COVID-19 RISK ESTIMATION FROM CONTACT TRACING DATA

LENKA ZDEBOROVÁ

CO-AUTHORS: A. BAKER, F. KRZAKALA, M. MÉZARD, M. REFINETTI, S. SARAO MANNELLI

COLLABORATING WITH: A. BRAUNSTEIN, LUCA DALL ASTA, ALESSANDRO INGROSSO, INDACO BIAZZO, ANNA PAOLA MUNTONI

HELP AND DISCUSSIONS FROM: YOSHUA BENGIO, IRINA RISH, OTHERS AT MILA, LUCA FERRETTI, IVAN BESTVINA

https://github.com/sphinxteam/sir_inference

CONTACT TRACING DATA

- Information about individuals (stored on the phone of the individual): Age, syndromes, health related-risks, etc.
- Information about contacts (stored on the phone of the two individuals): Time, duration, distance during the contact, barrier-measures used (mask etc.).

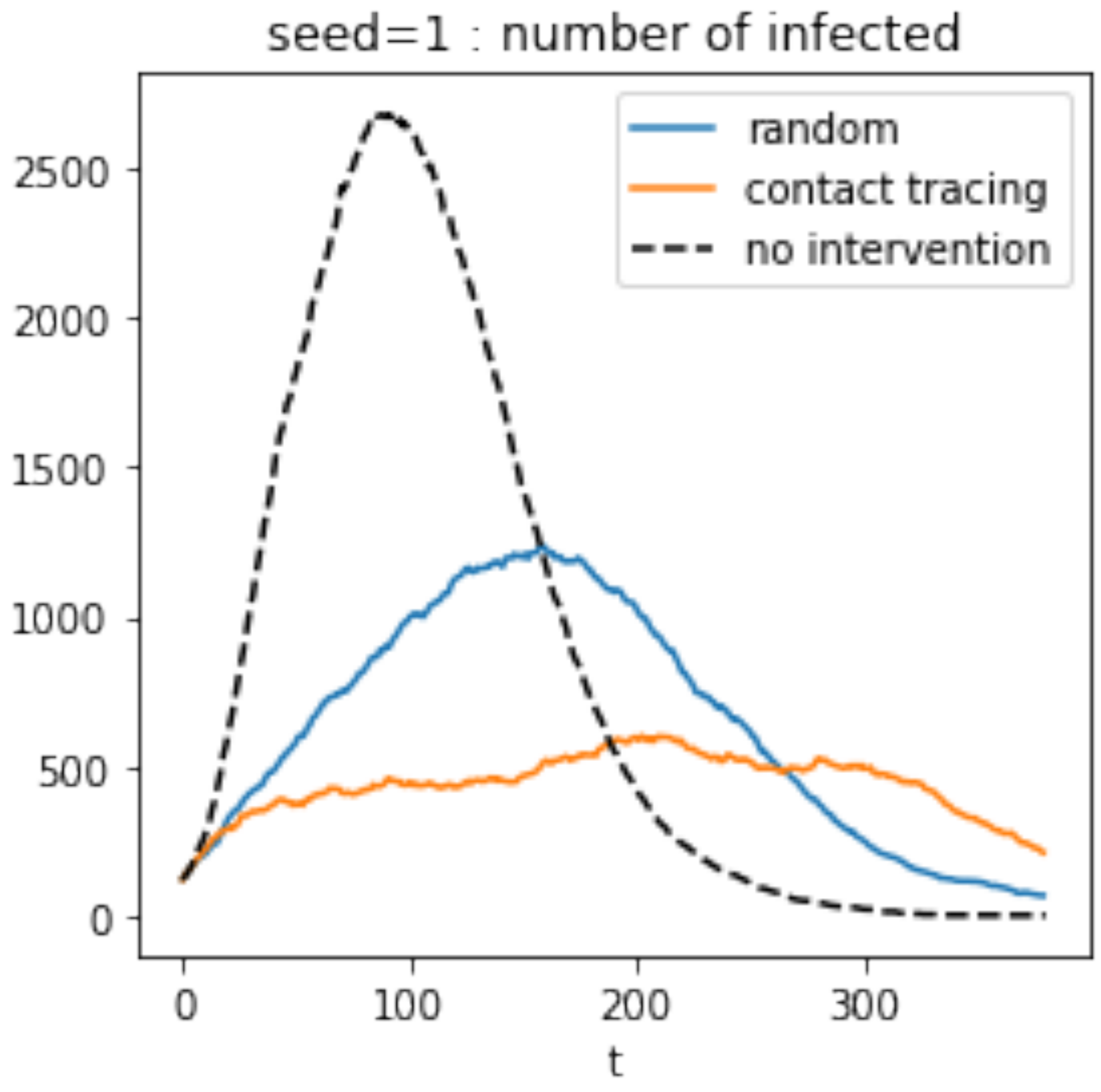
CONTACT TRACING: STATE-OF-THE-ART

- Contact tracing as implemented currently (Google & Apple, DP3T, etc.): Upon a positive test of an individual, his/her recent, sufficiently close, and long contacts are contacted and advised to be tested or to self-isolate.

