Simulation Phase

Prior Distribution
$p(\boldsymbol{\phi})$

Density

$\boldsymbol{\phi}$

Sampling

Generative Models For
Heterogeneous Populations

$$\mathrm{d}\boldsymbol{x}_t = f(\boldsymbol{x}_t, t, \boldsymbol{\phi})\,\mathrm{d}t + g(\boldsymbol{x}_t, t, \boldsymbol{\phi})\,\mathrm{d}\boldsymbol{B}_t$$

Simulate

Simulations
$\tilde{\boldsymbol{y}}_{\mathrm{sim}} \sim p(\boldsymbol{y} \mid \boldsymbol{\phi})$

Measurements

t

e.g., mRNA counts, drug
response level, ...

possible to store simulations

Up to 1 million simulations
(parallelization possible)

Training Phase

Simulations
$\tilde{\boldsymbol{y}}_{\mathrm{sim}} \sim p(\tilde{\boldsymbol{y}} \mid \boldsymbol{\phi})$

Measurements

t

Summary statistics

Latent
Distribution

Density

z

Training
Normalizing Flow

Individual-Specific Posteriors
$p(\boldsymbol{\phi} \mid \tilde{\boldsymbol{y}}_{\mathrm{sim}})$

Density

$\boldsymbol{\phi}$

4-6 hours per training
(utilizing GPUs possible)

Amortized Inference Phase

Real Data
$\tilde{\boldsymbol{y}}_{\mathrm{obs}} \sim p(\tilde{\boldsymbol{y}} \mid \boldsymbol{\phi})$

Measurements

t

Summary statistics

Latent
Distribution

Density

z

Inverting
Normalizing Flow

Individual-Specific Posteriors
$p(\boldsymbol{\phi} \mid \tilde{\boldsymbol{y}}_{\mathrm{obs}})$

Density

$\boldsymbol{\phi}$

Choice of population model $p(\boldsymbol{\phi} \mid \boldsymbol{\theta})$
with population parameter $\boldsymbol{\theta}$

$\boldsymbol{\phi}$

$$\boldsymbol{\theta}^* = \operatorname*{argmin}_{\boldsymbol{\theta}} - \sum_{i=1}^{n} \log\left(\frac{1}{M} \sum_{j=1}^{M} \frac{p(\boldsymbol{\phi}_j^{(i)} \mid \boldsymbol{\theta})}{p(\boldsymbol{\phi}_j^{(i)})}\right)$$

1. Generation of
starting values

$\boldsymbol{\phi}$

2. Sampling from
individual-specific posteriors

$\{\boldsymbol{\phi}_j^{(1)}\}_j^M \quad \{\boldsymbol{\phi}_j^{(2)}\}_j^M \quad \{\boldsymbol{\phi}_j^{(3)}\}_j^M$

3. Construction of
approximated likelihood

$\boldsymbol{\theta}^*$

4. Minimization of
negative log-likelihood

5. Performing
uncertainty analysis

Seconds to Minutes per new data set
(amortization of the previous phases)