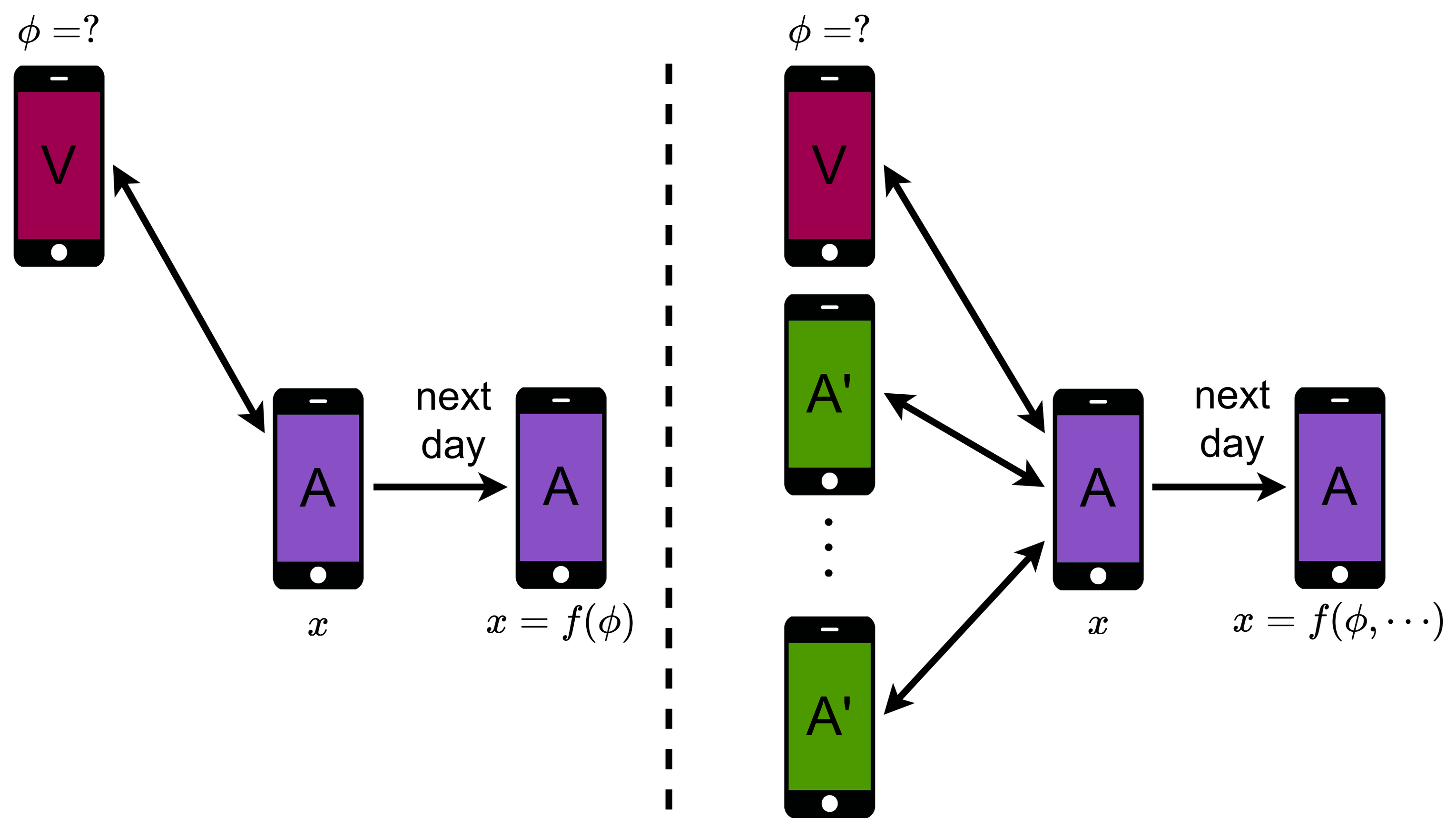
Private ML workshop 2024

DNA: Differentially private Neural Augmentation for contact tracing

