



# Epidemic modeling via cross-entropy approximate Bayesian computation

Americo Cunha Jr

Rio de Janeiro State University — UERJ

IV DYCAELS

Ponta Grossa - PR, Brazil, November 8, 2023



@AmericoCunhaJr

@AmericoCunhaJr

@AmericoCunhaJr

# Acknowledgments

## ► Collaborators



David A. W. Barton



Thiago G. Ritto

## ► Institutional and financial support



University of  
BRISTOL



Conselho Nacional de Desenvolvimento  
Científico e Tecnológico



CAPES



Fundação Carlos Chagas Filho de Amparo  
à Pesquisa do Estado do Rio de Janeiro

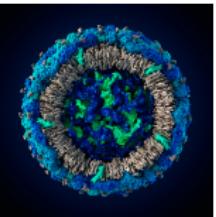


Engineering and Physical Sciences  
Research Council

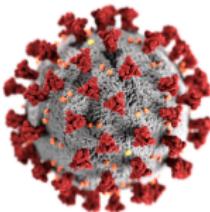
# Epidemics are recurrent in history ...



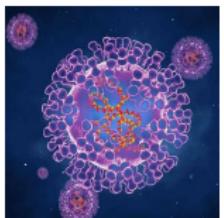
Dengue



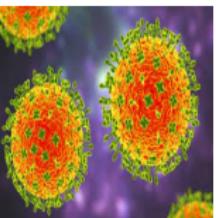
Zika virus



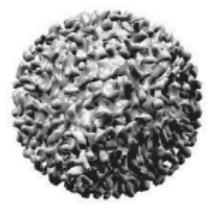
COVID-19



Monkeypox

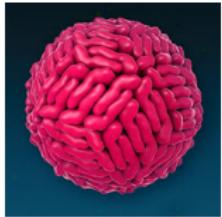


Langya

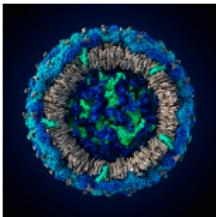


Next ?

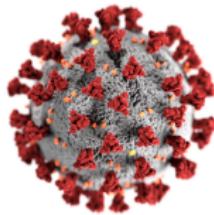
# Epidemics are recurrent in history ...



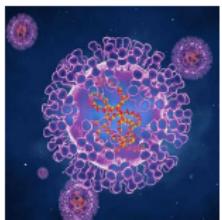
Dengue



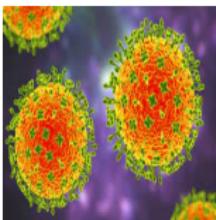
Zika virus



COVID-19



Monkeypox



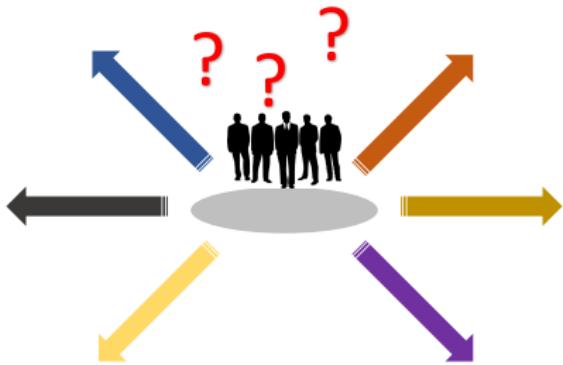
Langya



Next ?

## They may cause a lot of trouble!

# Tools to support for decision making



Computational Model + Real World Data = Tool to aid with Decision Making

# Mathematical models are approximations!



*"All models are wrong but some are useful"*

*George E. P. Box*

# Mathematical models are approximations!



*"All models are wrong but some are useful"*

*George E. P. Box*

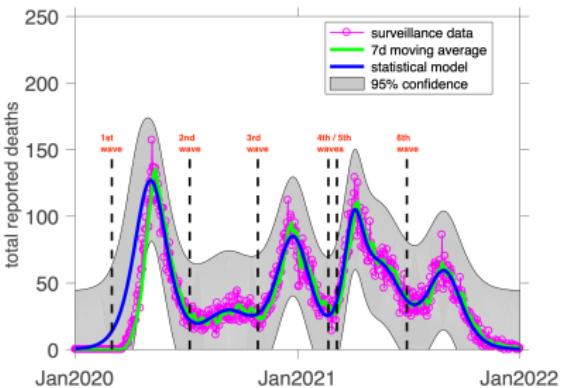


This idea has a more pronounced  
meaning in computational epidemiology than in physics!  
“First principles of epidemiology are unknown!”

# For deaths records an algebraic model is effective!



$$I(t) = \sum_{i=1}^N \frac{r_i K_i e^{-r_i(t-\tau_i)}}{\left(1 + e^{-r_i(t-\tau_i)}\right)^2}$$



Chaos ARTICLE citation.org/journals/chaos

## The starting dates of COVID-19 multiple waves

Cite as: Chaos **32**, 031101 (2022); doi:10.1063/5.0079904  
Submitted: 25 November 2021; Accepted: 9 February 2022;  
Published Online: 3 March 2022

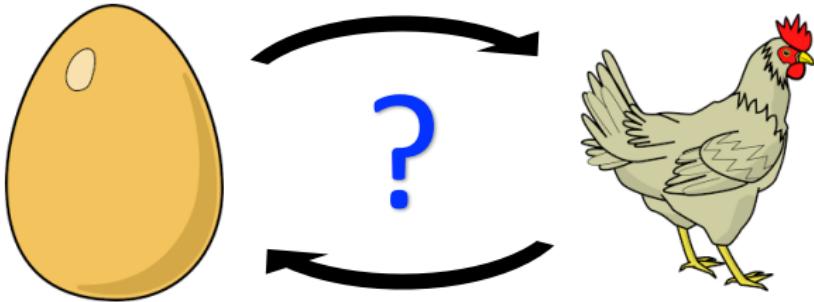
Paulo Roberto de Lima Gianfelice,<sup>1</sup> Ricardo Sovyk Oyarzabal,<sup>1</sup> Americo Cunha, Jr.,<sup>2,3,4</sup> Jose Mario Vicensi Grzybowski,<sup>5</sup> Fernando da Conceição Batista,<sup>6</sup> and Elbert E. N. Macau<sup>1</sup>



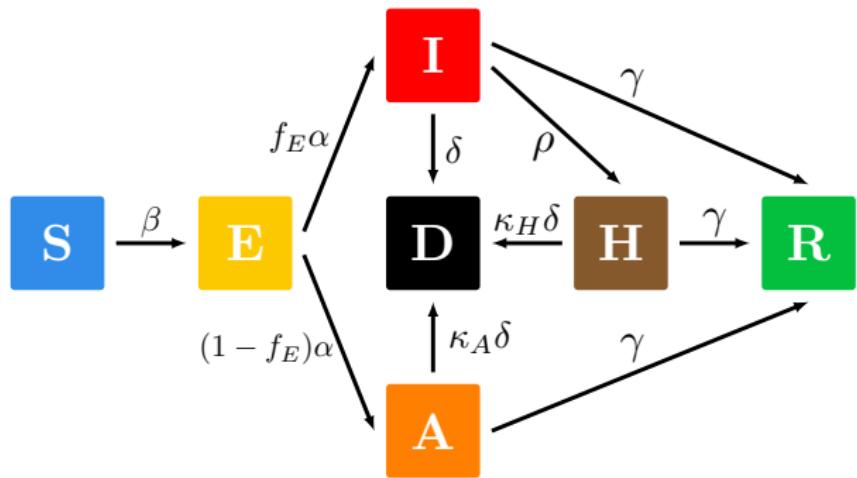
P. R. L. Gianfelice et al., *The starting dates of COVID-19 multiple waves*, Chaos, 32:031101, 2022

But when latent variables are of interest ?

Mechanistic models are necessary!  
(cause-effect relationship)



# SEIR(+AHD) mechanistic epidemic model



# SEIR(+AHD) $_{\beta}$ epidemic model

## ► Dynamic model:

$$\begin{aligned}\dot{S} &= -\beta(t) S(I + A + \epsilon_H H)/N \\ \dot{E} &= \beta(t) S(I + A + \epsilon_H H)/N - \alpha E \\ \dot{I} &= f_E \alpha E - (\gamma + \rho + \delta) I \\ \dot{R} &= \gamma (I + A + H) \\ \dot{A} &= (1 - f_E) \alpha E - (\kappa_A \delta + \gamma) \\ \dot{H} &= \rho I - (\gamma + \kappa_H \delta) H \\ \dot{D} &= \delta (I + \kappa_A A + \kappa_H H) \\ \dot{N} &= -\dot{D}\end{aligned}$$

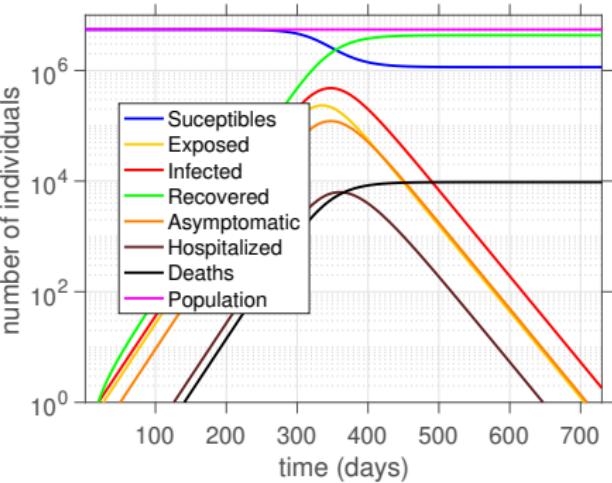
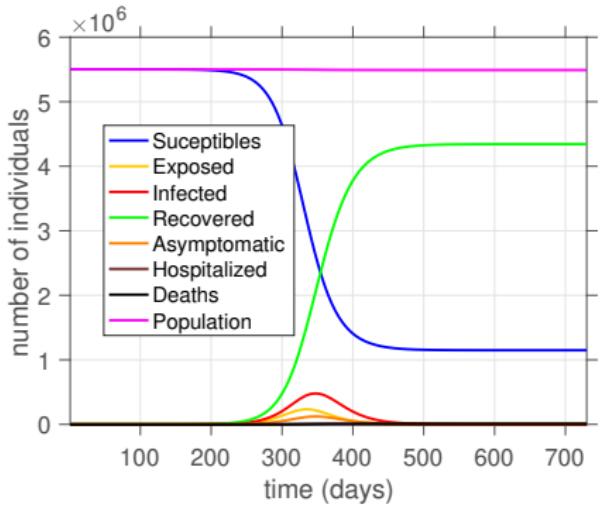
+ initial conditions

## ► Rate of transmission:

$$\beta(t) = \beta_0 + \frac{(\beta_{\infty} - \beta_0)}{2} \left( 1 + \tanh \left( \eta \frac{(t - t_{\beta})}{2} \right) \right)$$

# SEIR(+AHD) model response (virgin population)

In this example the transmission rate is fixed, i.e.,  $\beta(t) = \beta$



# Abstraction of the epidemic model

- Dynamic model:

$$\frac{d}{dt} \mathbf{u}(t) = F(t, \mathbf{u}(t), \mathbf{x})$$

- Vector field:

$$(t, \mathbf{u}(t), \mathbf{x}) \in \mathbb{R} \times \mathbb{R}^8 \times \mathbb{R}^{12} \mapsto F(t, \mathbf{u}(t), \mathbf{x}) \in \mathbb{R}^8$$

- State vector:

$$\mathbf{u}(t) = (S, E, I, A, H, R, D, N)$$

- Parameters vector:

$$\mathbf{x} = (\beta_0, \alpha, f_E, \gamma, \rho, \delta, \kappa_A, \kappa_H, \epsilon_H, \beta_\infty, \eta, t_\beta)$$

- QoI vector:

$$\mathbf{y} = [H(t_1), \dots, H(t_n), D(t_1), \dots, D(t_n)]$$

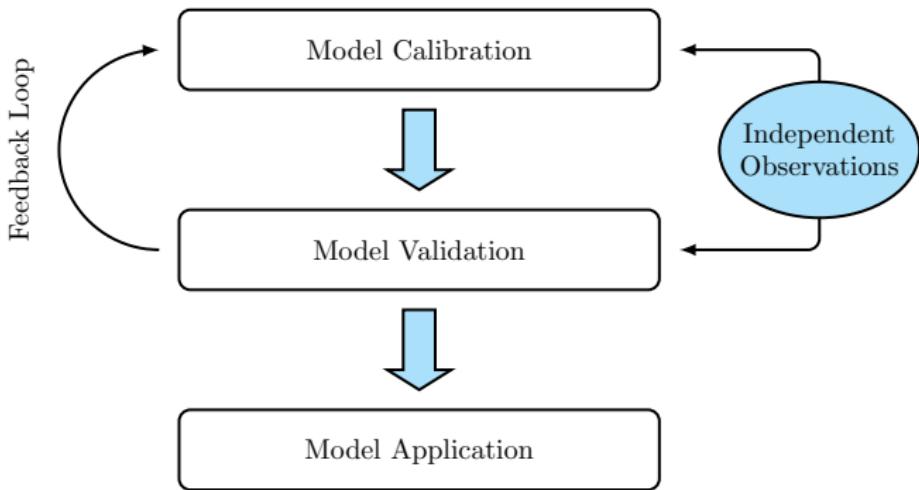
- Computational model:

$$\mathbf{y} = \mathcal{M}(\mathbf{x})$$

**Model Parametrization:** expert knowledge, literature, ansatz, etc

**Model Calibration:** data + model inversion

**Model Validation:** other data + error metrics



# How to obtain consistent initial conditions ?



**Initial conditions (almost surely) are unknown for epidemic systems!**

# How to obtain consistent initial conditions ?



**Initial conditions (almost surely) are unknown for epidemic systems!**

Ansatz                      →     biased and inconsistent

# How to obtain consistent initial conditions ?

**Initial conditions (almost surely) are unknown for epidemic systems!**

Ansatz → biased and inconsistent

Inversion → hard and expensive

# How to obtain consistent initial conditions ?

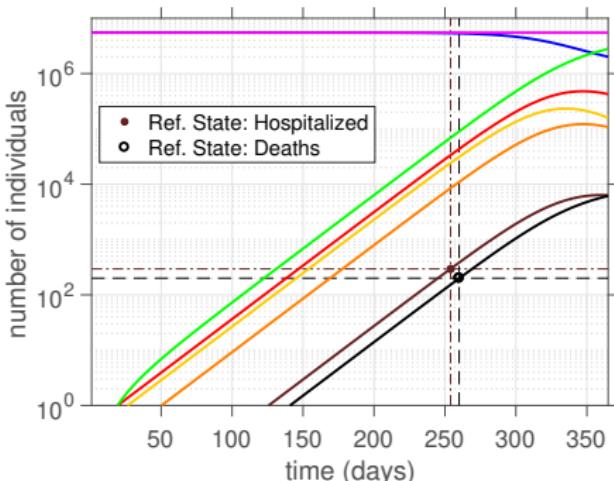
**Initial conditions (almost surely) are unknown for epidemic systems!**