- Effective Configurations of a Digital Contact Tracing App: A report to NHSX. Hinch, …., Ferretti, et al. (talk at 2nd Ellis-Covid), et. al. https://cdn.theconversation.com/static_files/files/1009/Report_-_Effective_App_Configurations.pdf?1587531217

HOW TRACING INFLUENCES EPIDEMIC SPREAD

- Uncontrolled epidemic vs. Random Tests & isolation vs. Tracing & isolation.



Random geometric contact graph in 2D, scale 1.1,
daily on average 7.4 contacts. Population size= 10000,
λ=0.02 , μ=0.03, Initially 20 infected + 10 time steps of uncontrolled evolution.
Tests: 1 random, 7 infected, and 21 from ranking, 1/2 of all infected, quarantine = 50 steps.

https://github.com/sphinxteam/sir_inference

https://colab.research.google.com/drive/1pRq13j8o6Y8GRWa_IDb_Erj-jqyjWDnM#scrollTo=bVi7bEAGgijX

BETTER THAN TRACING: SMART INFERENCE OF PEOPLE AT RISK (SIPAR)

- Risk can be estimated more accurately than the mere list of contacts. Individual should account for increased risks of their neighbours and spread the information to their neighbours.
- What is needed from the app? Communication between individuals who have been in contact (through a server, in an encrypted manner, only small bandwidth needed). Exchange of simple messages (probabilities) when in contact.
- Apps (we know of) that do estimate (or plan to) a more refined risk level:
  - Covi app, by Bengio & MILA (talk at 1st Ellis-Covid), implemented in Canada: https://docs.google.com/presentation/d/1uZ1-oiaE6LO7O0mxD34MTBD2xP68ED3wV9ed2Yi8fyQ/edit?usp=sharing.
  - ViraTrace, by Bestvina, implemented in India: https://github.com/ViraTrace/InfectionModel.

OUR WORK: DEVELOPMENT OF ALGORITHMS FOR SIPAR

- Mean-field risk estimation (basics in next slides): https://www.overleaf.com/read/tfhcpbvhmcwq
- Dynamical message passing approach (Lokhov, Mézard, Ohta, LZ, PRE '14 & PRE '15)
  In the present context: https://www.overleaf.com/read/tfhcpbvhmcwq
- Belief propagation on trajectories and the probabilistic model that conditions the SIR dynamics to the observations. Github repo: https://github.com/sibyl-team/epibench
  (Altarelli, Braunstein, Dall'Asta et al, PRL'14, Braunstein, Ingrosso Sci. Rep.'16)

SUSCEPTIBLE-INFECTED-RECOVERED (SIR) MODEL

- Population of N individuals

- Spreading of a virus
  - Susceptible individuals (S) → Can be infected
  - Infected individuals (I) → Can infect others
  - Removed individuals (R) → Cannot spread or be infected

Parameters:

- $\lambda_{ij}(t)$ attack rate = probability that if susceptible i meets infected j, j infects i. Depends on the duration and distance of contact, the barrier measures etc
- μ_i: Recovery rate = probability of person i becoming removed in one time-step. Depends on the individual (age, health, etc)

- What is the probability of person i to be in state S,I or R at time t? $P^i_S(t)$, $P^i_I(t)$, $P^i_R(t)$

MEAN-FIELD MESSAGE PASSING (MF)

- time evolution equations for $P^i_S(t)$, $P^i_I(t)$, and $P^i_R(t)$

$$P^i_S(t+1) = P^i_S(t)\left(1 - \sum_{j \in \partial i(t)} P^j_I(t)\lambda_{ij}(t)\right)$$

$$P^i_R(t+1) = P^i_R(t) + \mu_i P^i_I(t)$$

$$P^i_I(t+1) = P^i_I(t) + P^i_S(t) \sum_{j \in \partial i(t)} P^j_I(t)\lambda_{ij}(t) - \mu_i P^i_I(t)$$

Parameters:

- $\lambda_{ij}(t)$: Probability that if susceptible i meets infected j, j infects i:
  - depends on the individuals: barrier measures etc
  - depends on time: duration and distance of contact
- μ_i: Recovery probability of person i:
  - depends on the individual (age, health, etc)
- $\partial i(t)$: Sum over ALL the individuals i was in contact with at time t:
  - Tracked with App