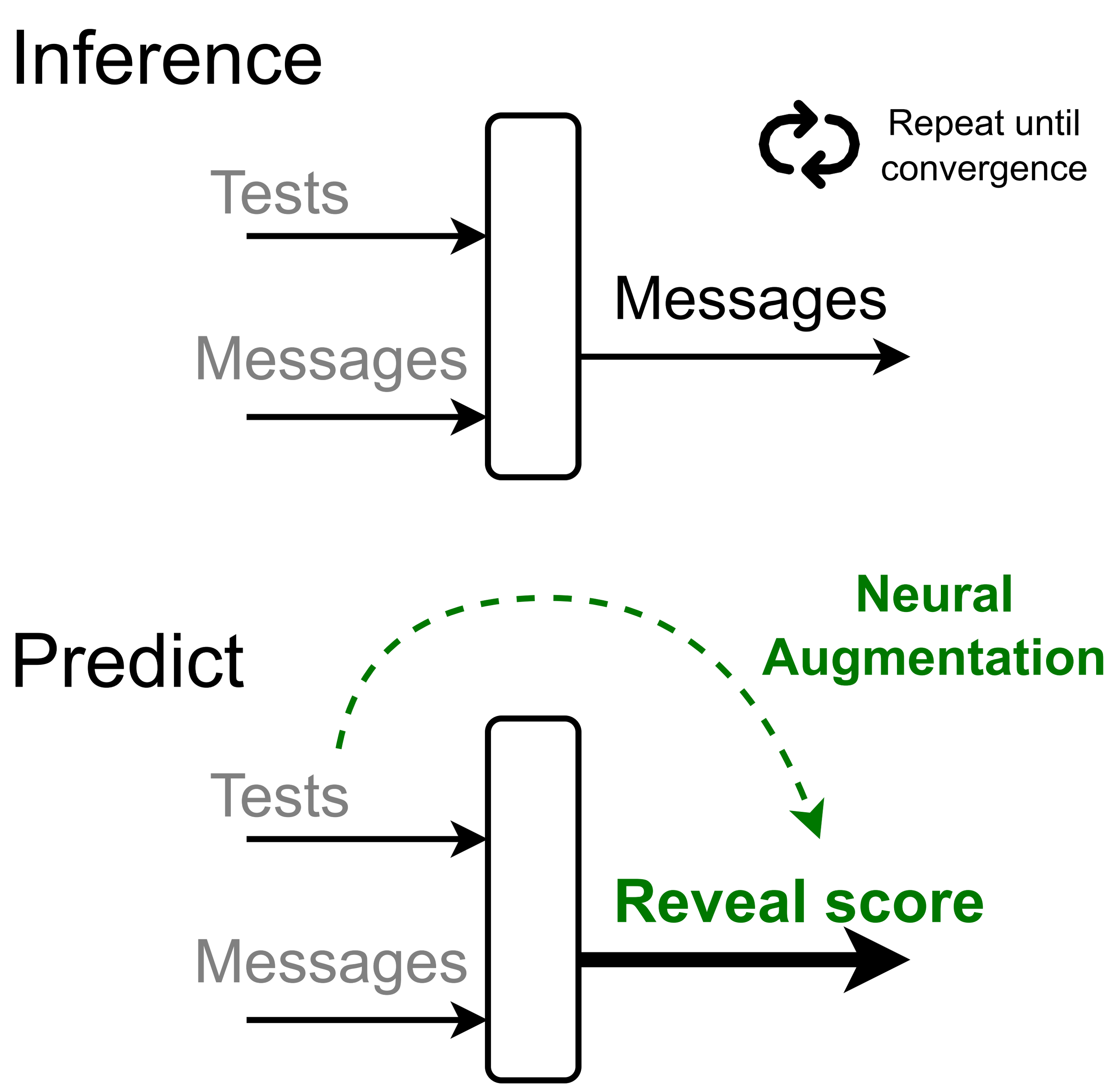
Rob Romijnders[1], Christos Louizos[2], Yuki M. Asano[1], Max Welling[1]