Ansatz → biased and inconsistent

Inversion → hard and expensive

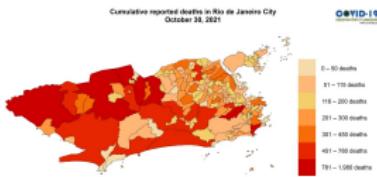
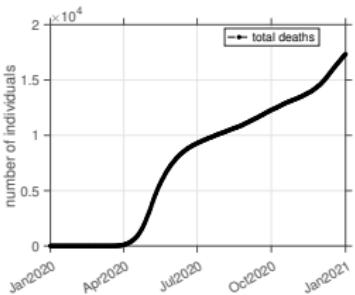
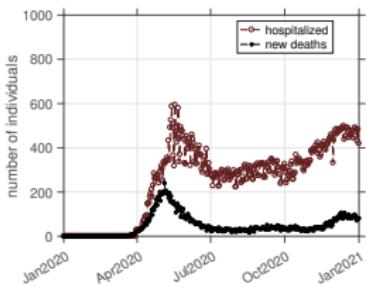
Dynamic inspired →  $\mathbf{u}(0) = \omega \mathbf{u}(0)_{ref}^H + (1 - \omega) \mathbf{u}(0)_{ref}^D$

"convex combination of admissible reference states"



# How to quantify model response and data similarity ?

- ▶ Surveillance data (Rio de Janeiro city):



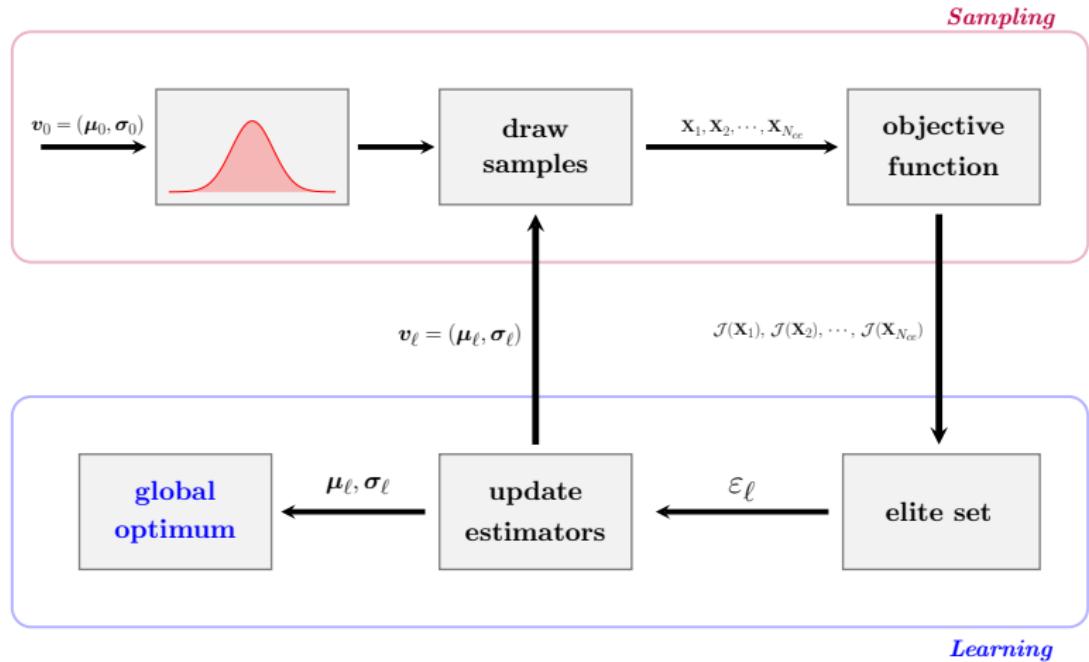
- ▶ Data partition:

$$\mathbf{y}_{\text{data}} = [\mathbf{y}_{\text{data}}^H \quad \mathbf{y}_{\text{data}}^D]$$

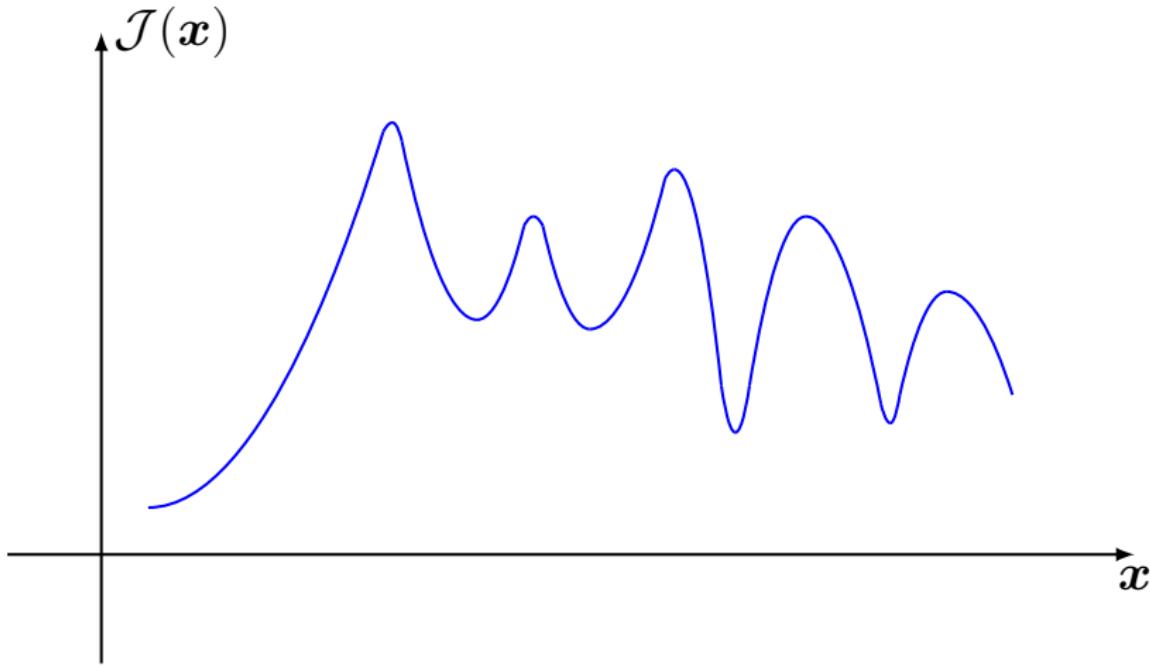
- ▶ Misfit function:

$$\mathcal{J}(\mathbf{x}) = \omega \frac{\|\mathbf{y}_{\text{data}}^H - \mathbf{y}^H(\mathbf{x})\|^2}{\|\mathbf{y}_{\text{data}}^H\|^2} + (1 - \omega) \frac{\|\mathbf{y}_{\text{data}}^D - \mathbf{y}^D(\mathbf{x})\|^2}{\|\mathbf{y}_{\text{data}}^D\|^2}$$

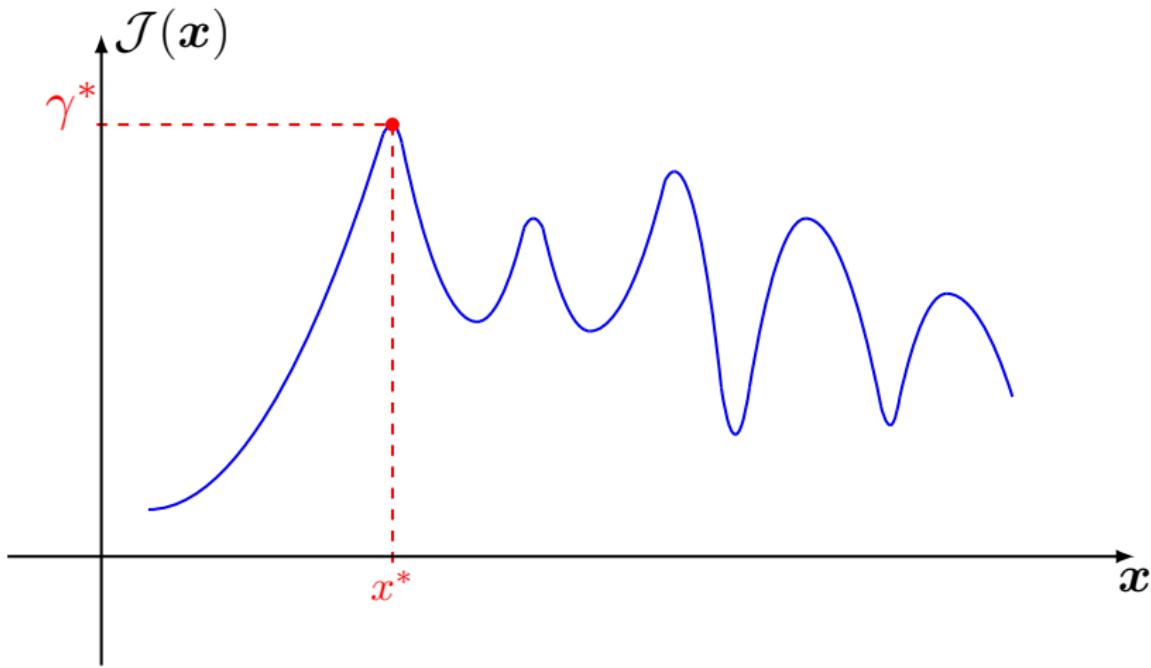
# Non-convex optimization via Cross-Entropy method



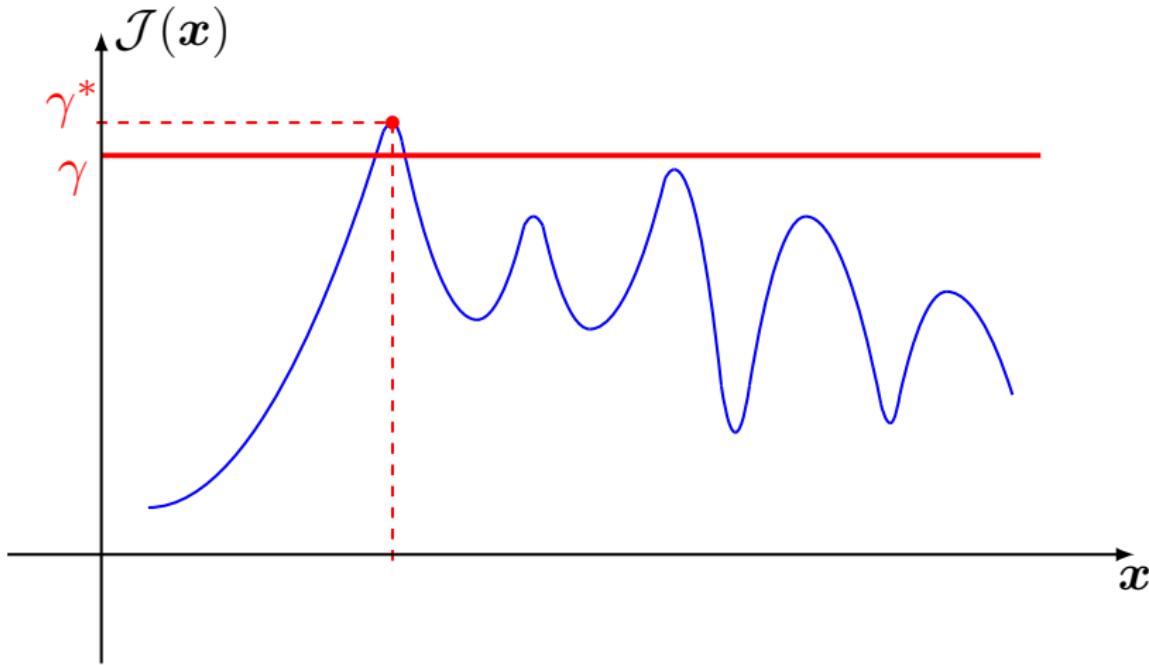
# Cross-entropy method in action!