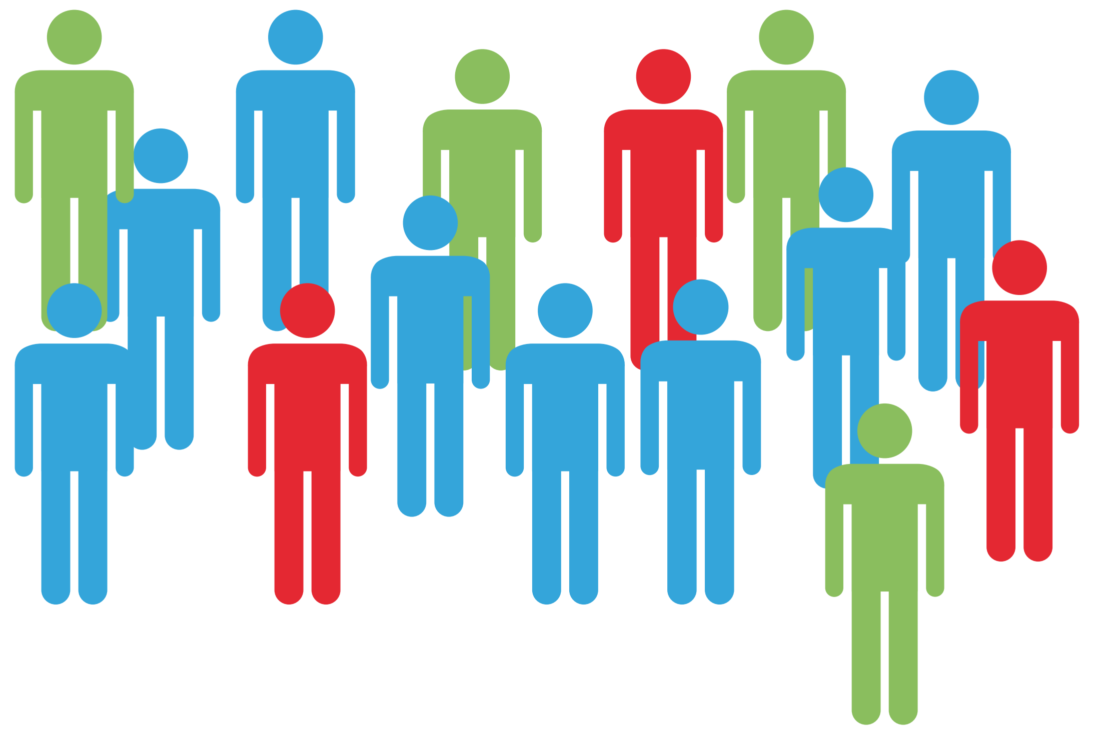
- Given an initial conditions $\{P^i_S(0), P^i_I(0), P^i_R(0)\}$ + Parameters $\Rightarrow$ $\{P^i_S(t), P^i_I(t), P^i_R(t)\}$

Needed: Use test results + symptoms to better estimate the probabilities

https://www.overleaf.com/read/tfhcpbvhmcwq

https://github.com/sphinxteam/sir_inference

FEED BACK LOOP: USE TEST RESULTS

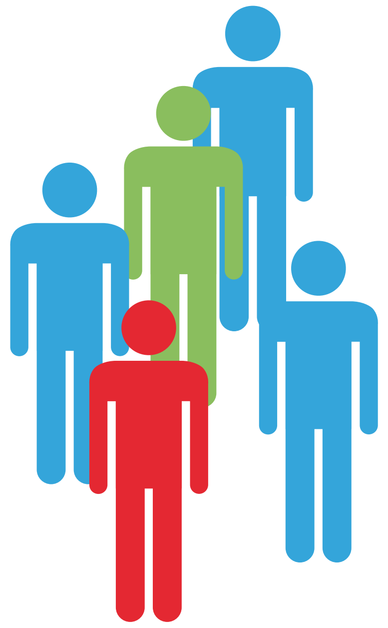


Time evolution

Time t

Test M individuals

Know if they are S, I or R



$\{P_S^i(0), P_I^i(0), P_R^i(0)\}$ — MF equations → $\{P_S^i(t), P_I^i(t), P_R^i(t)\}$

- An infected person has been infectious for the time τ: $P_I^i(t') = 1 \text{ for } t - \tau \leq t' \leq t$
- A recovered person cannot change state: $P_R^i(t') = 1 \text{for} t' \geq t$
- A susceptible individual has not yet been infected: $P_S^i(t') = 1 \text{for} t' \leq t$

$\{P_S^i(t-\tau), P_I^i(t-\tau), P_R^i(t-\tau)\}$ ← Use Information to update probabilities at t-Δt

$\{P_S^i(t-\tau), P_I^i(t-\tau), P_R^i(t-\tau)\}$ — MF equations → $\{P_S^i(t), P_I^i(t), P_R^i(t)\}$

