Hierarchical models for SBI

- split hidden parameters $Y = [Y^G, Y^I]$

global: same for an entire population

individual: different for every instance

$\Bigg\}$ mixed effects model

- example virology Y^G properties of a patient (& disease)

Y^I —||— of individual cells we analyse

epidemiology Y^G properties of a region

Y^I —||— of individual persons

- <u>simulation</u>

$Y^G \sim p^S(Y^G)$ once for entire population

$Y^I_i \sim p^S(Y^I \mid Y^G)$ once for every individual / instance

$X_i = \phi(Y_G, Y^I_i, \eta_i)$ $\quad \eta_i \sim p^S(\eta \mid Y^G, Y^I_i)$ or $p^S(\eta)$

$l = \phi(Y^I_i, \eta_i)$ $\quad$ common simplification $] \Rightarrow$ TS estimation of a non-linear mixed effects model [Arruda et al 2023]

- given $p^s(Y^I \mid Y^G)$ and $p^s(Y^G)$ as part of simulation
- calculate marginal $p^s(Y^I) = \int p^s(Y^I \mid Y^G)\, p^s(Y^G)\, dY^G$ (either analytically or by learned generative model)
- use simulated TS to train a cNF for individual effects

$$p(Y_i^I \mid X_i) \approx p^s(Y_i^I \mid X_i)$$

- determine individual effects for a test set $\{X_i\}_{i=1}^N$ with fixed (unknown) Y^G

$$\{\hat{Y}_{ik}^I \sim p(Y_i^I \mid X_i)\}_{k=1}^M$$

(M posterior samples for Y_i^I for each i)

- find Y^G via numeric optimisation over $p(\{X_i\}_{i=1}^N \mid Y^G)$ (marginal likelihood)
  - under the conditional ($\hat{=}$ given Y^G) i.i.d assumption, marginal likelihood is

$$p^s(\{X_i\}_{i=1}^N \mid Y_G) = \prod_{i=1}^N \int \underbrace{p^s(X_i \mid Y^I)}_{\text{simulation}}\, \underbrace{p^s(Y^I \mid Y^G)}_{\text{hierarchical prior for fixed } Y^G}\, dY^I$$

Bayes rule: $p^s(X_i \mid Y_i^I) = \dfrac{p^s(X_i)\; p^s(Y_i^I \mid X_i)}{p^s(Y_i^I)}$

$$p^s(\{x_i\}_{i=1}^N \mid Y_G) = \prod_{i=1}^N p^s(x_i) \int p^s(Y^I \mid x_i) \cdot \frac{p^s(Y^I \mid Y^G)}{p^s(Y^I)} \, dY^I$$

$$\hat{Y}^G = \underset{Y^G}{\arg\min} \; -\log p^s(\{X_i\}_{i=1}^N \mid Y_G)$$ (standard maximum likelihood estim.

$$= \underset{Y^G}{\arg\min} \sum_{i=1}^N -\log \underbrace{\mathbb{E}_{Y^I \sim p^s(Y^I \mid X_i)} \left[\frac{p^s(Y^I \mid Y^G)}{p^s(Y^I)} \right]}_{\text{approximate by empirical average over } \underbrace{p(Y^I \mid X_i)}_{cNF}}$$

$$\approx \underset{Y^G}{\arg\min} \sum_{i=1}^N -\log \left[\frac{1}{M} \sum_{k=1}^M \frac{p^s(Y_{ik}^I \mid Y^G)}{p^s(Y_{ik}^I)} \right] \qquad Y_{ik}^I \sim p(Y^I \mid X_i)$$