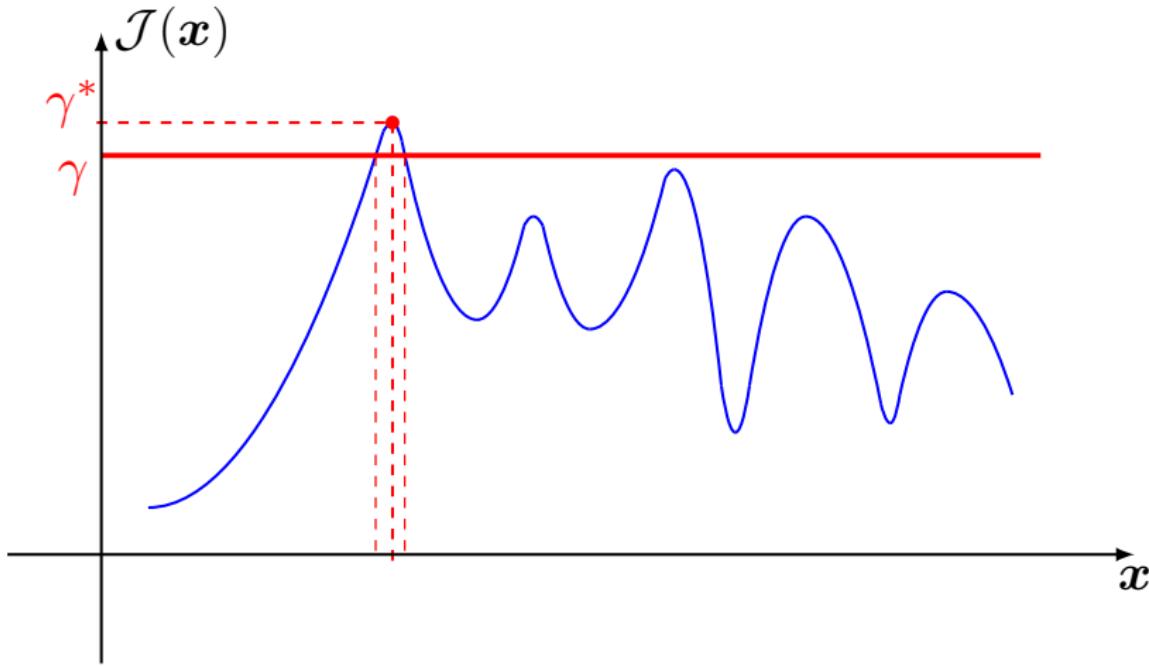
# Cross-entropy method in action!



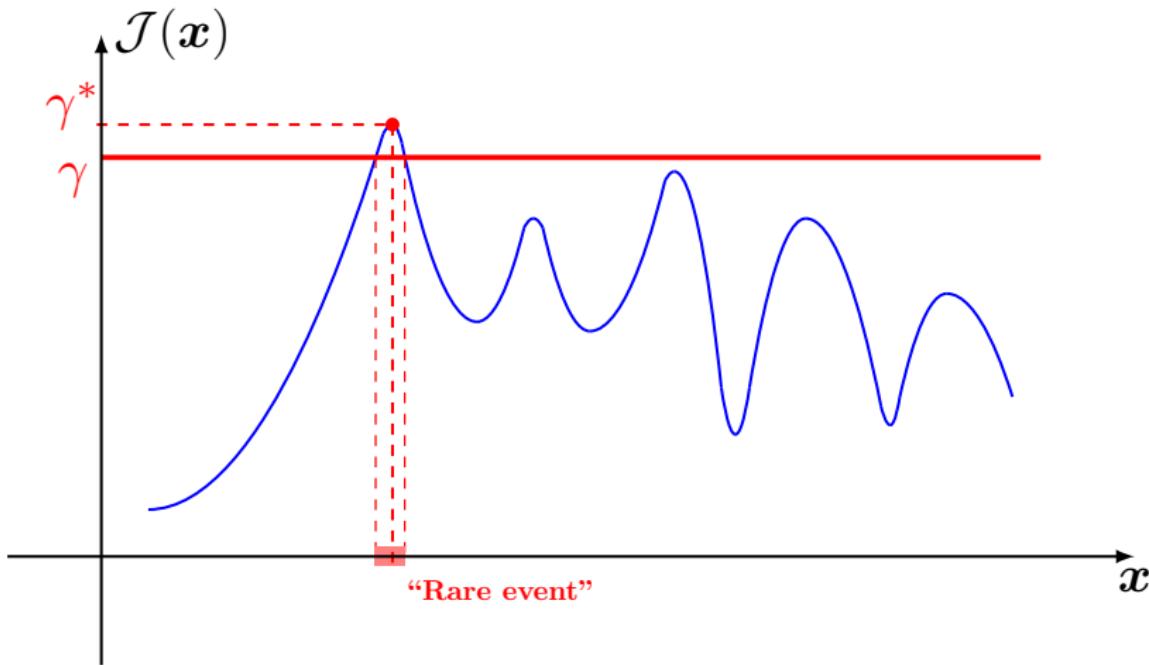
# Cross-entropy method in action!



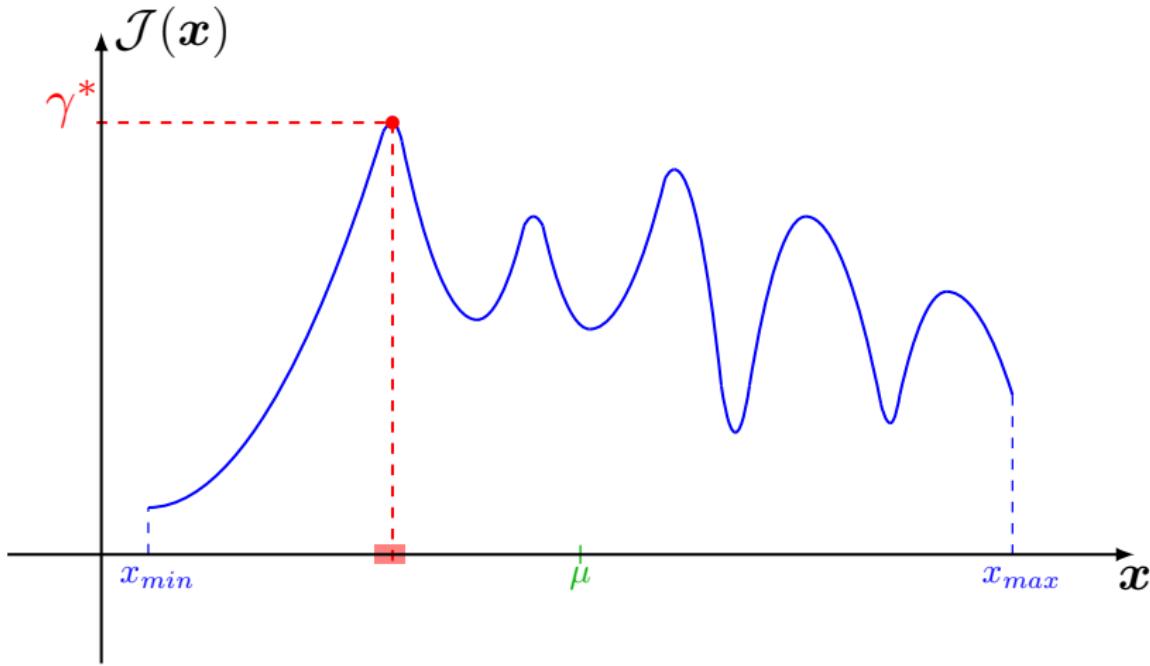
# Cross-entropy method in action!



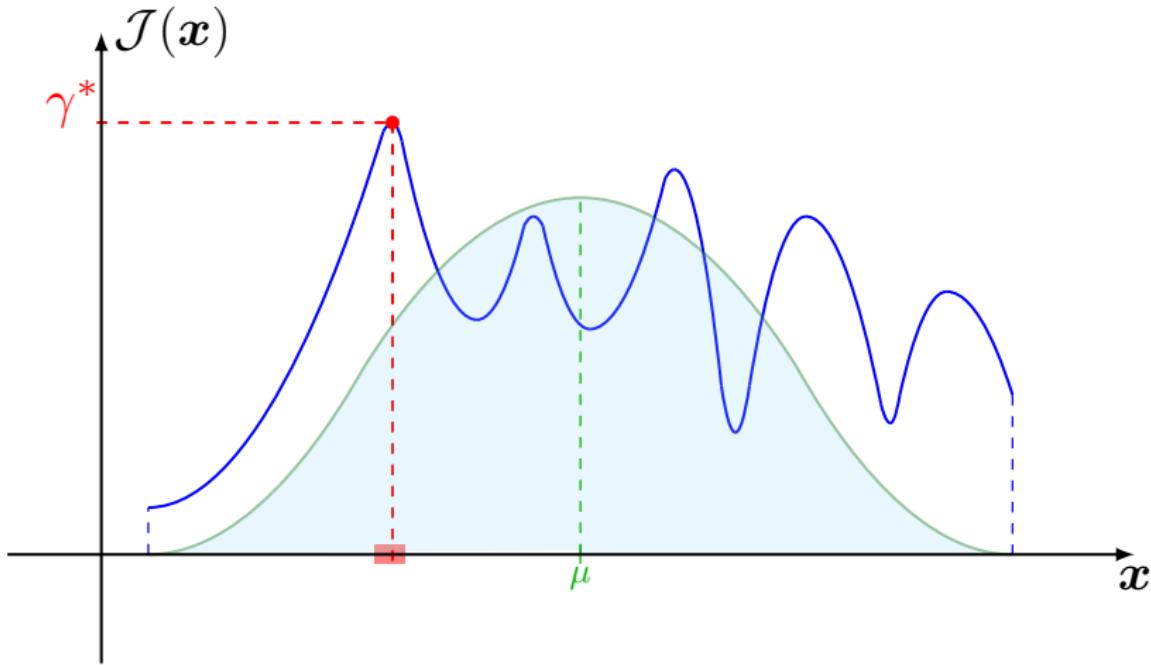
# Cross-entropy method in action!



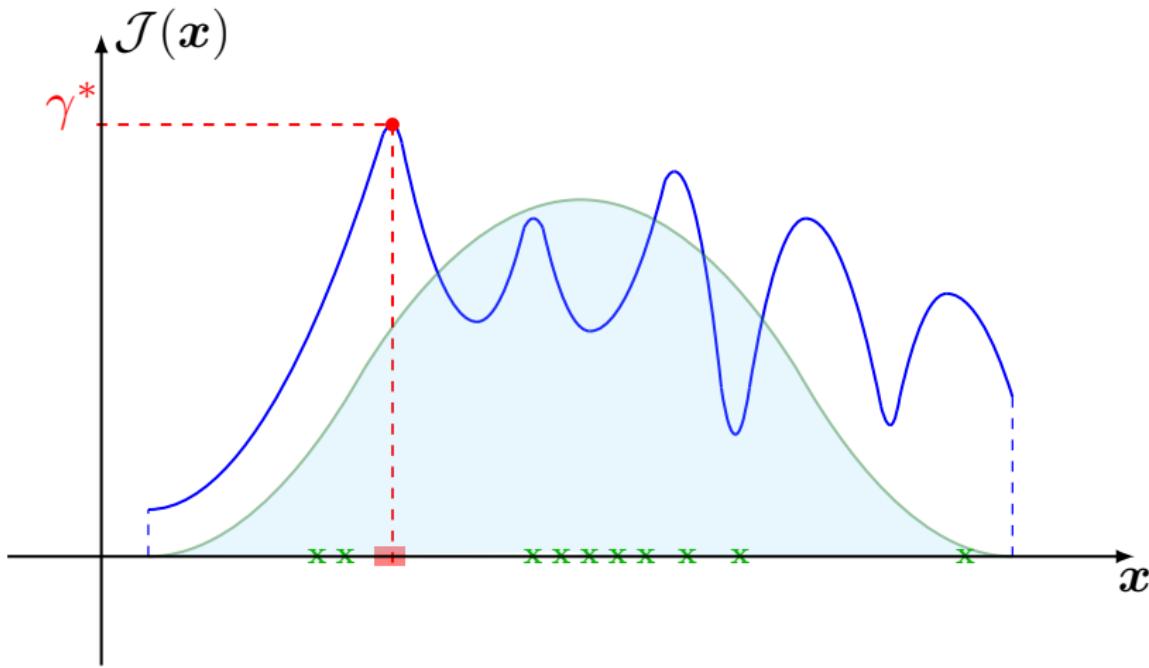
# Cross-entropy method in action!



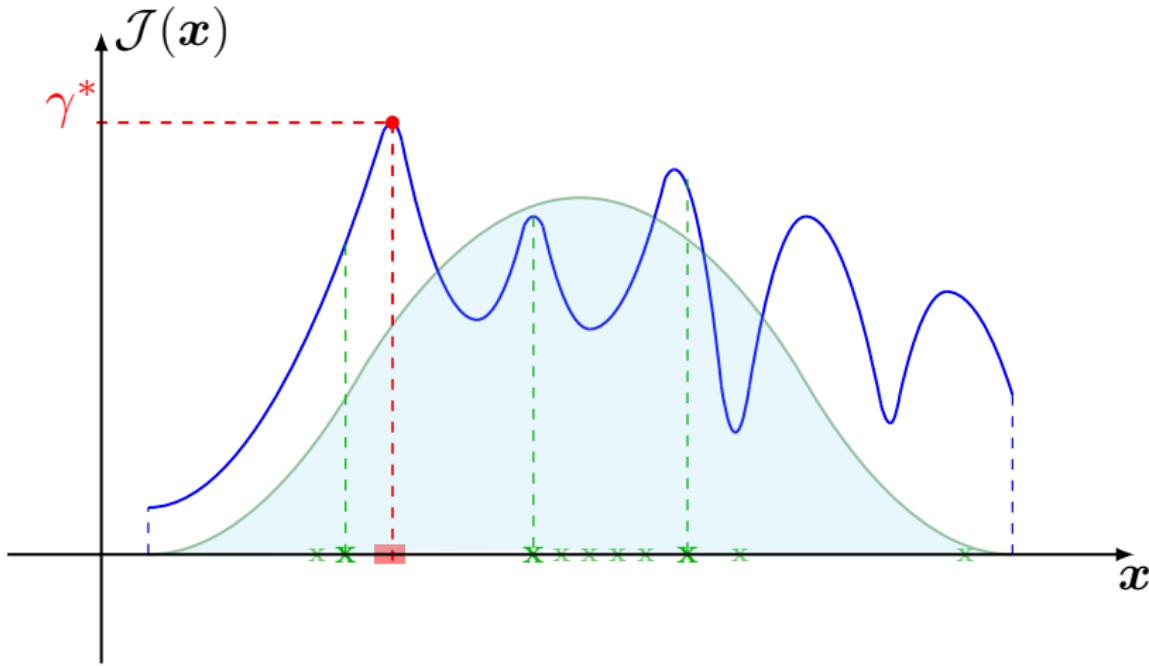
# Cross-entropy method in action!