Get better estimates of the probabilities at time t

→ Repeat the process at every time step → Final Inference

→ NB: Symptoms can be incorporated in this feedback loop analogously

Overleaf https://www.overleaf.com/read/tfhcpbvhmcwq

https://github.com/sphinxteam/sir_inference

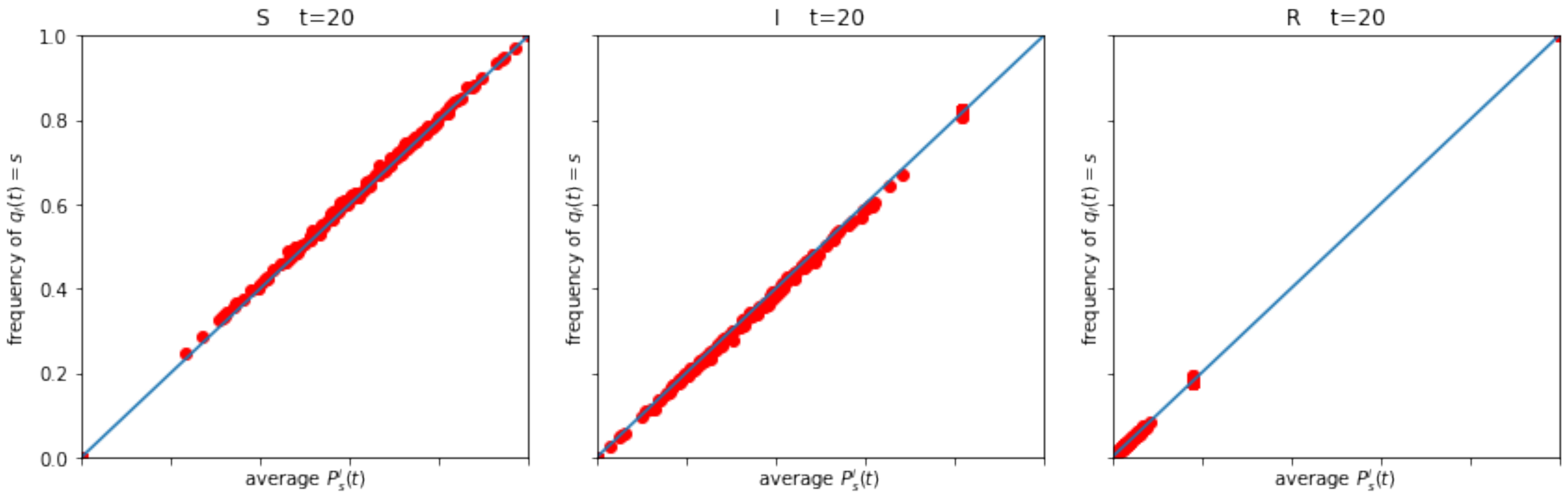
OUR WORK: DEVELOPMENT OF ALGORITHMS FOR SIPAR

- Mean-field risk estimation (just explained): https://www.overleaf.com/read/tfhcpbvhmcwq
- Dynamical message passing approach (Lokhov, Mézard, Ohta, LZ, PRE '14 & PRE '15)
  In the present context: https://www.overleaf.com/read/tfhcpbvhmcwq
- Belief propagation on trajectories and the probabilistic model that conditions the SIR dynamics to the observations. Github repo: https://github.com/sibyl-team/epibench (Altarelli, Braunstein, Dall'Asta et al, PRL'14, Braunstein, Ingrosso Sci. Rep.'16)

PROBABILITY ESTIMATION WORKS

- Validation of the tree-like approximation: Inference (without feed-back) correlates well with ground truth

Scatter plot: Y-axes average over 10000 simulations, X-axes MF risks.



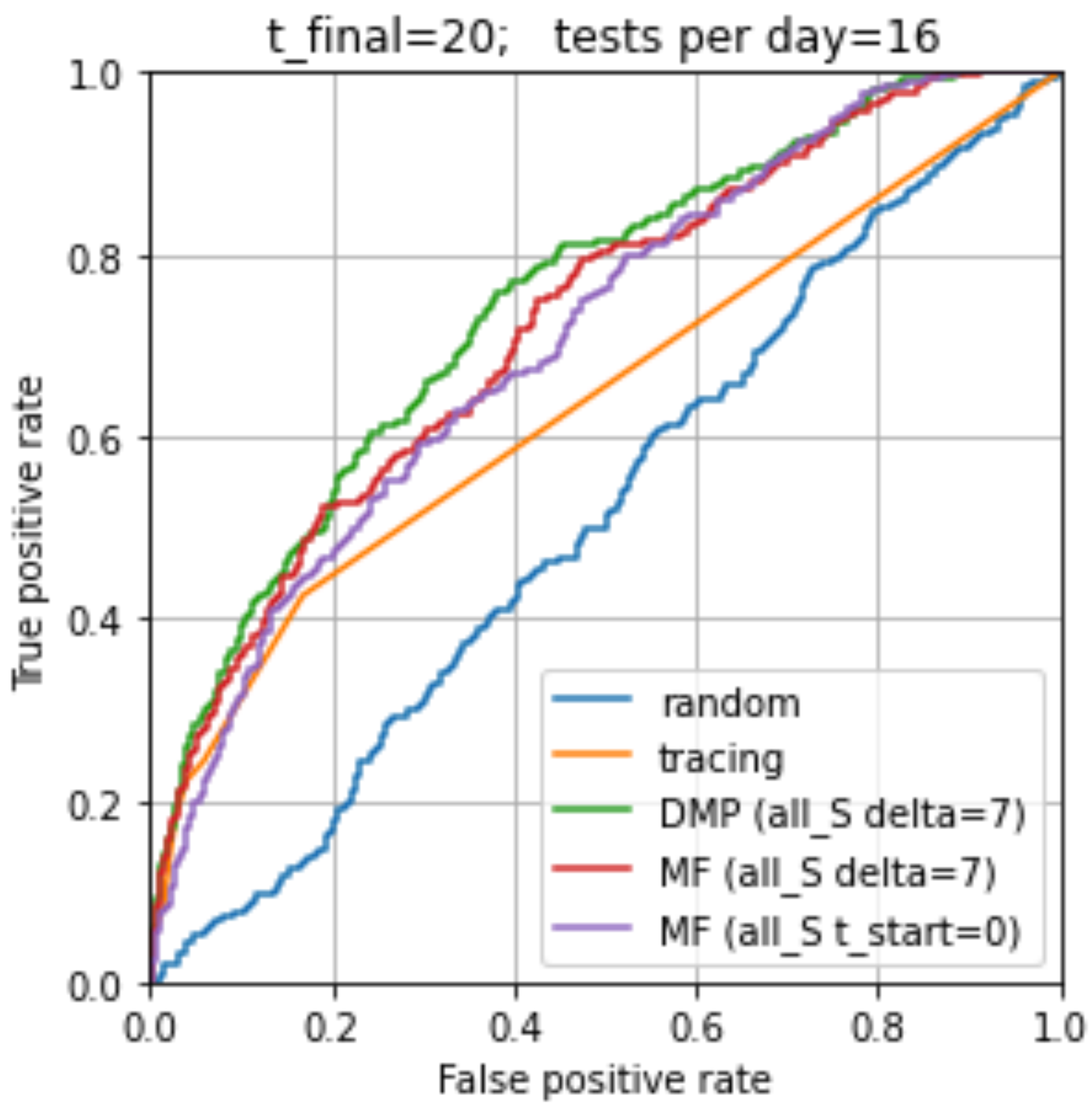
Random geometric graph. Population size= 500, λ=0.02 , μ=0.01 , 2% Infected Individuals at t=0