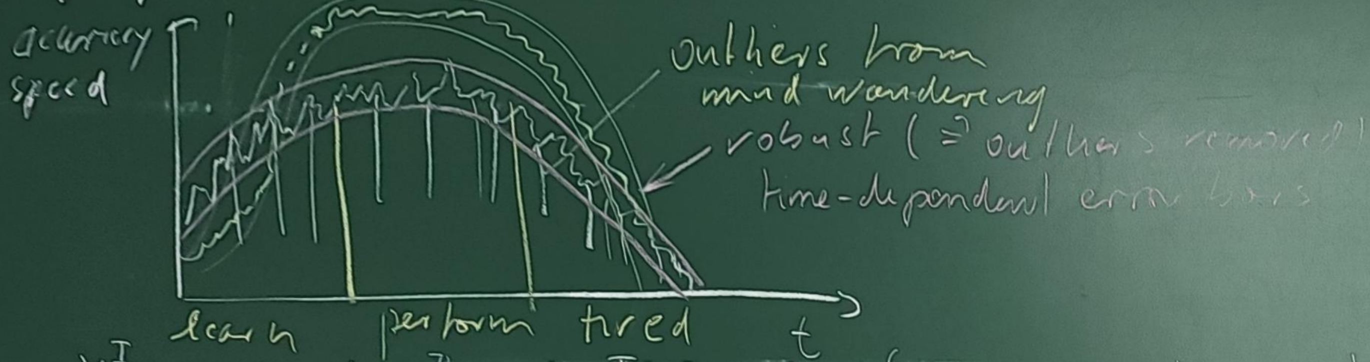
solve by standard non-linear optimazation (gradient decent, Newton, BFGS)

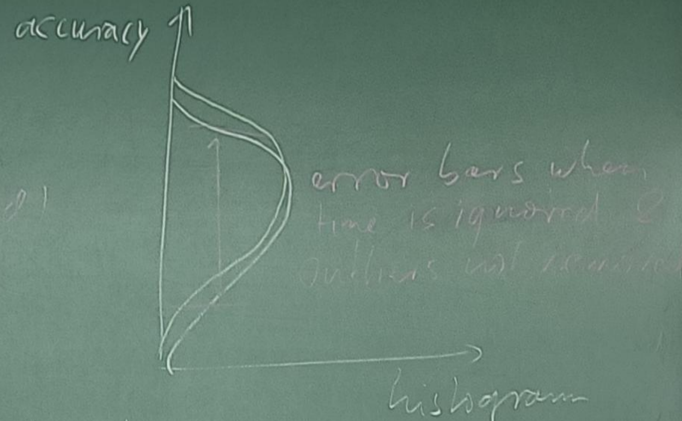
- application: mRNA transfection ($\hat{=}$ how mRNA vaccines enter the cells)

variant: non-stationary time series: Y^G parameters constant over time

Y_t^I —‖— vary over time

e.g. psychological experiment:





e.g. Y_{it}^I as random walk: $Y_{it}^I = Y_{i,t-1}^I + f(Y_i^G, Y_{i,1:t-1}, \eta_t)$ most information is lost

- train SBI model with recurrent summary network (e.g. LSTM network)

$$\mathbb{R}^d \ni h(X_{i,1:t})$$

observations for instance i from timesteps 1 … t

$$\hat{p}, \hat{h} = \underset{p,h}{\arg\min} \sum_{i=1}^{N} \sum_{t=1}^{T_i} -\log p\left(Y_{it}^I, Y_i^G \mid h(X_{i,1:t})\right)$$

cNF summary network

Y_i^G: stationary parameters of instance i, Y_{it}^I: non-stationary parameters of instance i at time t

- simulation: ⓪ $Y_i^G \sim p^S(Y^G)$ $\quad Y_{i0}^I \sim p^S(Y_0^I \mid Y_i^G)$

① for $t = 1, \dots, T_i$

ⓐ $Y_{i:t}^I = Y_{i:t-1}^I + f(Y_i^G, Y_{i,1:t-1}^I, \eta_i^I)$ $\quad \eta_i^I$ noise for Y_{it}^I

ⓑ $X_{it} = \phi(Y_{i:t}^I, \eta_i^X)$ $\quad \eta_i^X$ noise for X_{it}

$$\hat{p}(Y_{it}^I, Y_i^G \mid \hat{h}(X_{1:t}^{obs}))$$