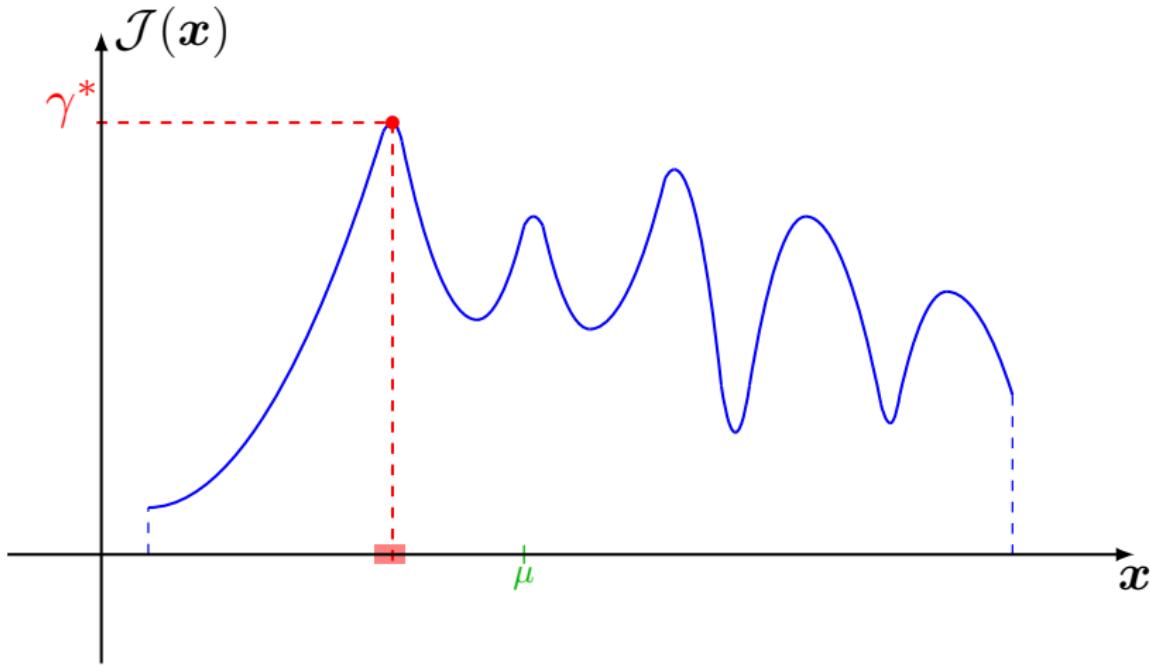
# Cross-entropy method in action!



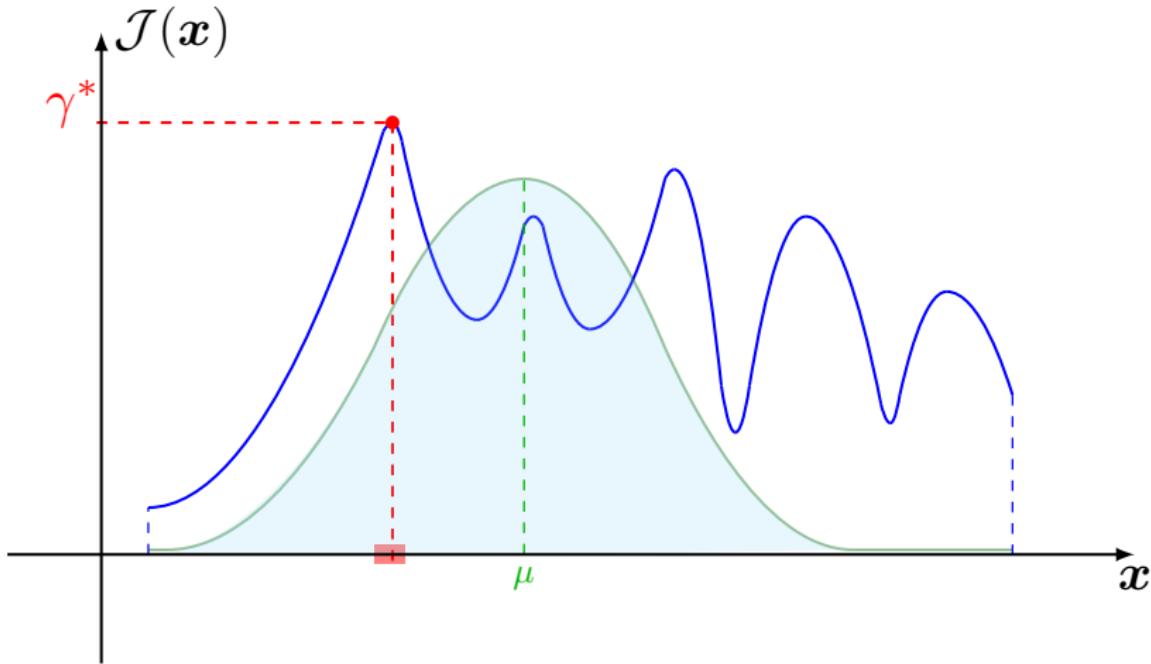
# Cross-entropy method in action!



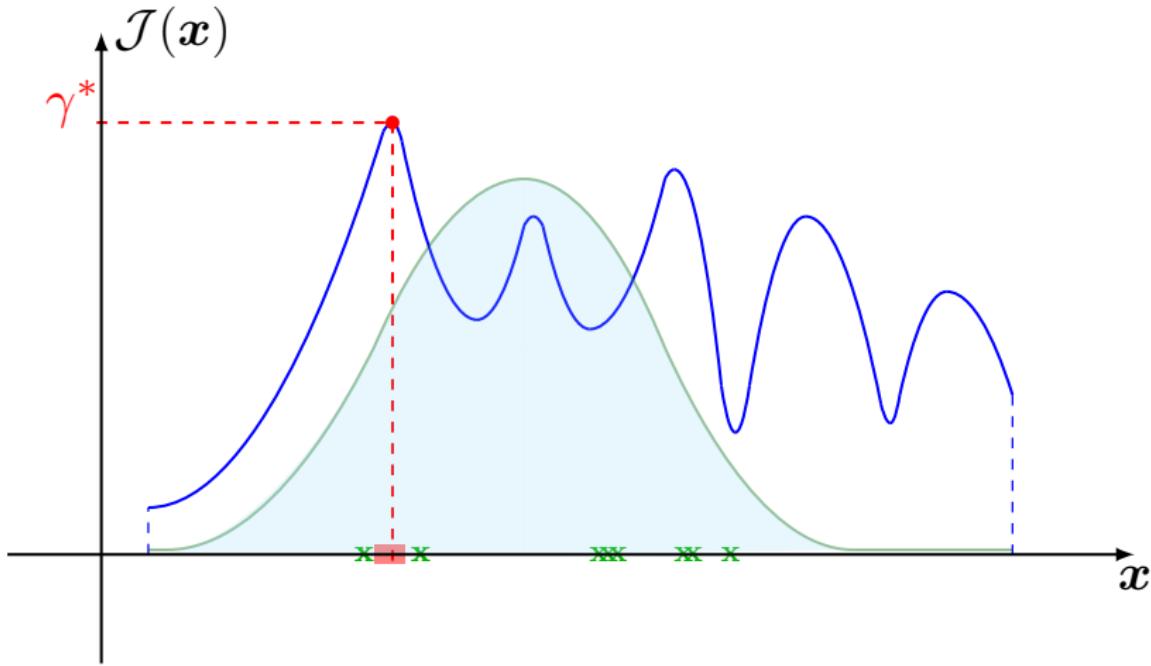
# Cross-entropy method in action!



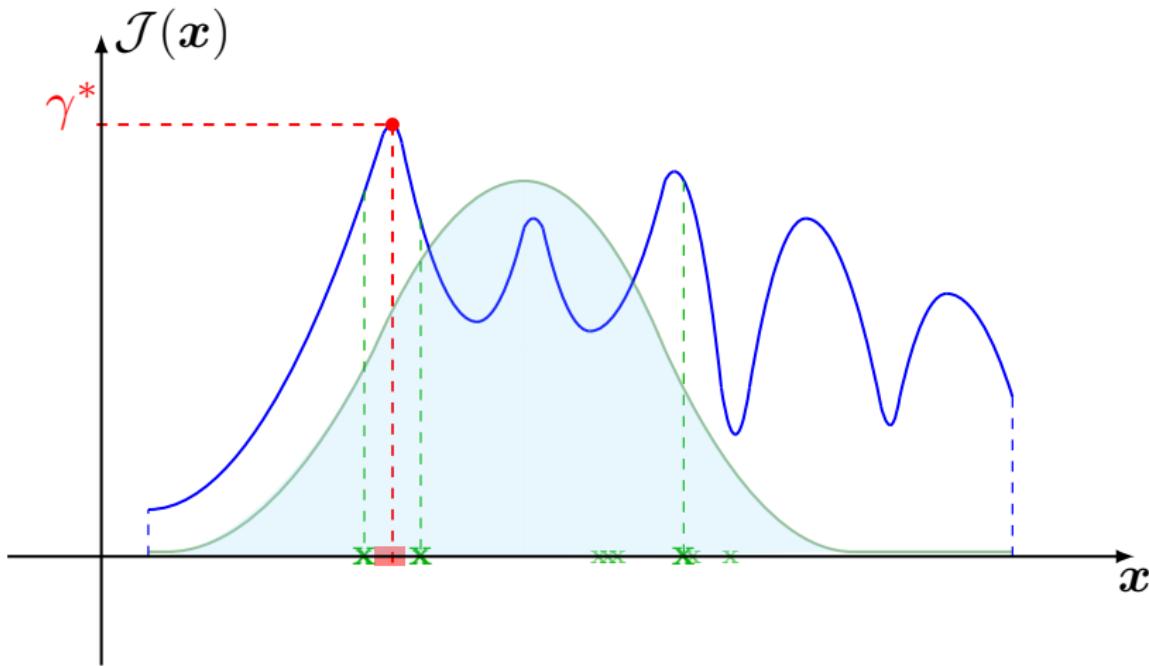
# Cross-entropy method in action!



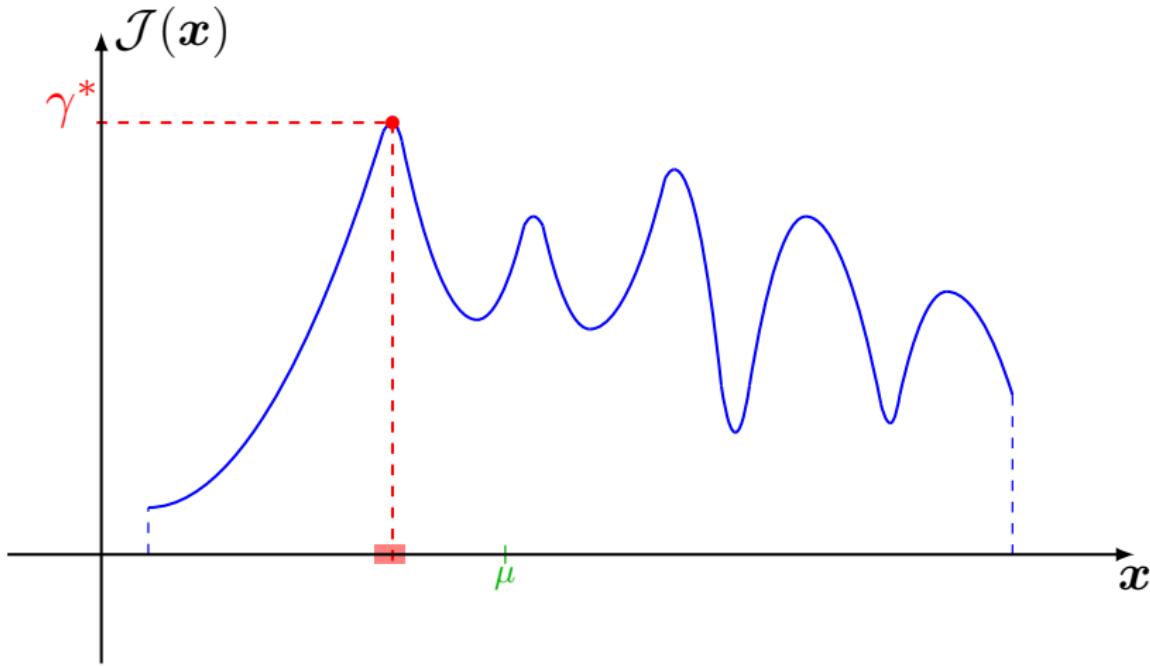
# Cross-entropy method in action!



