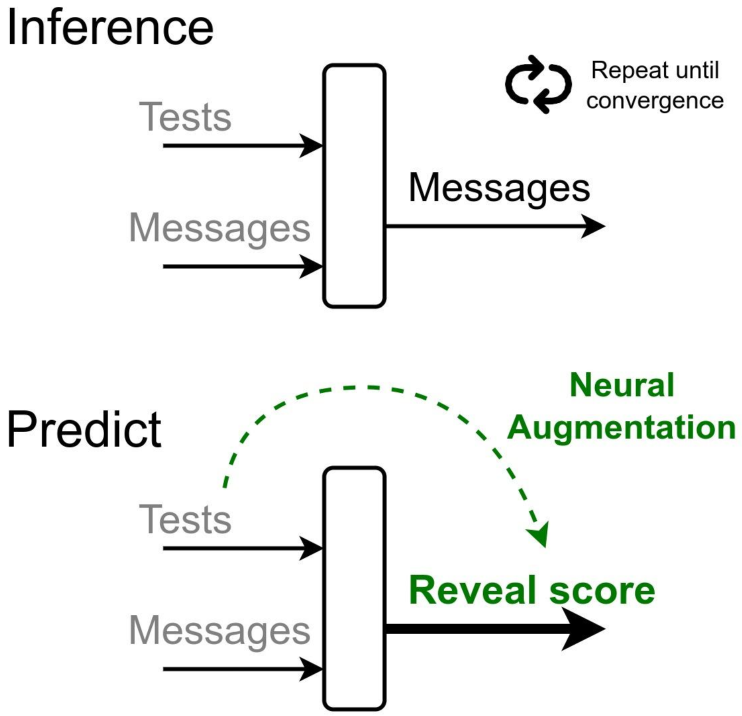
Units are number of infections per thousand agents, ± standard deviation

FPR/FNR = False Positive/Negative Rate

Conclusion

- Novel view of Lipschitz Neural Augmentation as providing Differential Privacy w.r.t. input
- Neural augmentation increases sensitivity, but compares favourably with better predictions
- Future work:
  - Decentralized reinforcement learning, partial adoption