https://github.com/sphinxteam/sir_inference

https://colab.research.google.com/drive/1QcK2ilzqRiE7bnHqrfxjHhXbwEs_12WU

ROC CURVES (UNCONTROLLED EPIDEMIC, NO INTERVENTIONS)

Contact graph provided by L. Ferretti & R. Hinch, Population size= 10000; 12.7 contacts on average a day,



Epidemic spread and inference with: λ=0.02 , μ=0.07 , 10 Infected Individuals at t=0, τ=7

https://github.com/sphinxteam/sir_inference

https://colab.research.google.com/drive/15qClUFJl_mWTVL6e2VG9mgsAgRQ9Armb

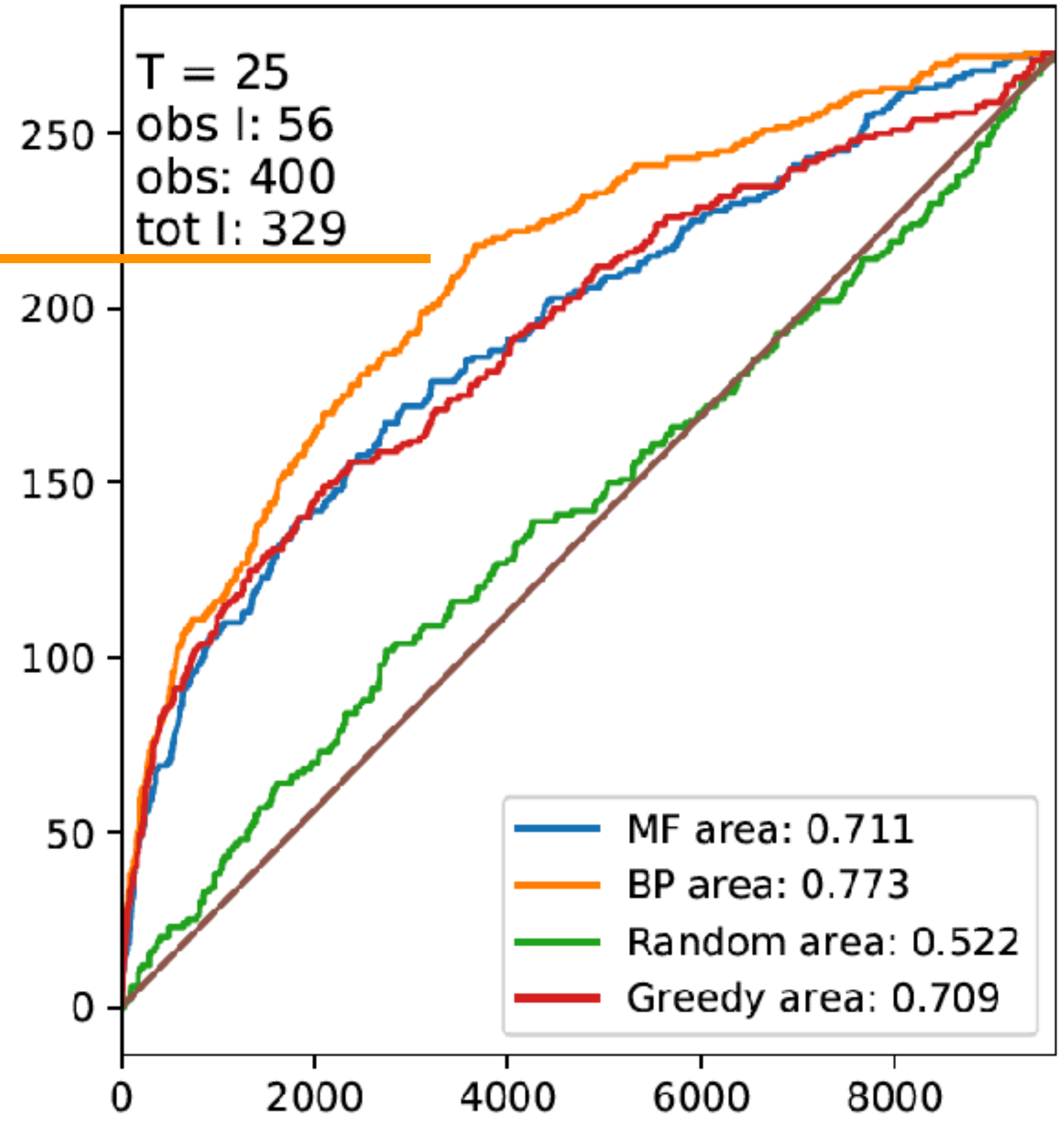
ROC CURVES (UNCONTROLLED EPIDEMIC, NO INTERVENTIONS)