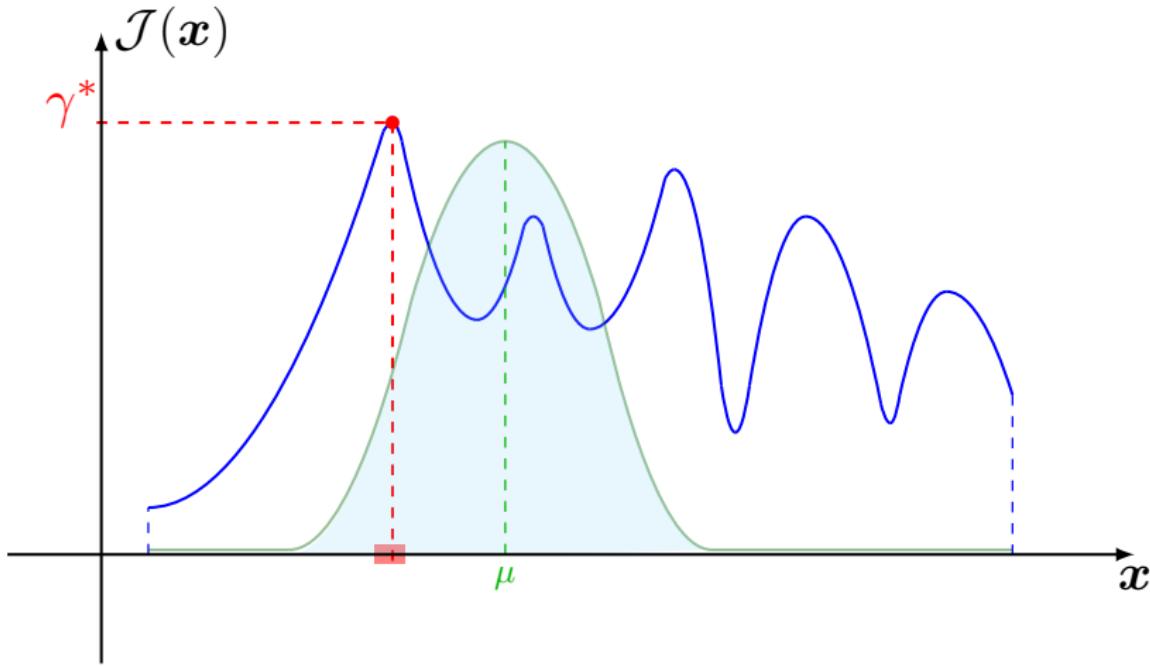
# Cross-entropy method in action!



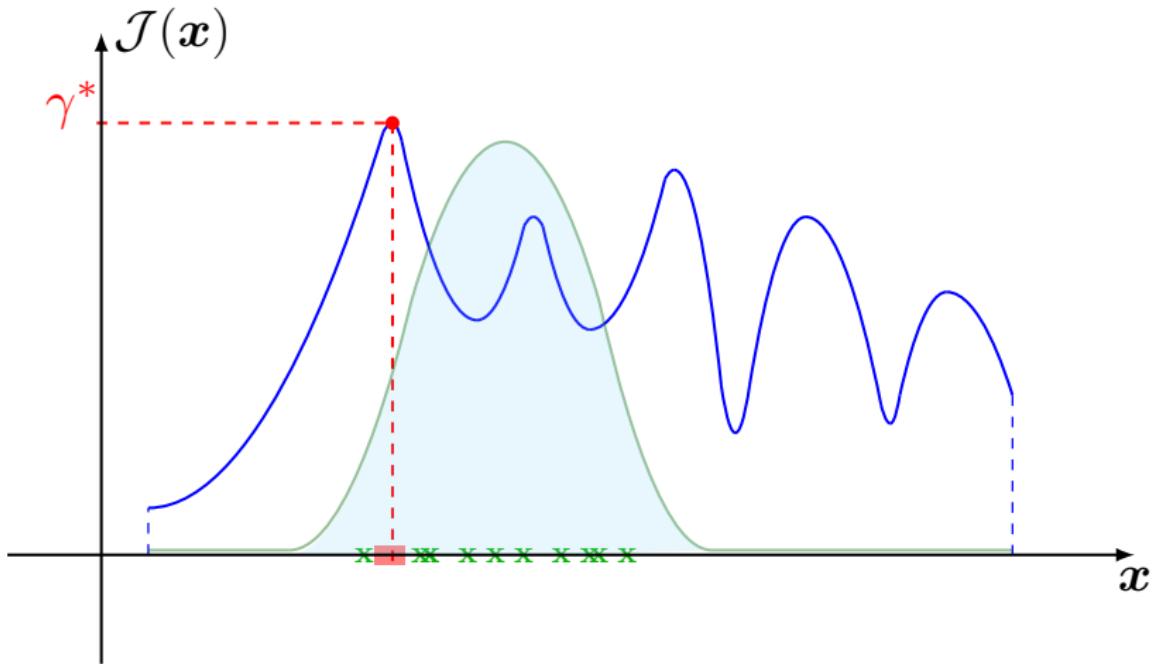
# Cross-entropy method in action!



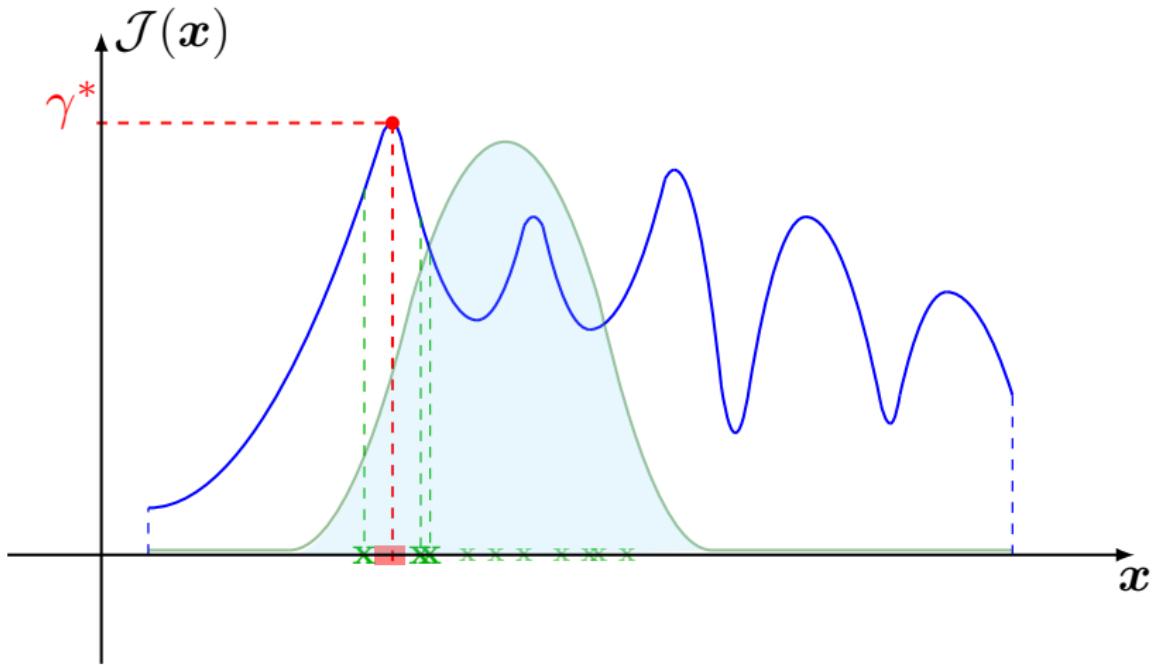
# Cross-entropy method in action!



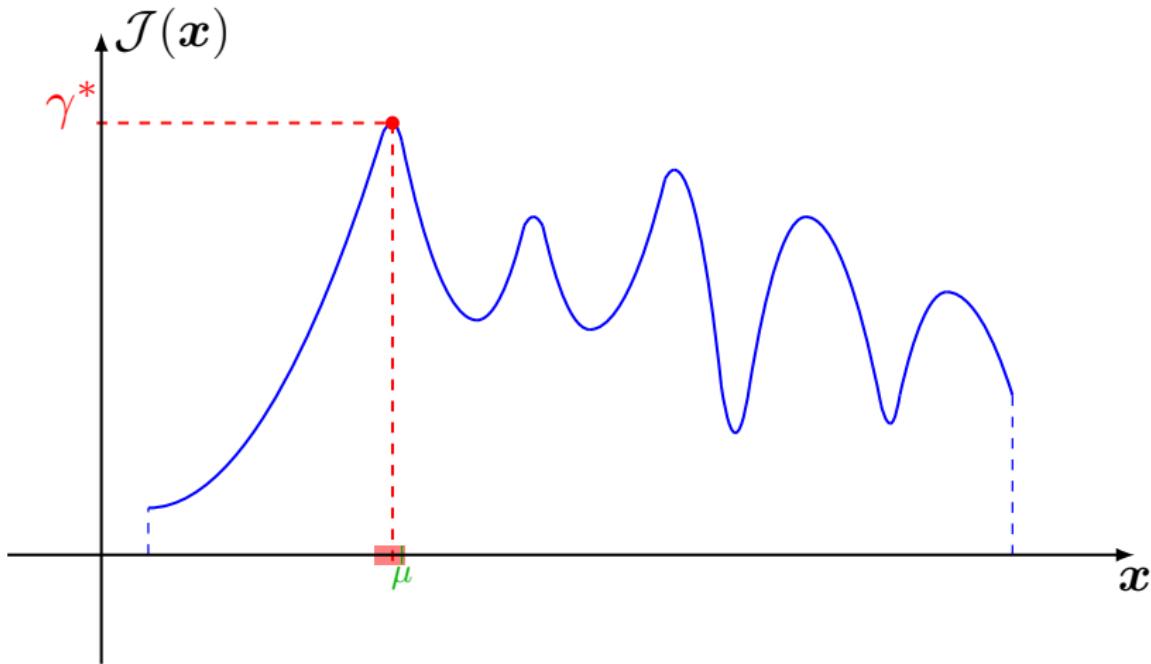
# Cross-entropy method in action!



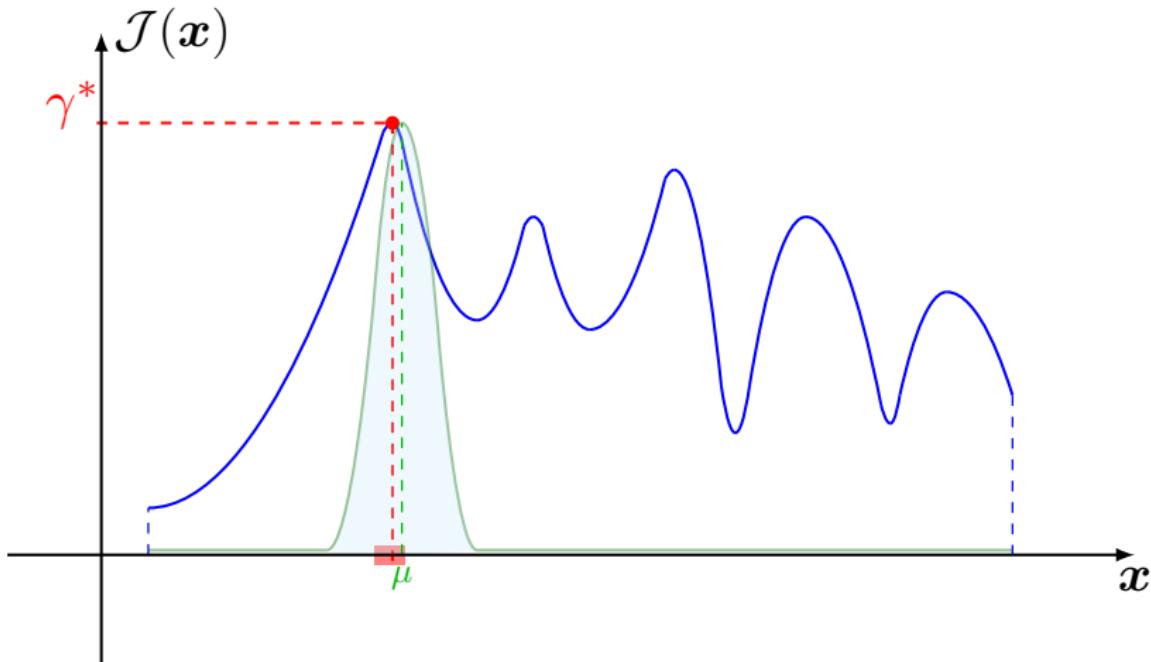
# Cross-entropy method in action!