University of Amsterdam
Informatics Institute



DNA: DP Neural Augmentation for Contact Tracing

Questions

r.romijnders@uva.nl; romijndersrob@gmail.com
github.com/robromijnders/dna

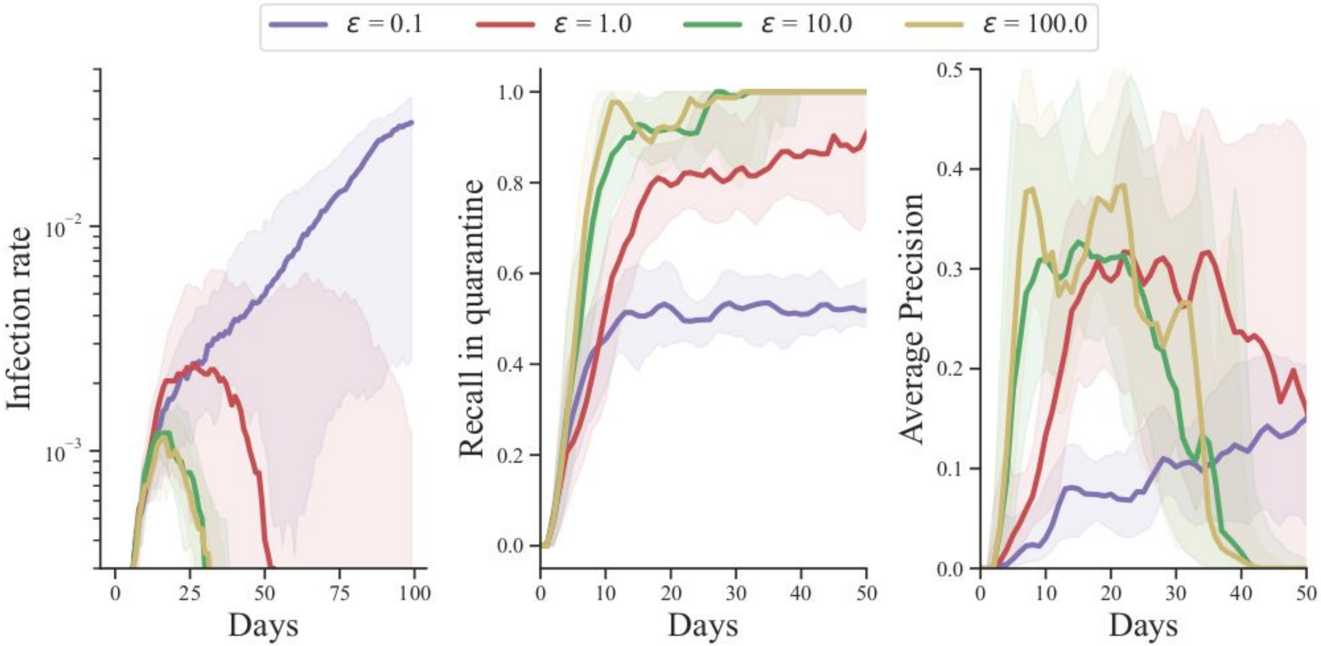


Figure 4: Recall and average precision during a simulation on OpenABM. The shaded regions indicate the 20-80 quantiles of twenty random restarts. The curve $\varepsilon = 1$ achieves a low peak infection rate, which is reflected by a high recall and high average precision in the crucial first month of the epidemic simulation, compared to $\varepsilon = 0.1$. The recall and average precision diagrams only plot 50 days of simulation, which is the crucial phase for a pandemic (Perra 2021).

Practical use of Differential Privacy

- Emoji suggestions at **Apple** (eps=4.)
- QuickType suggestions at **Apple** (eps=8.)
- **US Census** releases data under DP (eps=12.2 per person)
- Executive order US gov. mentions Differential Privacy multiple times
- **Governments** releasing birth rate data (eps=9.98)
- **Facebook** releases mobility data of users during covid pandemic (eps=2.)
- **Google GBoard** language next word prediction (eps=8.9, device level)
- **LinkedIn** user analytics (eps=1.0, record level)
- Telemetry on **Windows** (every six hours, eps=1.0)

Ponomareva et al. JAIR 2023
Apple blog on DP
Executive order US white house
Facebook Mobility, 2020
LinkedIn (slides, post)
Windows telemetry (NeurIPS paper)
Israeli Birthrate release
Other great overview

Inference

Gibbs sampling

$$p(z_u | \hat{z}_{\neg u}, \mathcal{O}).$$

Belief Propagation

$$\mu_{f_s \to z_{u,t}}(z_{u,t}) = \sum_{z_s} f_s(z_s, z_{u,t}) \prod_{k \in \mathrm{Nb}(f_s) \setminus z_{u,t}} \mu_{z_k \to f_s}$$

$$\mu_{z_{u,t} \to f_s}(z_{u,t}) = \prod_{k \in \mathrm{Nb}(z_{u,t}) \setminus f_s} \mu_{f_k \to z_{u,t}}$$

Belief propagation finds lagrange multipliers of ELBO

Claim 1: Let $\{m_{ij}\}$ be a set of BP messages and let $\{b_{ij}, b_i\}$ be the beliefs calculated from those messages. Then the beliefs are fixed-points of the BP algorithm if and only if they are zero gradient points of the Bethe free energy, F_β:

$$F_\beta(\{b_{ij}, b_i\}) = \sum_{ij} \sum_{x_i, x_j} b_{ij}(x_i, x_j) \left[\ln b_{ij}(x_i, x_j) - \ln \phi_{ij}(x_i, x_j)\right] - \sum_i (q_i - 1) \sum_{x_i} b_i(x_i) \left[\ln b_i(x_i) - \ln \psi_i(x_i)\right] \quad (4)$$

subject to the normalization and marginalization constraints: $\sum_{x_i} b_i(x_i) = 1$, $\sum_{x_i} b_{ij}(x_i, x_j) = b_j(x_j)$. ($q_i$ is the number of neighbors of node i.)

Yedidia et al. “Generalised Belief Propagation.” 2000

PGM book, Koller & Fridman

nonnegative, and thus we do not need to enforce these constraints actively. We therefore obtain the following Lagrangian:

$$
\begin{aligned}
\mathcal{J} &= \tilde{F}[\tilde{P}_\Phi, \boldsymbol{Q}] \\
&\quad - \sum_{i \in \mathcal{V}_\mathcal{T}} \lambda_i \left(\sum_{\boldsymbol{c}_i} \beta_i(\boldsymbol{c}_i) - 1 \right) \\
&\quad - \sum_i \sum_{j \in \mathrm{Nb}_i} \sum_{\boldsymbol{s}_{i,j}} \lambda_{j \to i}[\boldsymbol{s}_{i,j}] \left(\sum_{\boldsymbol{c}_i \sim \boldsymbol{s}_{i,j}} \beta_i(\boldsymbol{c}_i) - \mu_{i,j}[\boldsymbol{s}_{i,j}] \right),
\end{aligned}
$$

where Nb_i is the neighbors of $\boldsymbol{C}_i$ in the clique tree. We introduce Lagrange multipliers λ_i for each beliefs factor β_i to ensure that it sums to 1. We also introduce, for each pair of neighboring cliques i and j and assignment to their sepset $\boldsymbol{s}_{i,j}$, a Lagrange multiplier $\lambda_{j \to i}[\boldsymbol{s}_{i,j}]$ to ensure

