Cranmer, Kyle, Johann Brehmer, and Gilles Louppe. "The frontier of simulation-based inference." PNSA 117.48 (2020): 30055-30062.
Radev, Stefan T., et al. "BayesFlow: Learning complex stochastic models with invertible neural networks." IEEE trans on neural networks and learning systems 33.4 (2020): 1452-1466.

Neural posterior estimation

Likelihood-free inference

Train a *conditional* neural network w/ *in silico* simulations on pairs $\{(\mathbf{x}^{(m)}, \mathbf{y}^{(m)})\}_{m=1}^{M}$

$$\mathbf{y} \sim p(\mathbf{y} \,|\, \mathbf{x}) \quad \Longleftrightarrow \quad \mathbf{y} = \mathrm{g}(\mathbf{x}, \zeta)$$

where $\zeta \sim p(\zeta \,|\, \mathbf{x})$ with $\mathbf{x} \sim p(\mathbf{x})$

Given these pairs train via

$$\underset{\phi,\psi}{\text{minimize}} \frac{1}{M} \sum_{m=1}^{M} \left(\frac{1}{2} \left\| f_\phi\left(\mathbf{x}^{(m)}; h_\psi(\mathbf{y}^{(m)})\right) \right\|_2^2 - \log \left| \det J_{f_\phi}^{(m)} \right| \right)$$

- f_ϕ is the *inference network* defined by a *Conditional Normalizing Flow* (CNF)
- h_ψ is a learned *summary* statistic yielding $MI\left(\mathbf{x}, h_\psi(\mathbf{y})\right) \approx MI\left(\mathbf{x}, \mathbf{y}\right)$

Radev, Stefan T., et al. "BayesFlow: Learning complex stochastic models with invertible neural networks." *IEEE transactions on neural networks and learning systems* 33.4 (2020): 1452-1466.
Orozco, Rafael, et al. "Adjoint operators enable fast and amortized machine learning based Bayesian uncertainty quantification." Medical Imaging 2023: Image Processing. Vol. 12464. SPIE, 2023.
Alsing, Justin, Benjamin Wandelt, and Stephen Feeney. "Massive optimal data compression and density estimation for scalable, likelihood-free inference in cosmology." Monthly Notices of the Royal Astronomical Society 477.3 (2018): 2874-2885.

SLIM ML4Seismic

Summary statistic

physics-based

Assume Gaussian log-likelihood:

$$\log p(\mathbf{y} \mid \mathbf{x}) = \|\mathscr{H}(\mathbf{x}) - \mathbf{y}\|_2^2$$

Use *score* function at *fiducial* point $\bar{\mathbf{x}}$:

$$p(\mathbf{x} \mid \nabla_{\mathbf{x}} \log p(\mathbf{y} \mid \mathbf{x}) \big|_{\bar{\mathbf{x}}}) \approx p(\mathbf{x} \mid \mathbf{y})$$

in combination w/ learned *summary network* yields

$$\bar{\mathbf{y}} = h_\psi(\mathbf{J}_{\mathscr{H}}^\top(\bar{\mathbf{x}})(\mathscr{H}(\bar{\mathbf{x}}) - \mathbf{y}))$$

such that

$$p(\mathbf{x} \mid \bar{\mathbf{y}}) \approx p(\mathbf{x} \mid \mathbf{y})$$