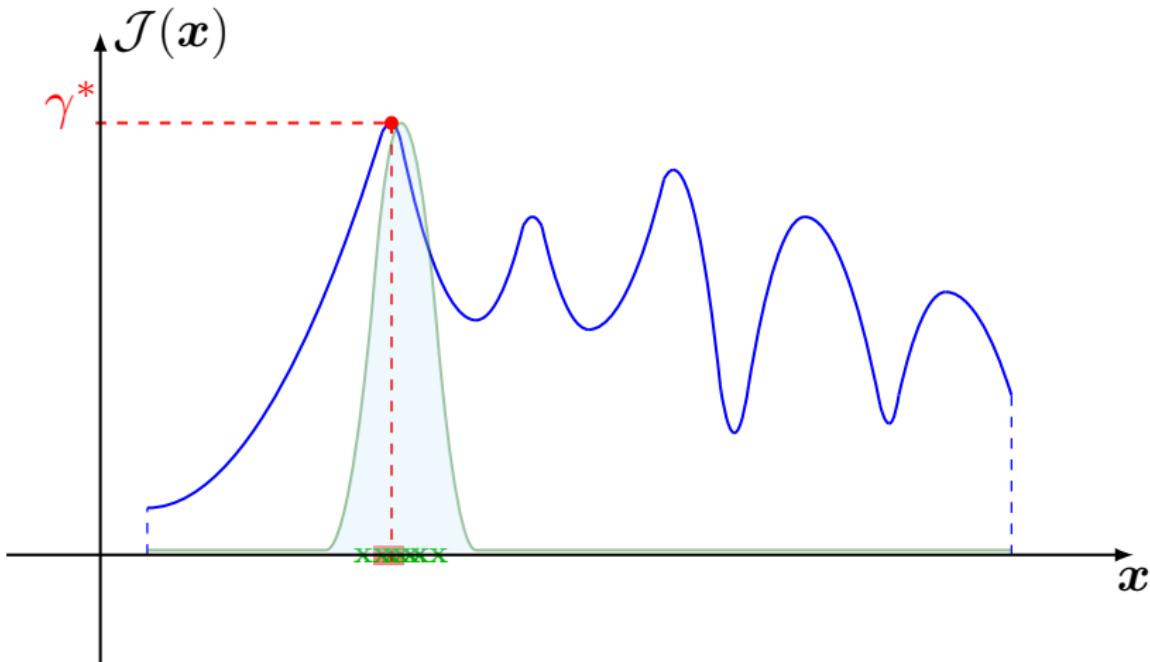
# Cross-entropy method in action!



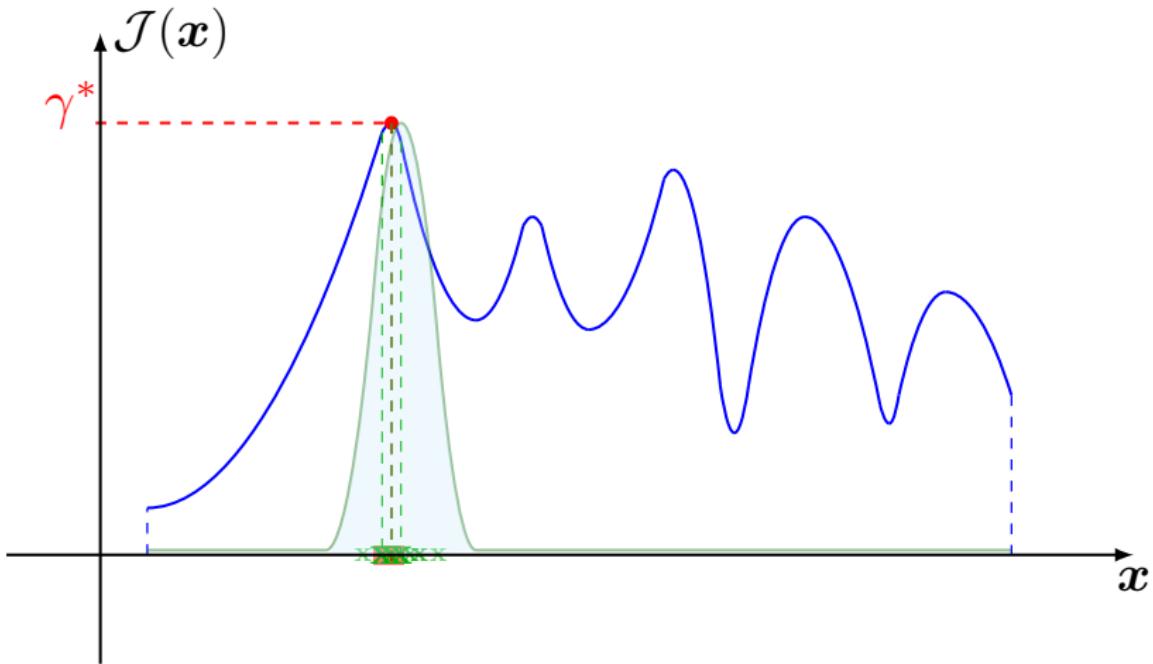
# Cross-entropy method in action!



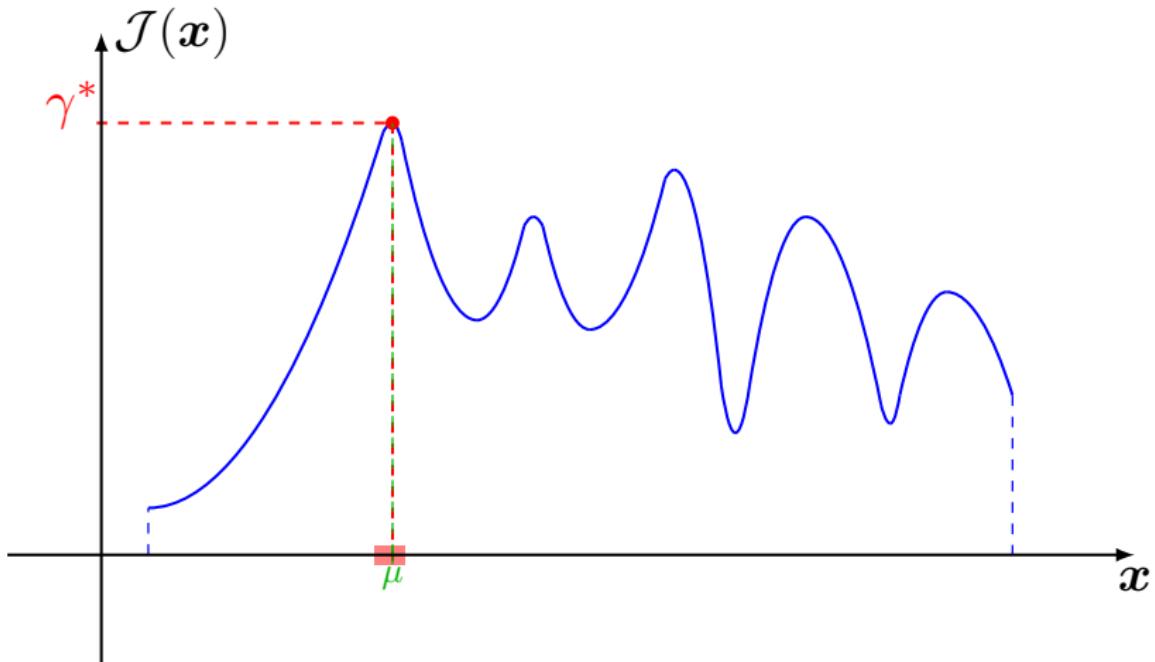
# Cross-entropy method in action!



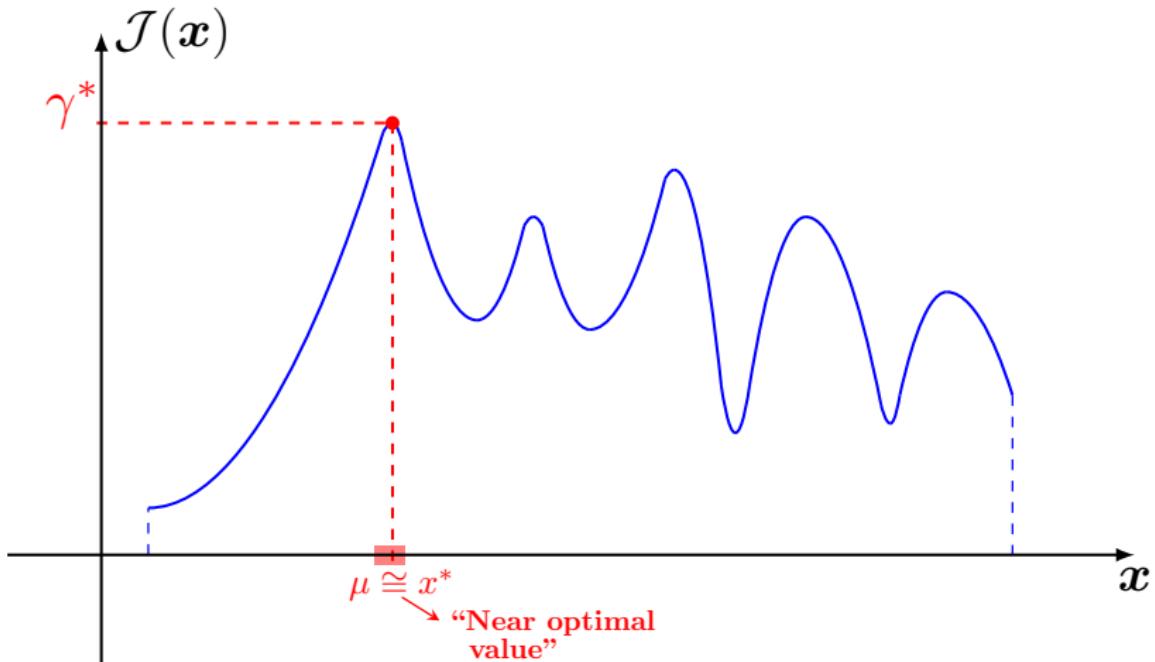
# Cross-entropy method in action!



# Cross-entropy method in action!



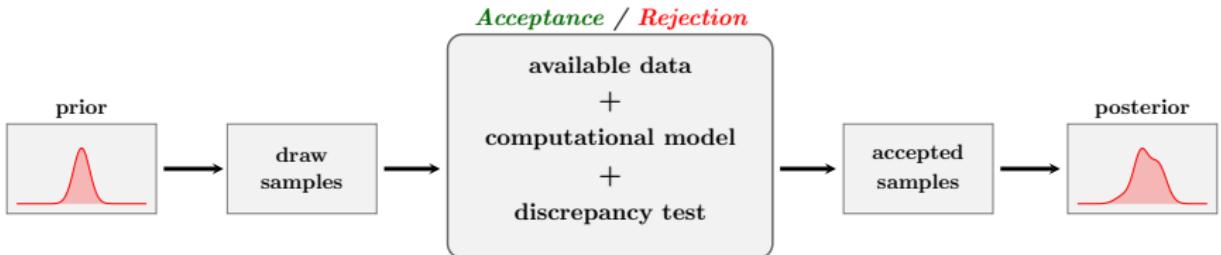
# Cross-entropy method in action!



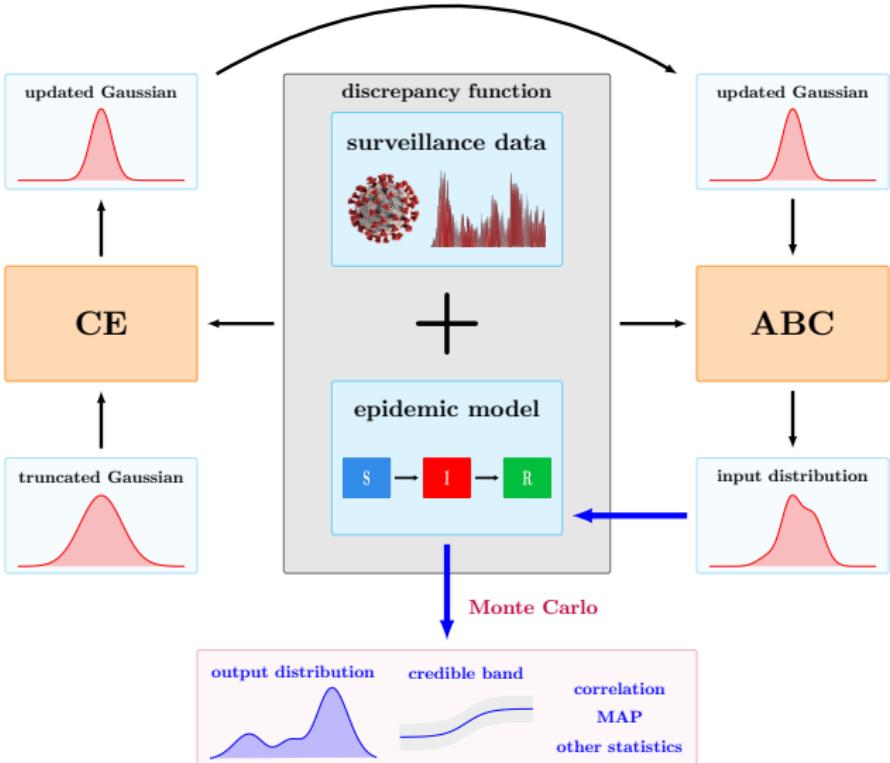
- ▶ Traditional Bayesian Computation

$$\underbrace{\pi(\text{model} \mid \text{data})}_{\text{posterior}} \propto \underbrace{\pi(\text{data} \mid \text{model})}_{\text{likelihood}} \times \underbrace{\pi(\text{model})}_{\text{prior}}$$

- ▶ Approximate Bayesian Computation

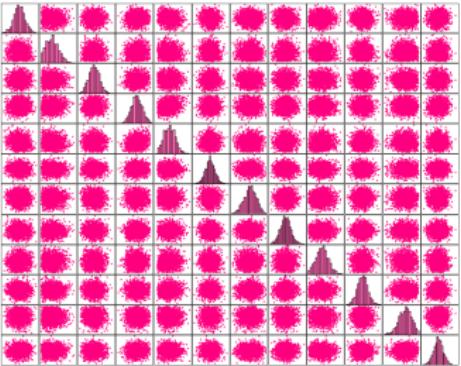


# Cross-Entropy Approximate Bayesian Computation



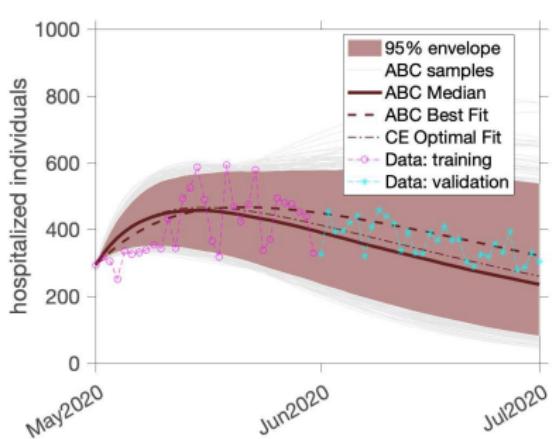
# Calibration of the SEIR(+AHD) $\beta$ model parameters