PGM, Koller & Fridman

Definition 11.1

factored energy functional

Given a cluster tree $\mathcal{T}$ with a set of beliefs $\boldsymbol{Q}$ and an assignment α that maps factors in P_Φ to clusters in $\mathcal{T}$, we define the factored energy functional*:*

$$\tilde{F}[\tilde{P}_\Phi, \boldsymbol{Q}] = \sum_{i \in \mathcal{V}_\mathcal{T}} \boldsymbol{E}_{\boldsymbol{C}_i \sim \beta_i}[\ln \psi_i] + \sum_{i \in \mathcal{V}_\mathcal{T}} \boldsymbol{H}_{\beta_i}(\boldsymbol{C}_i) - \sum_{(i-j) \in \mathcal{E}_\mathcal{T}} \boldsymbol{H}_{\mu_{i,j}}(\boldsymbol{S}_{i,j}), \quad (11.6)$$

Koller et al. “Probabilistic Graphical Models.” 2009

Factorised neighbours

$$b_u(z_u) = \sum_{z_{N(u)}} P(z_u | z_{N(u)}, \mathcal{O}) B_{N(u)}(z_{N(u)})$$
$$= E_{B_{N(u)}(z_{N(u)})}[P(z_u | z_{N(u)}, \mathcal{O})].$$

Rosen-zvi et al. “The dlr hierarchy of approximate inference.” 2012

Factorised neighbours

We obtain an approximation that we refer to as the *factorized neighbors* (FN) algorithm by defining the neighborhood distribution as the factorized expression $B_{N(i)}(x_{N(i)}) = \prod_{j\in N(i)} b_j(x_j)$.

Given this definition, the reduced DLR equations are satisfied when:

$$b_i(x_i) = \sum_{x_{N(i)}} P(x_i|x_{N(i)})B(x_{N(i)}) \ , \qquad (4)$$

where here and elsewhere in the paper we drop the subscript on the B variables to avoid cluttering the notation. The corresponding iterative update takes the form:

$$\begin{aligned} b_i^{t+1}(x_i) &= \sum_{x_{N(i)}} P(x_i|x_{N(i)})B^t(x_{N(i)}) \\ &= \sum_{x_{N(i)}} P(x_i|x_{N(i)}) \prod_{j\in N(i)} b_j^t(x_j) \ . \qquad (5) \end{aligned}$$

It is clear that fixed points of the update rules (5) are solutions of the reduced DLR equations (4).

Equation (4) has appeared previously in the physics literature as the basis for the so-called "hard spin" mean field equations (see references in Pretti and Pelizzola (2003)).