Contact graph provided by L. Ferretti & R. Hinch, Population size= 10000; 12.7 contacts on average a day,

Epidemic: Realistic uncontrolled epidemic spread from Hinch et al. report ($\lambda \sim 0.04$, $\mu \sim 1/12$)

BP on trajectories, Braunstein et al.

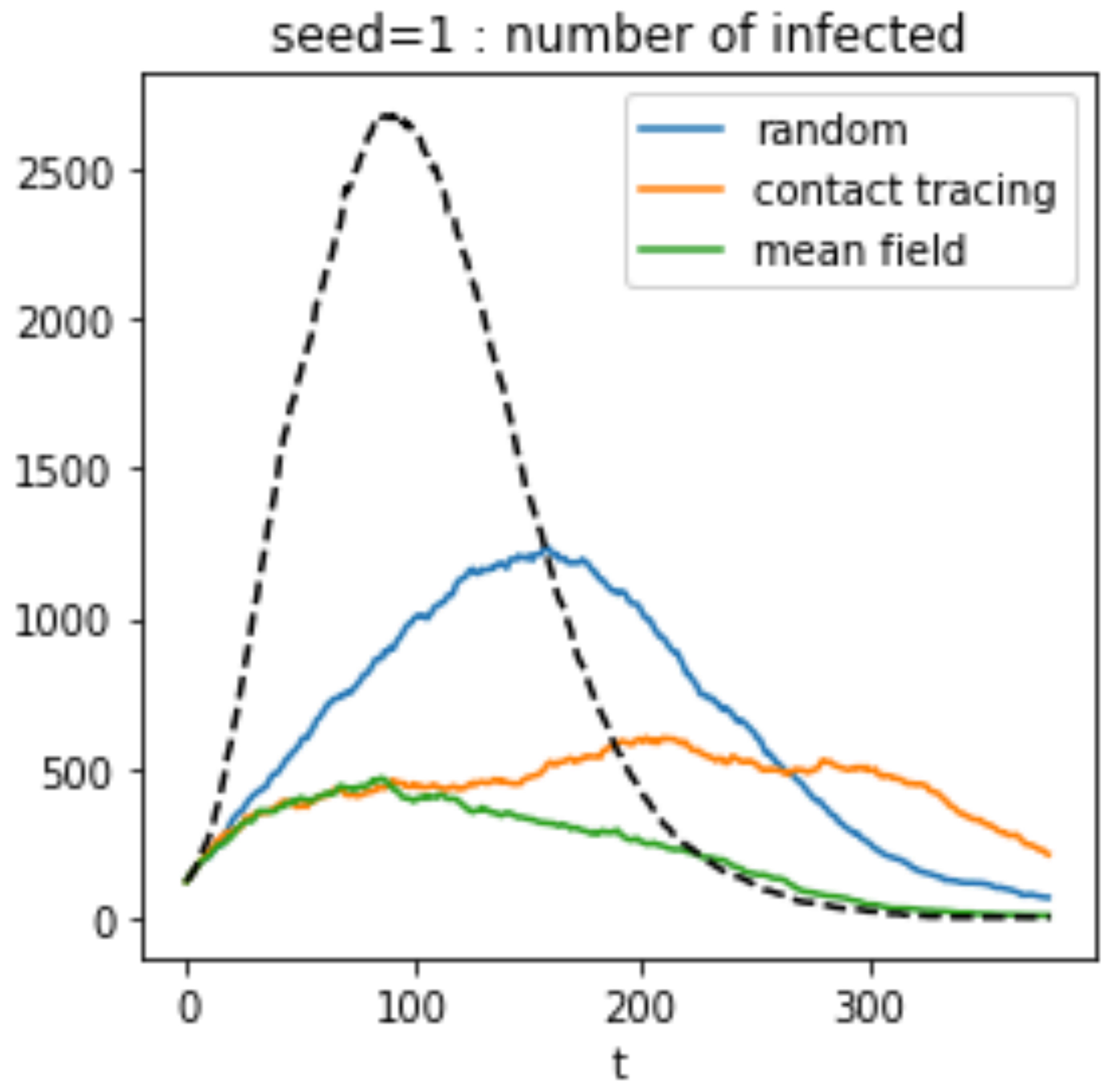
https://github.com/sibyl-team/epibench



MF inference with: λ=0.02 , μ=0.01 , τ=5, δ = 8 (note the mismatch in λ, μ after cross-validation)