[1]University of Amsterdam, [2]Qualcomm AI research

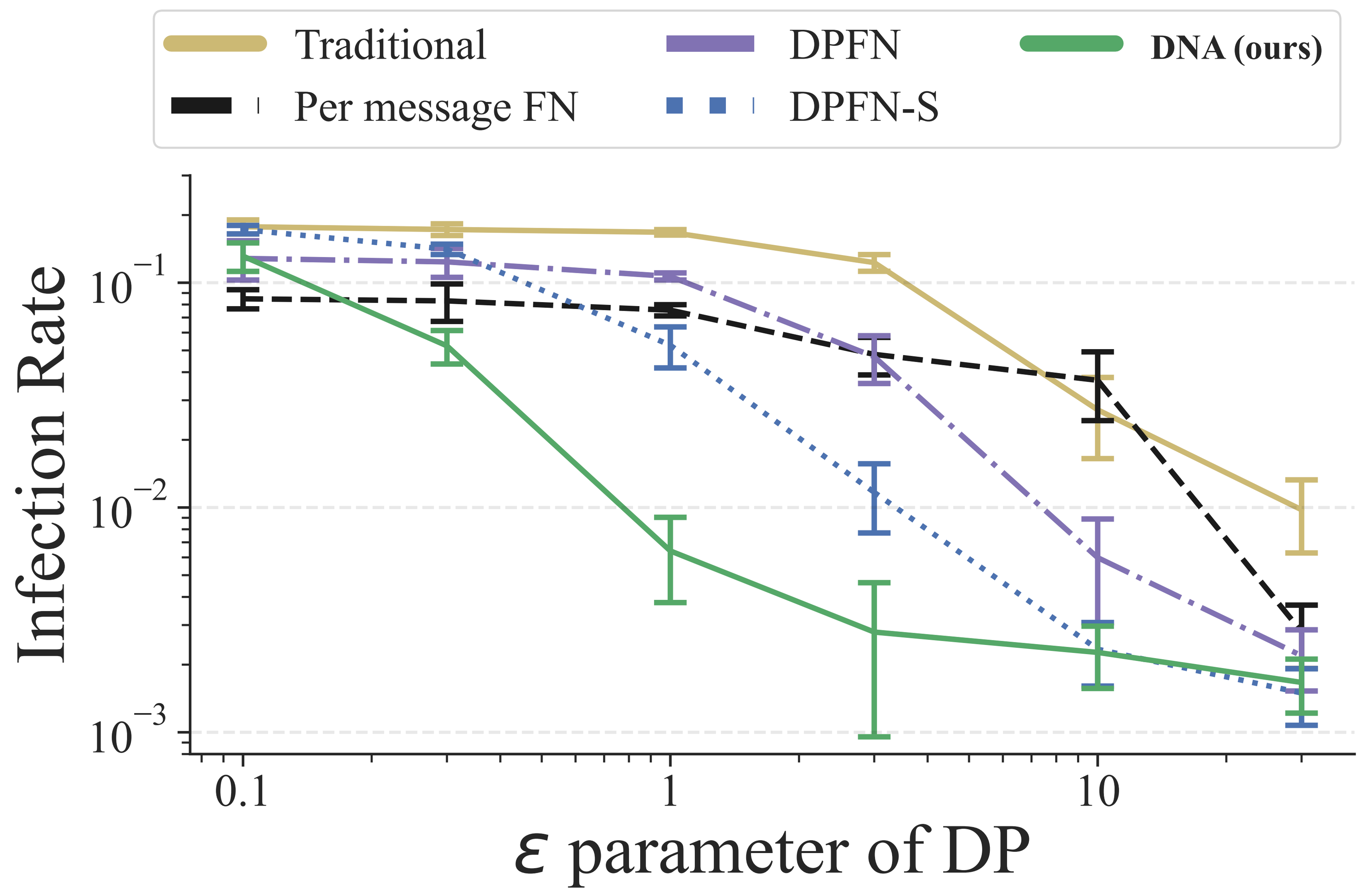
Concerns about privacy are the main reason for low adoption of contact-tracing algorithms, even though they are shown to be effective [1,2,3]. We present a **Differentially Private (DP)** version of Neural Augmentation to improve predictions in decentralized contact tracing.



In the **attack scenario,** the Attacker can reconstruct the score of the Victim -- even in the presence of multiple contacts. Our method reveals the score under DP.



The reveal of the risk score is a DP function (DPFN), but the predictions can be improved with Neural Augmentation.



In the **trade-off** between privacy and the peak infection rate [4], our method achieves a significantly lower infection rate at the crucial setting of ε = 1 DP.

DP definition
For $\varepsilon > 0$, $\delta \in [0,1]$, a function $f(\cdot)$, for any outcome Φ, and any two adjacent data sets D, D', satisfies [6]:

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

Sensitivity
Maximal change with respect to one message, score μ:

$$\Delta = \max_{\mu_1, \mu_1'} |f\big(\{(\mu_1, t_1)\} \cup D\big) - f\big(\{(\mu_1', t_1)\} \cup D\big)| \leq p_1 \gamma_u \;\forall\, D.$$

Neural Augmentation
The Lipschitz-constrained model has a bounded sensitivity [5]:

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

Algorithm 1 DNA: Differentially private Neural Augmentation

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

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

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

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

The neural augmentation (operations in green) increases the sensitivity, but the required additional noise compares favorably in predictions.

romijndersrob@gmail.com
github.com/RobRomijnders/dna

[1] "Epidemic mitigation by statistical inference from contact tracing data," Baker et al. PNAS 2020
[2] "CRISP: A Probabilistic Model for Individual-Level COVID-19 Infection Risk Estimation Based on Contact Data," Herbrich et al. 2020
[3] "Protect Your Score: Contact Tracing With Differential Privacy Guarantees," Romijnders et al. AAAI 2024
[4] "OpenABM-Covid19—An agent-based model for non-pharmaceutical interventions against COVID-19... ," Hinch et al. PLOS 2021
[5] "Deep sets," Zaheer et al. NeurIPS 2017.
[6] "The algorithmic foundations of differential privacy," Dwork et al. Foundations and Trends in Theoretical Computer Science 2014

University of Amsterdam
Faculty of Science

Deep Vision Lab