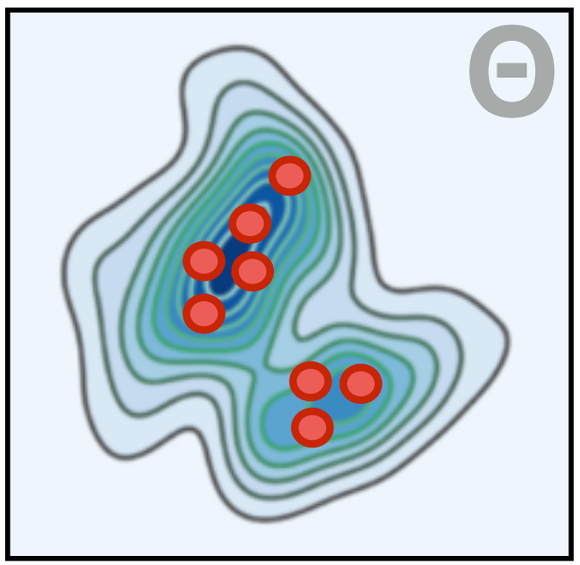
Variational Inference with Stein Mixtures

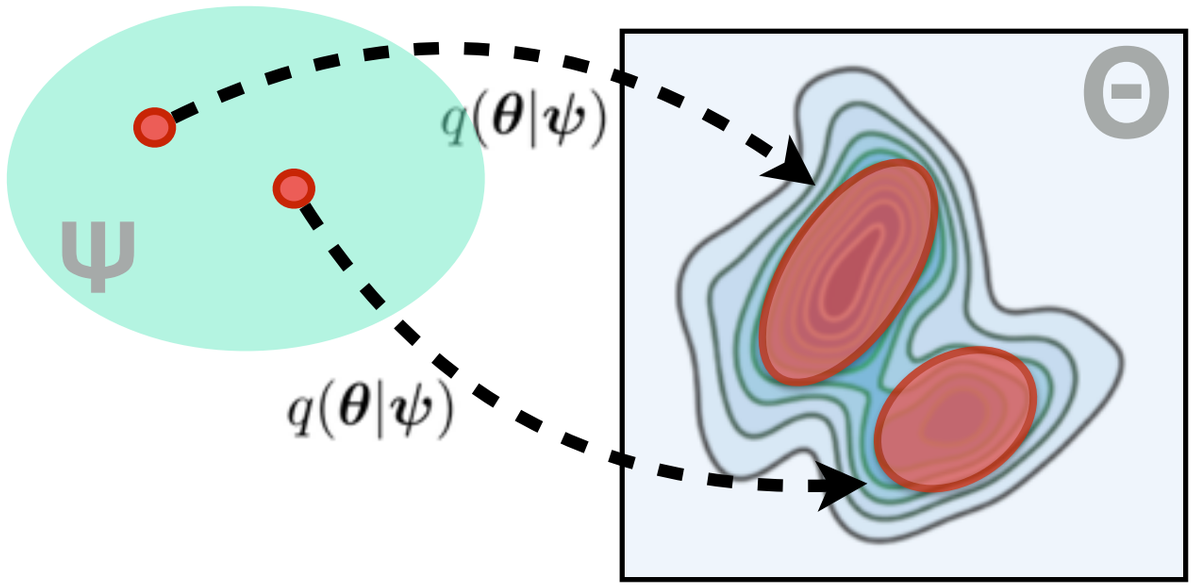
We extend *Stein variational gradient descent* [Liu & Wang, NIPS 2016] to operate on the parameters of the variational approximation.

STEIN PARTICLE APPROXIMATION



$$p(\boldsymbol{\theta}|\mathbf{X}) \approx \frac{1}{K}\sum_k \delta[\boldsymbol{\theta}_k]$$

STEIN MIXTURE APPROXIMATION



$$p(\boldsymbol{\theta}|\mathbf{X}) \approx \frac{1}{K}\sum_k q(\boldsymbol{\theta}|\boldsymbol{\psi}_k)$$

Eric Nalisnick

Padhraic Smyth

Variational Inference with Stein Mixtures

We extend *Stein variational gradient descent* [Liu & Wang, NIPS 2016] to operate on the parameters of the variational approximation.

*K*TH PARTICLE UPDATE

KERNEL

MODEL JOINT

$$\boldsymbol{\phi}[\boldsymbol{\psi}_k^t] = \frac{1}{K}\sum_{j=1}^{K} k(\boldsymbol{\psi}_j, \boldsymbol{\psi}) \sum_s \tilde{w}_s \nabla_{\boldsymbol{\psi}_j} \log \frac{p(\mathbf{X}, \hat{\boldsymbol{\theta}}_s)}{q(\hat{\boldsymbol{\theta}}_s|\boldsymbol{\psi}_j)} + \nabla_{\boldsymbol{\psi}_j} k(\boldsymbol{\psi}_j, \boldsymbol{\psi})$$

IMPORTANCE WEIGHT

Eric Nalisnick

Padhraic Smyth