https://github.com/sphinxteam/sir_inference

EXPERIMENTS ON CONTROL OF EPIDEMIC

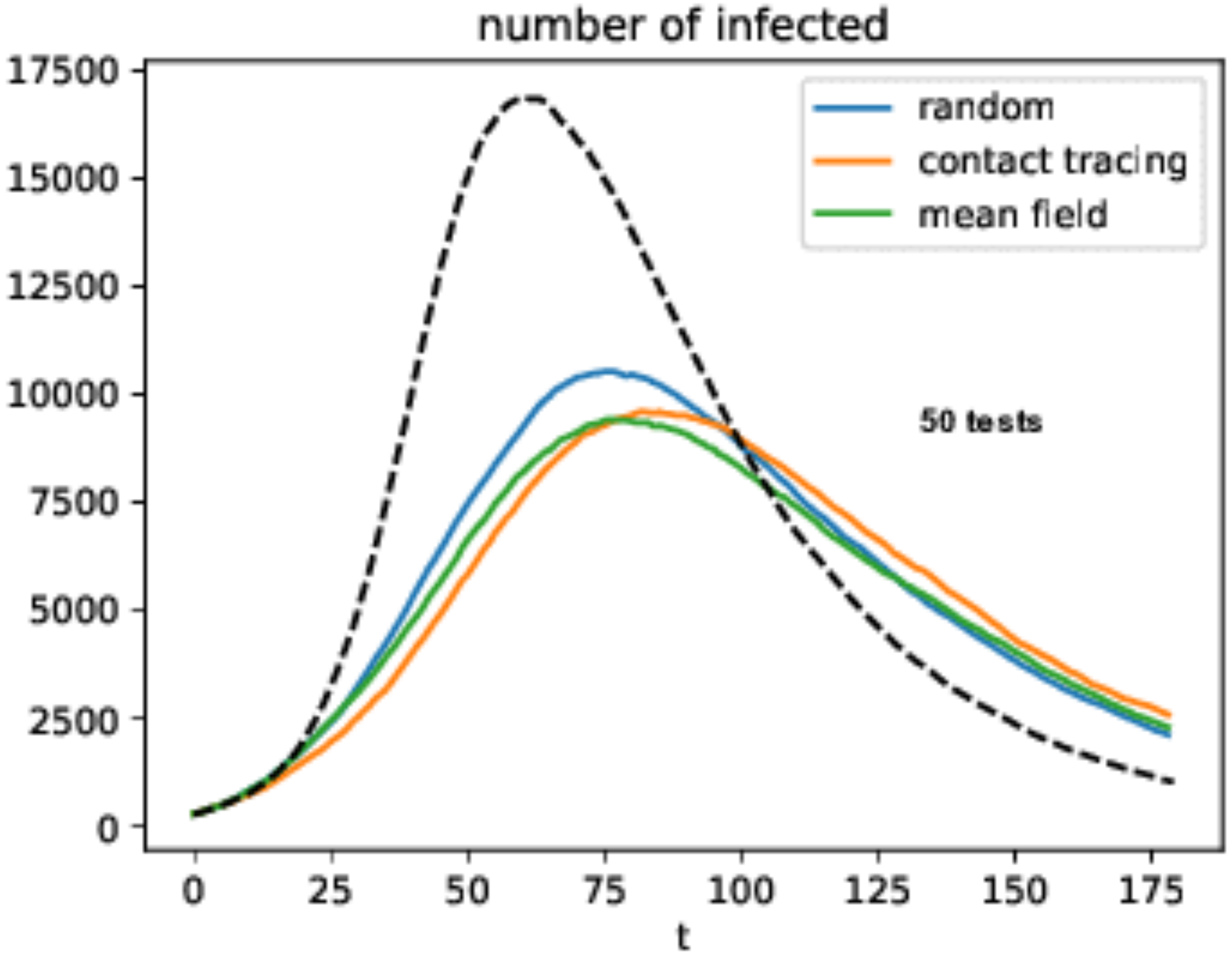


Random geometric contact graph in 2D, scale 1.1, daily on average 7.4 contacts. Population size= 10000, λ=0.02 , μ=0.03. Initially 20 infected + 10 time steps of uncontrolled evolution, τ=5, δ=15.
Tests: symptomatic = 1/2 of all infected, 1 at random, 7 among infected, and 21 from ranking, quarantine = 50 time steps,

https://github.com/sphinxteam/sir_inference

https://colab.research.google.com/drive/1pRq13j8o6Y8GRWa_IDb_Erj-jqyjWDnM#scrollTo=bVi7bEAGgijX

