Solution: Bayesian experimental design

Chose acquisition design $\mathbf{W}[\mathbf{w}]$ that allows for **maximal** information gain

$$\mathbf{y} = \mathbf{W}(\mathbf{x})$$

where $\mathbf{W}$ is a binary sampling mask derived from density $\mathbf{w}$

Collect data by maximizing the Kullback-Leibler divergence:

$$\max_{\mathbf{W}} D_{KL}(p(\mathbf{x}\,|\,\mathbf{y})\,||\,p(\mathbf{x}))\,.$$

***Maximize Expected information gain* (EIG) *averages* over *all* possible designs**

$$\max_{\mathbf{W}} \{EIG(\mathbf{W}) = \mathbb{E}_{p(\mathbf{y}|\mathbf{W})}\,\left[D_{KL}(p(\mathbf{x}\,|\,\mathbf{y})\,||\,p(\mathbf{x}))\right]\}\,.$$

Go, Jinwoo, and Tobin Isaac. "Robust expected information gain for optimal Bayesian experimental design using ambiguity sets." *Uncertainty in Artificial Intelligence*. PMLR, 2022.

SLIM ML4Seismic

Hoffmann, Till, and Jukka-Pekka Onnela. "Minimizing the Expected Posterior Entropy Yields Optimal Summary Statistics." arXiv preprint arXiv:2206.02340 (2022).

SLIM ML4Seismic

Relation

conditional neural density & EIG

Maximizing the expected *posterior density* is *equivalent* to maximizing the expected *information gain*

$$\max_{\mathbf{W}} \; EIG(\mathbf{W}) = \mathbb{E}_{p(\mathbf{y}|\mathbf{W})} \left[D_{KL}(p_\theta(\mathbf{x}\,|\,\mathbf{y}) \,||\, p(\mathbf{x})) \right] = \mathbb{E}_{p(\mathbf{y}|\mathbf{W})} \left[\mathbb{E}_{p(\mathbf{x}|\mathbf{y})} \left[\log p_\theta(\mathbf{x}\,|\,\mathbf{y}) - \log p(\mathbf{x}) \right] \right]$$

$$= \mathbb{E}_{p(\mathbf{y}|\mathbf{W})} \left[\mathbb{E}_{p(\mathbf{x}|\mathbf{y})} \left[\log p_\theta(\mathbf{x}\,|\,\mathbf{y}) \right] \right]$$ law of total expectation

$$= \mathbb{E}_{p(\mathbf{x},\mathbf{y}|\mathbf{W})} \left[\log p_\theta(\mathbf{x}\,|\,\mathbf{y}) \right]$$ same as neural posterior objective!

Thus optimizing under the posterior density objective will increase the EIG!

Foster, Adam, et al. "A unified stochastic gradient approach to designing bayesian-optimal experiments." PMLR, 2020.
Hoffmann, Till, and Jukka-Pekka Onnela. "Minimizing the Expected Posterior Entropy Yields Optimal Summary Statistics." arXiv preprint arXiv:2206.02340 (2022).