	Unit	CE Optimal	CE std dev	ABC Best	ABC std dev
$\beta$ or $\beta_0$	1/day	0.12	0.02	0.13	0.02
$\alpha$	1/day	0.20	0.07	0.27	0.06
$f_E$	—	0.81	0.03	0.84	0.03
$\gamma$	1/day	0.13	0.01	0.12	0.01
$\rho$	1/day	0.0006	0.0001	0.0005	0.0001
$\delta$	1/day	0.0021	0.0004	0.0015	0.0004
$\kappa_A$	—	0.0026	0.0008	0.0027	0.0008
$\kappa_H$	—	0.0563	0.0130	0.0575	0.0128
$\epsilon_H$	—	0.25	0.07	0.33	0.07
$\beta_\infty$	1/day	0.31	0.06	0.43	0.06
$\eta$	1/day	5.8	1.9	6.2	1.8
$\tau_\beta$	day	146	7	153	7

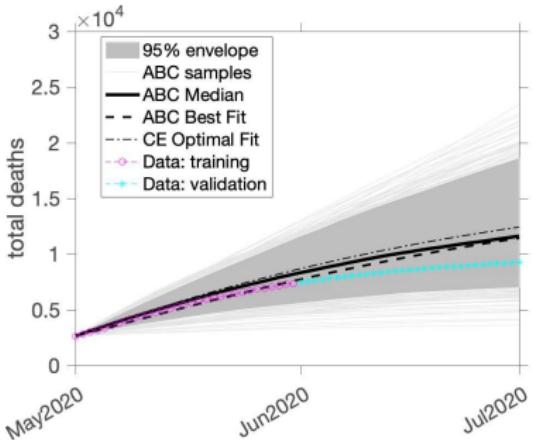


# Model validation and uncertainty quantification

hospitalized



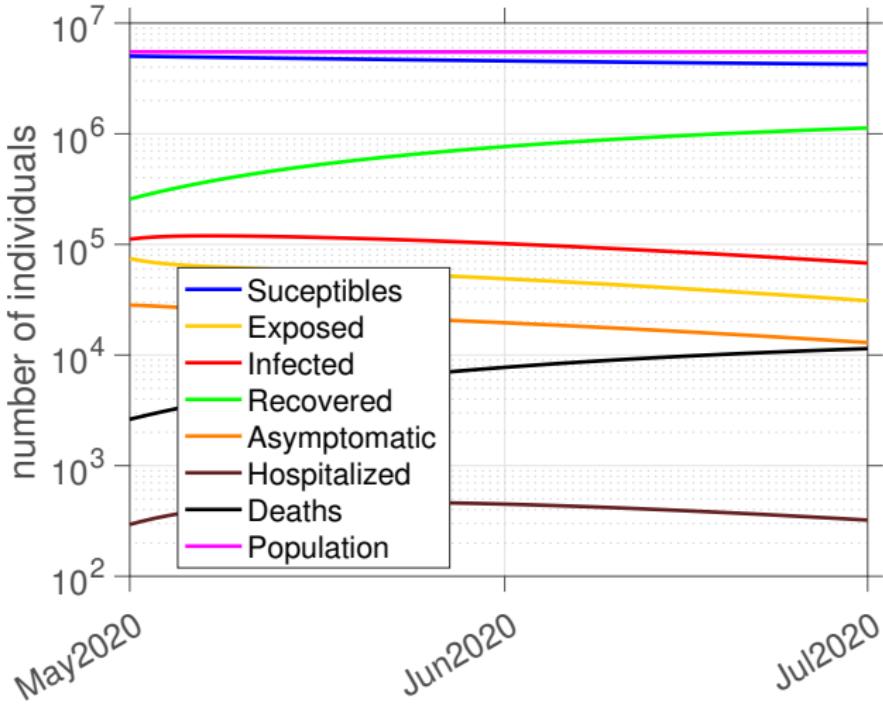
total deaths



$$\omega = 0.75 \quad N_{ce} = 100 \quad N_{abc} = 2000$$

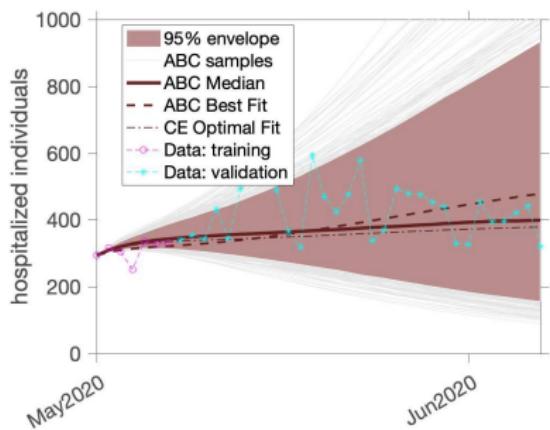
Acceptance rate: 87%

# Model exploration of latent variables

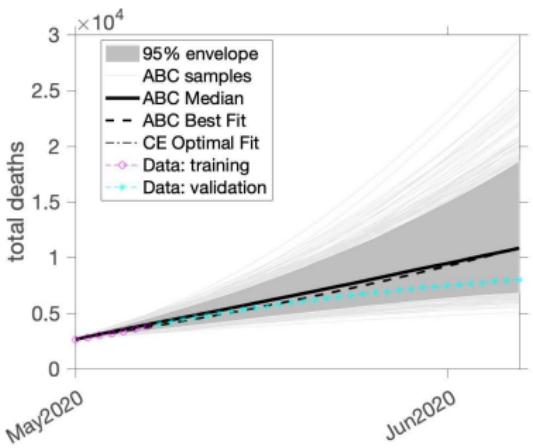


# Influence of the dataset size

hospitalized



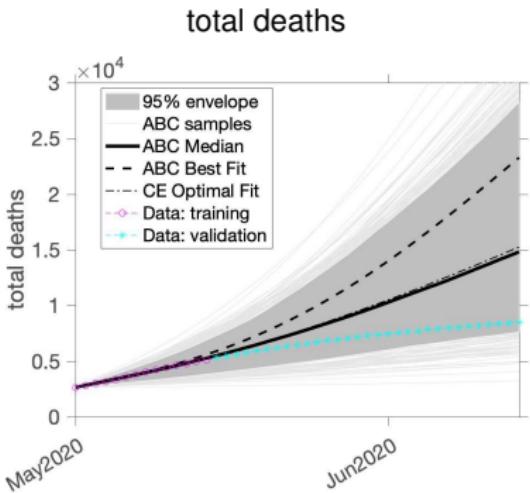
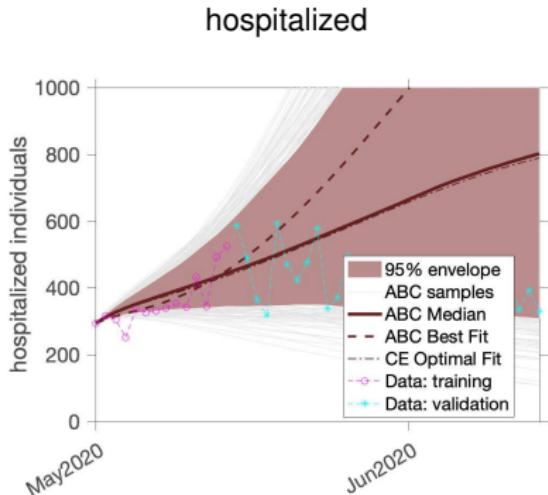
total deaths