EXPERIMENTS ON CONTROL OF EPIDEMIC



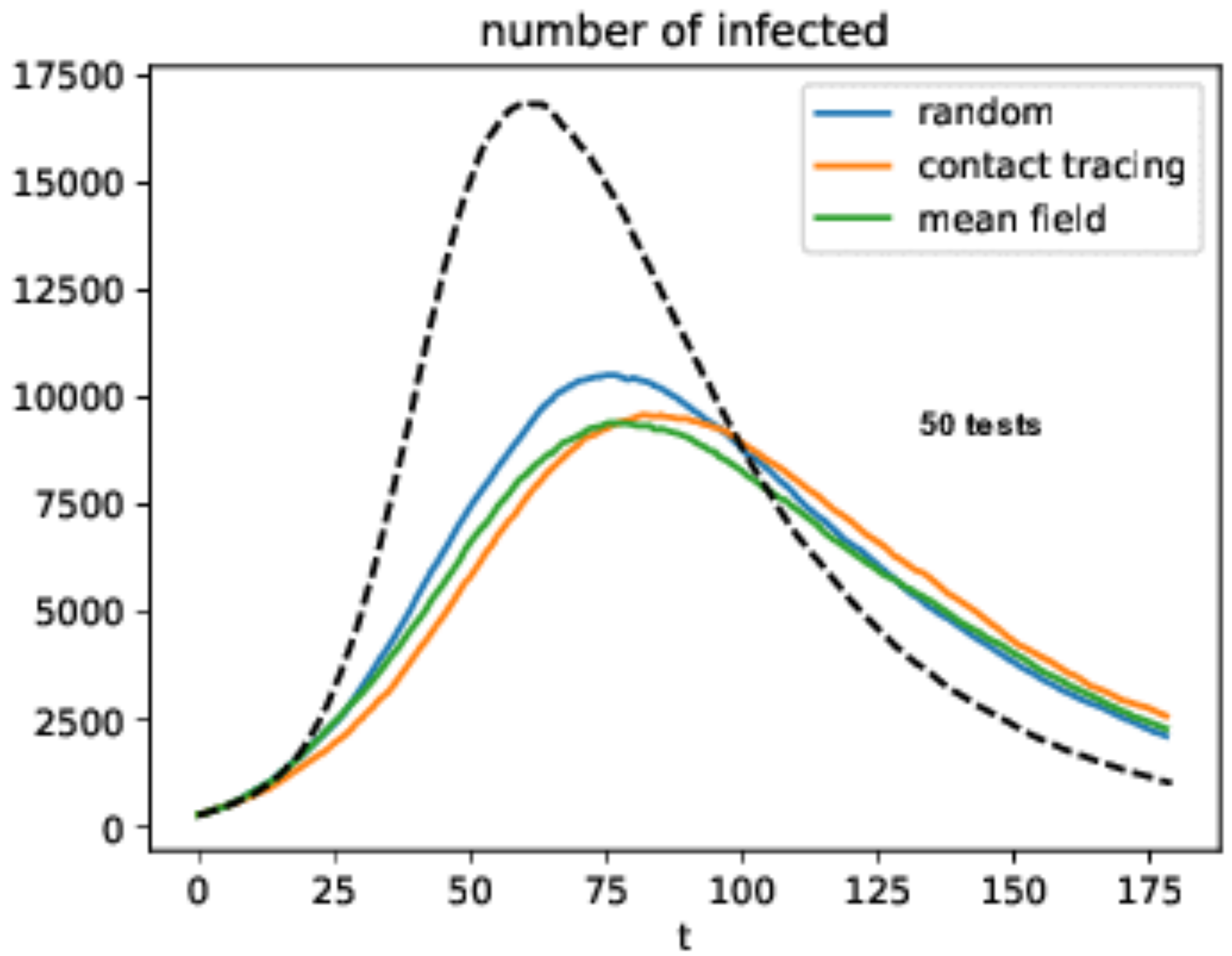
X tests a day + symptomatic

Contact graph from R. Hinch, … L. Ferretti, et al model. Population size= 50000; 11.3 contacts on average a day, λ=0.01, μ=0.033. Initially 50 infected + 20 time steps of uncontrolled evolution, τ=5, δ=15. Tests: symptomatic = 1/2 of all infected, and X from ranking, quarantine = 50 time steps,

https://github.com/sphinxteam/sir_inference

https://colab.research.google.com/drive/1TxwzU59qPBAvk8X-X5Ft21C6nVW7c_pO

EXPERIMENTS ON CONTROL OF EPIDEMIC



X tests a day + symptomatic

Contact graph from R. Hinch, … L. Ferretti, et al model. Population size= 50000; 11.3 contacts on average a day, λ=0.01, μ=0.033. Initially 50 infected + 20 time steps of uncontrolled evolution, τ=5, δ=15. Tests: symptomatic = 1/2 of all infected, and X from ranking, quarantine = 50 time steps,

https://github.com/sphinxteam/sir_inference

https://colab.research.google.com/drive/1TxwzU59qPBAvk8X-X5Ft21C6nVW7c_pO

ONGOING WORK:

- Epidemic control with BP SIPAR strategy of Braunstein et al.
- Explore the role of the various parameters.
- Larger networks, realistic epidemic spread (<1% infected, millions of nodes),
- Include more info from the simulator (age groups, households, details about syndromes, …)
- Learn parameters with expectation maximisation from observed data.

WHAT'S NEXT?

- Implementation respecting privacy constraints (limited bandwidth of communication, privacy consistent user identification, etc.)
- Learn a generative model with graph neural networks? In progress in MILA group. Max Welling's call-for-interest on Facebook.

Our progress can be followed at: https://github.com/sphinxteam/sir_inference/
with notes on the methods: https://www.overleaf.com/read/tfhcpbvhmcwq

We are interested to collaborate.

WHAT'S NEXT?

- Implementation respecting privacy constraints (limited bandwidth of communication, privacy consistent user identification, etc.)
- Learn a generative model with graph neural networks? In progress in MILA group. Max Welling's call-for-interest on Facebook.

Our progress can be followed at: https://github.com/sphinxteam/sir_inference/
with notes on the methods: https://www.overleaf.com/read/tfhcpbvhmcwq

We are interested to collaborate.