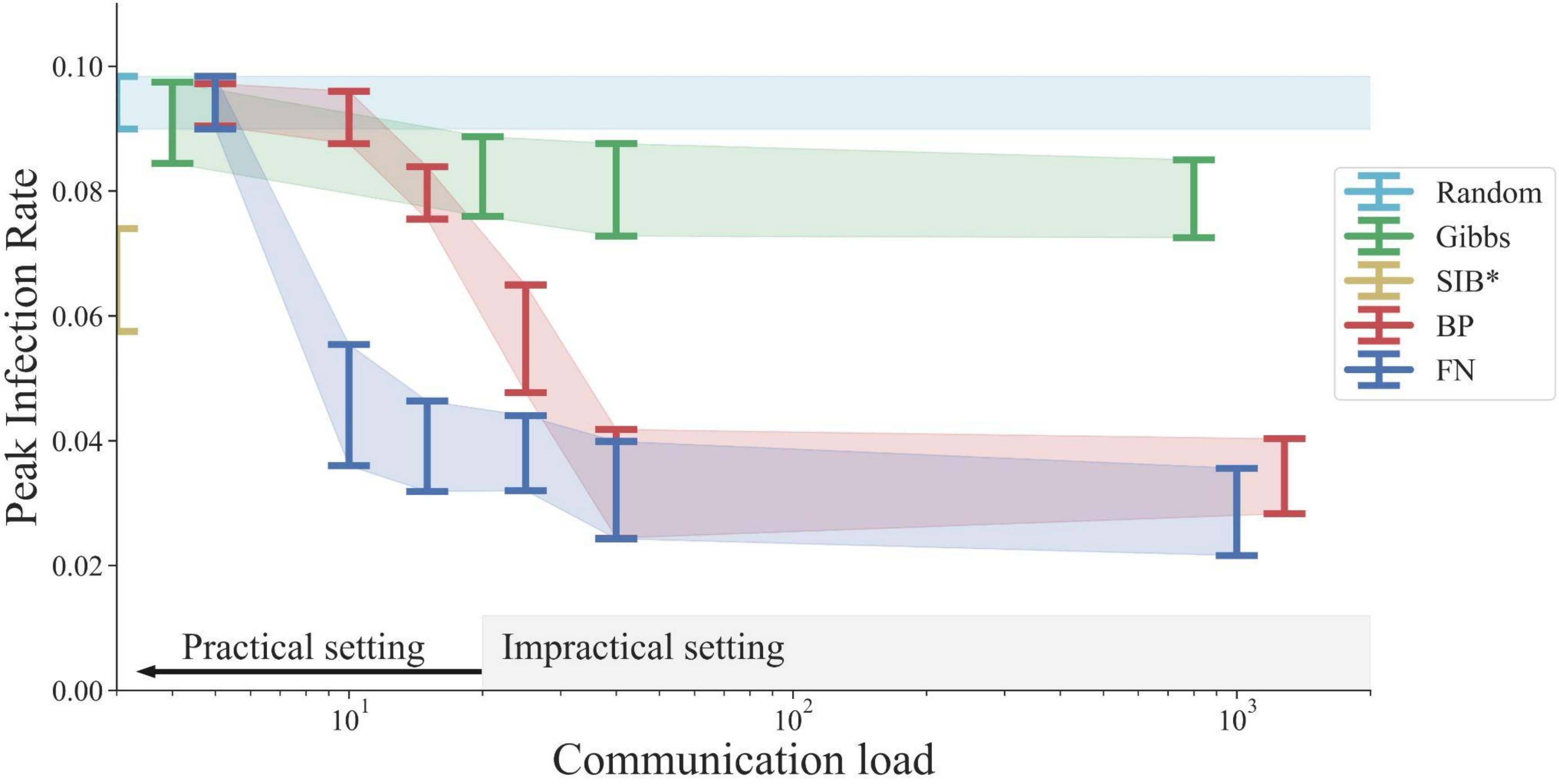
Factorised neighbours

$$E_{B_{N(u)}(z_{N(u)})}\Big[$$

$$p(z_{v,\tau+1} = S | z_{v,\tau} = S, \{z_{v_c,\tau}\}_{c=0}^{C-1})\Big]$$

$$= E_{B_{N(u)}(z_{N(u)})}\left[(1-p_0)\prod_{c=0}^{C-1}(1-p_1)^{\mathbf{1}[z_{v_c},\tau]}\right]$$

$$= (1-p_0)\prod_{c=0}^{C-1} E_{b_c(z_{v_c},\tau)}\left[(1-p_1)^{\mathbf{1}[z_{v_c},\tau]}\right]$$



Romijnders et al. AISTATS 2023

Multiple seeds

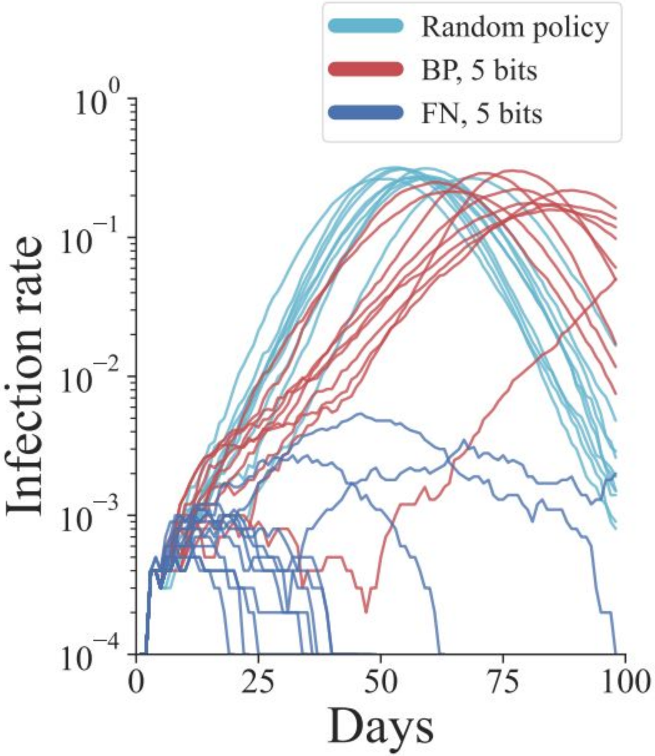


Figure 14: Repeating the experiments on the Open-ABM simulator with 10 random seeds each.

Romijnders et al. AISTATS 2023