$$\omega = 0.75 \quad N_{ce} = 100 \quad N_{abc} = 2000$$

May 1 - May 7, 2020

# Influence of the dataset size

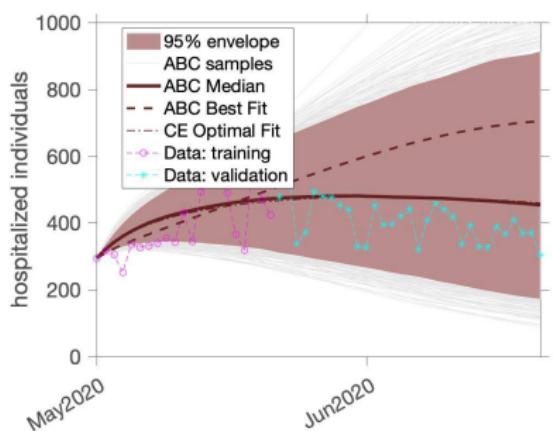


$$\omega = 0.75 \quad N_{ce} = 100 \quad N_{abc} = 2000$$

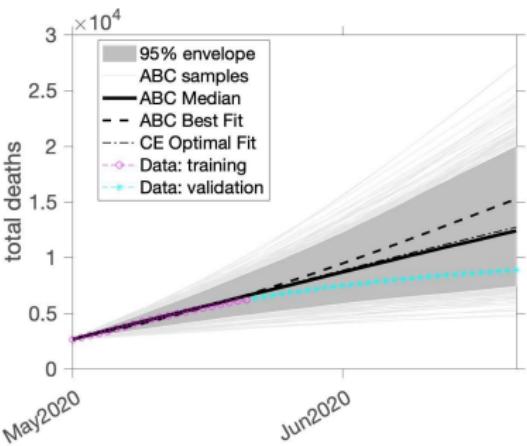
May 1 - May 14, 2020

# Influence of the dataset size

hospitalized



total deaths

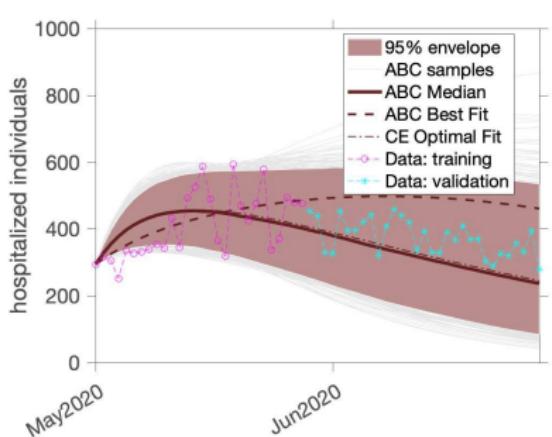


$$\omega = 0.75 \quad N_{ce} = 100 \quad N_{abc} = 2000$$

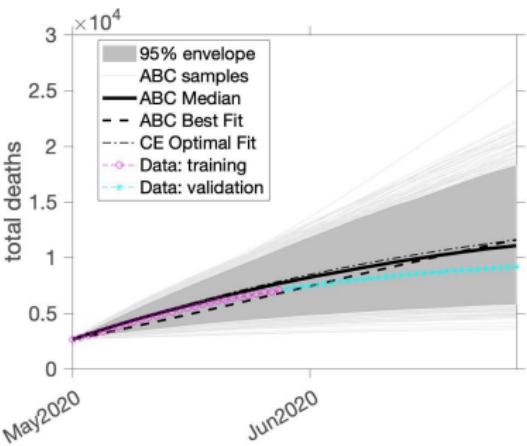
May 1 - **May 21**, 2020

# Influence of the dataset size

hospitalized



total deaths

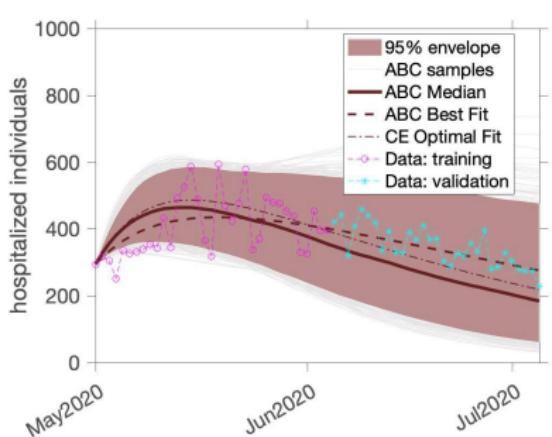


$$\omega = 0.75 \quad N_{ce} = 100 \quad N_{abc} = 2000$$

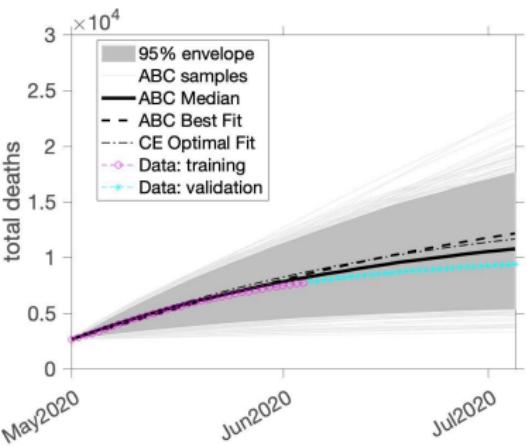
May 1 - **May 28**, 2020

# Influence of the dataset size

hospitalized



total deaths

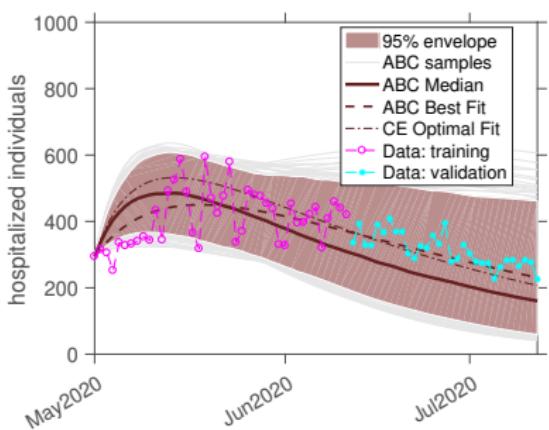


$$\omega = 0.75 \quad N_{ce} = 100 \quad N_{abc} = 2000$$

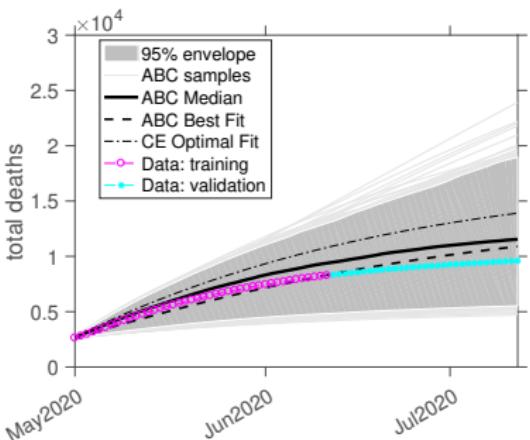
May 1 - June 4, 2020

# Influence of the dataset size

hospitalized



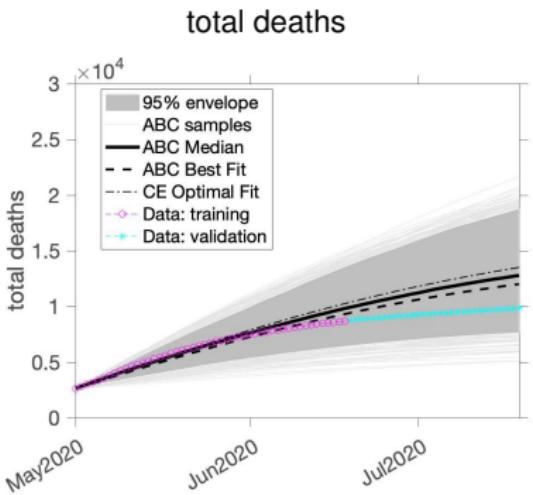
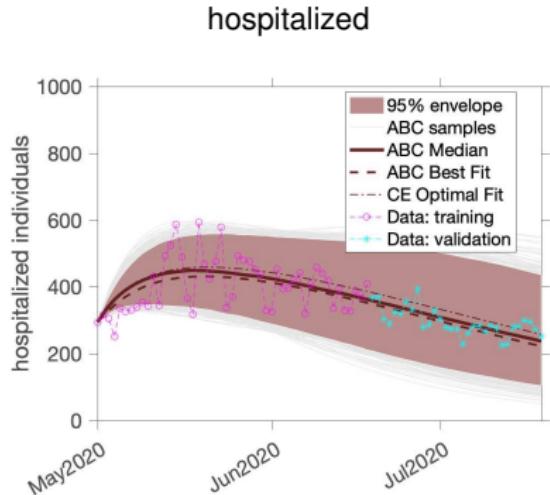
total deaths



$$\omega = 0.75 \quad N_{ce} = 100 \quad N_{abc} = 2000$$

May 1 - June 11, 2020

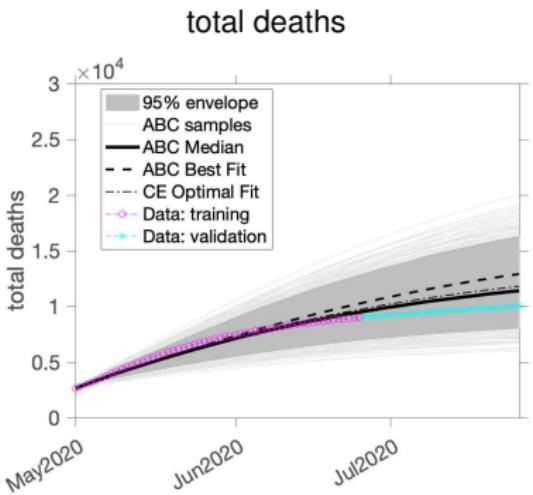
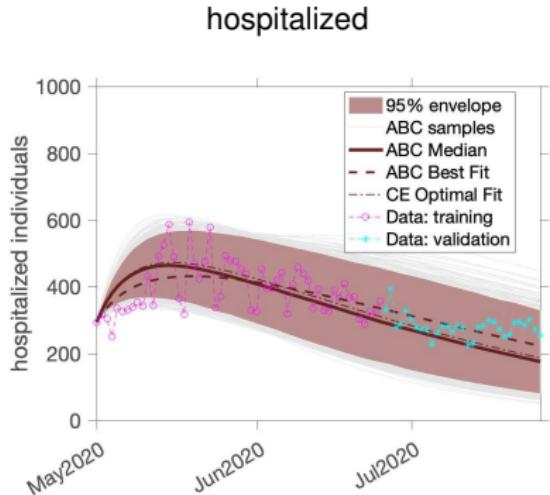
# Influence of the dataset size



$$\omega = 0.75 \quad N_{ce} = 100 \quad N_{abc} = 2000$$

May 1 - June 18, 2020

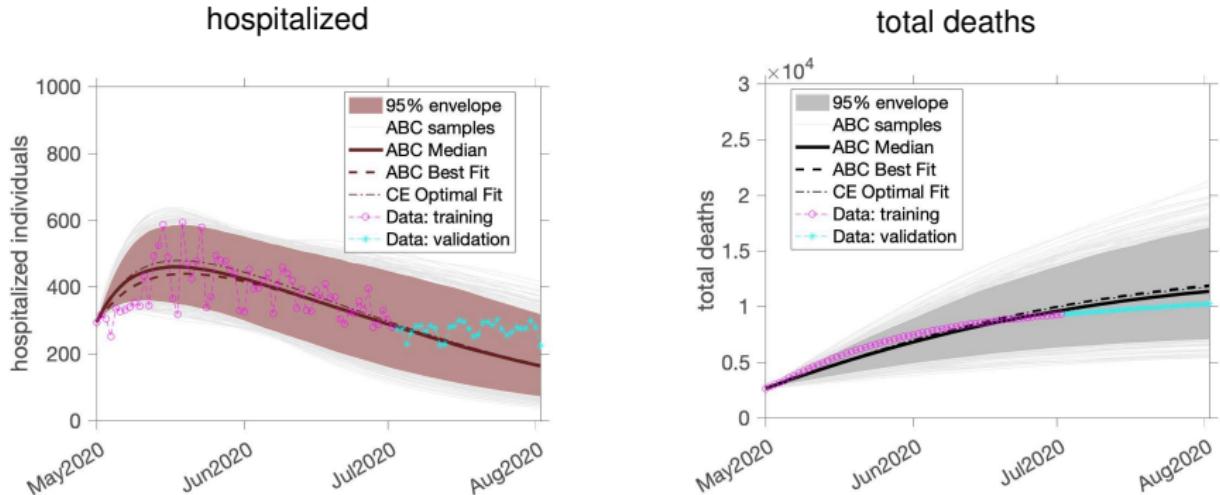
# Influence of the dataset size



$$\omega = 0.75 \quad N_{ce} = 100 \quad N_{abc} = 2000$$

May 1 - June 25, 2020

# Influence of the dataset size

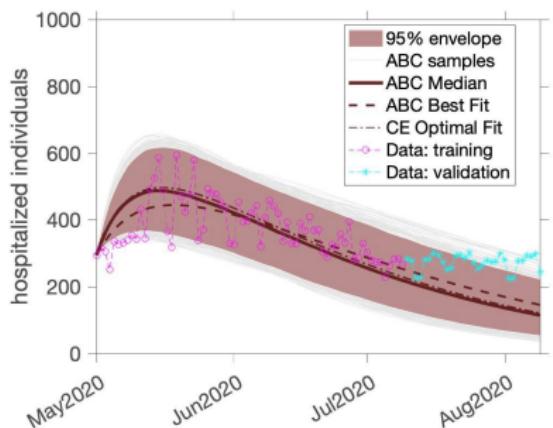


$$\omega = 0.75 \quad N_{ce} = 100 \quad N_{abc} = 2000$$

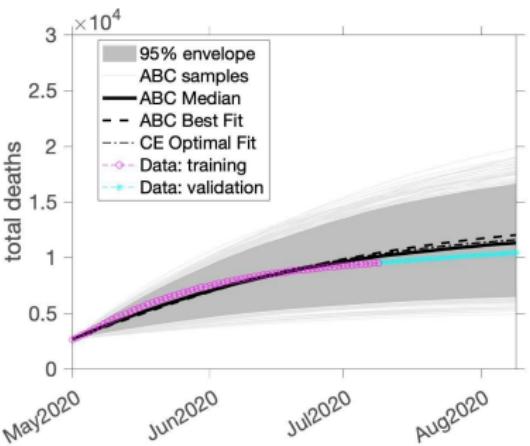
May 1 - July 2, 2020

# Influence of the dataset size

hospitalized



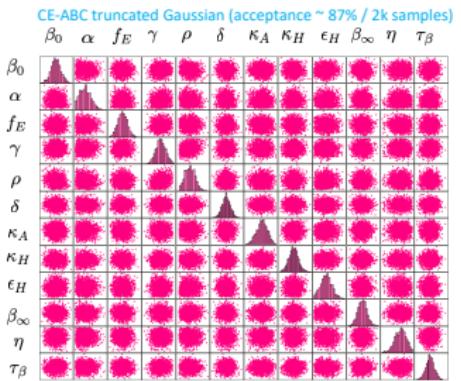
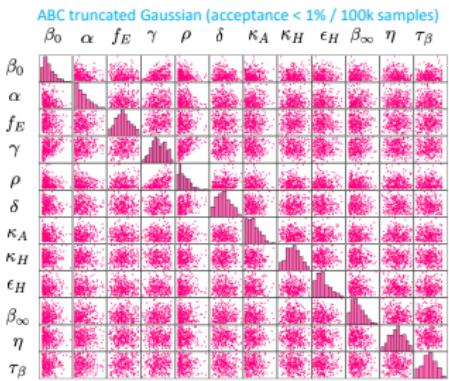
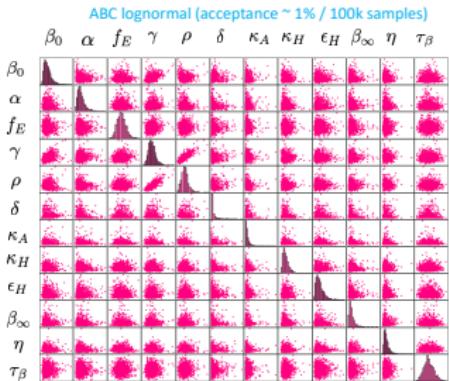
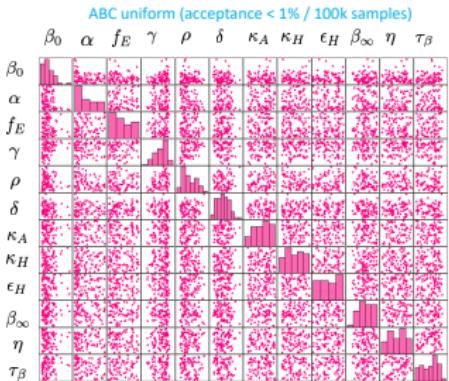
total deaths



$$\omega = 0.75 \quad N_{ce} = 100 \quad N_{abc} = 2000$$

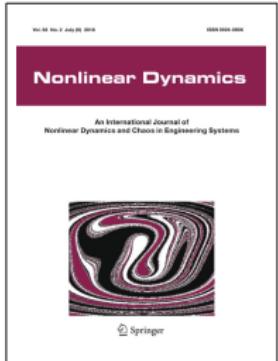
May 1 - July 9, 2020

# Comparing CE-ABC and classical ABC



## Final remarks

- ▶ Data-driven computational models may be a useful tool to aid with decision making during the course of an epidemic outbreak
- ▶ A Cross-Entropy Approximate Bayesian Computation (CE-ABC) framework for model calibration and uncertainty quantification is proposed
- ▶ CE-ABC + suitable dynamic model + reliable epidemic data may provide a very helpful insights (quantitative and qualitative) about an ongoing outbreak
- ▶ Our data-driven COVID-19 dynamic model for Rio de Janeiro:
  - ▶ “high-fidelity” predictability horizon: until one-week
  - ▶ “reasonable” predictability horizon: until one-month
- ▶ Model oriented decisions demand an interdisciplinary panel of experts:
  - ▶ Scientific
  - ▶ Ethical
  - ▶ Humanistic



Nonlinear Dyn (2023) 111:9649–9679  
<https://doi.org/10.1007/s11071-023-08327-8>

ORIGINAL PAPER



## Uncertainty quantification in mechanistic epidemic models via cross-entropy approximate Bayesian computation

Americo Cunha Jr  · David A. W. Barton  ·  
Thiago G. Ritto 



A. Cunha Jr, D. A. W. Barton and T. G. Ritto, *Uncertainty quantification in mechanistic epidemic models via cross-entropy approximate Bayesian computation*, *Nonlinear Dynamics*, 111:9649–9679, 2023

# Thank you for your attention!

[www.americocunha.org](http://www.americocunha.org)



 @AmericoCunhaJr

 @AmericoCunhaJr

 @AmericoCunhaJr