A Differentiable POGLM with Forward-Backward Message Passing

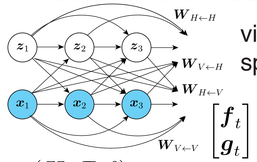
Chengrui Li, Weihan Li, Yule Wang, Anqi Wu @ GaTech CSE

Georgia Tech

ICML 2024 International Conference On Machine Learning

Partially observable generalized linear model (POGLM)

V visible neurons and $H = N - V$ hidden neurons.



$p(\boldsymbol{X}, \boldsymbol{Z}; \theta)$

visible/hidden spike counts $\begin{bmatrix}\boldsymbol{x}_t\\ \boldsymbol{z}_t\end{bmatrix} \sim \mathrm{Pois}\left(\begin{bmatrix}\boldsymbol{f}_t\\ \boldsymbol{g}_t\end{bmatrix}\right)$ visible/hidden firing rates

$$\begin{bmatrix}\boldsymbol{f}_t\\ \boldsymbol{g}_t\end{bmatrix} = \sigma\left(\underbrace{\begin{bmatrix}\boldsymbol{b}_V\\ \boldsymbol{b}_H\end{bmatrix}}_{\text{bias}} + \underbrace{\begin{bmatrix}\boldsymbol{W}_{V\leftarrow V} & \boldsymbol{W}_{V\leftarrow H}\\ \boldsymbol{W}_{H\leftarrow V} & \boldsymbol{W}_{H\leftarrow H}\end{bmatrix}}_{\text{weight}} \underbrace{\left(\sum_{l=1}^{L}\psi_l \begin{bmatrix}\boldsymbol{x}_{t-l}\\ \boldsymbol{z}_{t-l}\end{bmatrix}\right)}_{\text{convolved history}}\right)$$

Observed variable: visible spike train $\boldsymbol{X} = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_T]^{\mathrm{T}} \in \mathbb{N}^{T\times V}$

Latent variable: hidden spike train $\boldsymbol{Z} = [\boldsymbol{z}_1, \ldots, \boldsymbol{z}_T]^{\mathrm{T}} \in \mathbb{N}^{T\times H}$

Generative parameter set: $\theta = \left\{\boldsymbol{b} \in \mathbb{R}^{V+H}, \boldsymbol{W} \in \mathbb{R}^{(V+H)\times(V+H)}\right\}$

Variational inference (VI)

Variational model $q(\boldsymbol{Z}|\boldsymbol{X};\phi)$ parameterized by ϕ :

$\boldsymbol{z}_t \sim \mathrm{Pois}(\boldsymbol{g}_t)$ related to **inference methods**

$\boldsymbol{g}_t = \mathrm{function}(\boldsymbol{X};\phi)$ related to **sampling schemes**

Evidence lower bound: $\mathrm{ELBO}(\boldsymbol{X};\theta,\phi) = \mathbb{E}_q\left[\ln p(\boldsymbol{X},\boldsymbol{Z};\theta) - \ln q(\boldsymbol{Z}|\boldsymbol{X};\phi)\right]$

VI: Inference methods

Gradient estimators

	score function	pathwise
distribution	any distribution	continuous distribution with reparameterization trick $\boldsymbol{Z}\|\boldsymbol{X};\phi = r(\boldsymbol{\epsilon}\|\boldsymbol{X};\phi)$
samples	$\{\boldsymbol{Z}^{(k)}\}_{k=1}^{K} \sim q(\boldsymbol{Z}\|\boldsymbol{X};\phi)$	$\{\boldsymbol{\epsilon}^{(k)}\}_{k=1}^{K} \sim R(\boldsymbol{\epsilon})$
$\frac{\partial\,\mathrm{ELBO}(\boldsymbol{X};\theta,\phi)}{\partial\phi} \approx$	$\frac{\partial}{\partial\phi}\frac{-1}{2K}\sum_{k=1}^{K}\left[\ln p\left(\boldsymbol{X},\boldsymbol{Z}^{(k)};\theta\right) - \ln q\left(\boldsymbol{Z}^{(k)}\middle\|\boldsymbol{X};\phi\right)\right]^2$	$\frac{\partial}{\partial\phi}\frac{1}{K}\sum_{k=1}^{K}\left[\ln p\left(\boldsymbol{X}, r\left(\boldsymbol{\epsilon}^{(k)}\middle\|\boldsymbol{X};\phi\right);\theta\right) - \ln q\left(r\left(\boldsymbol{\epsilon}^{(k)}\middle\|\boldsymbol{X};\phi\right)\middle\|\boldsymbol{X};\phi\right)\right]$

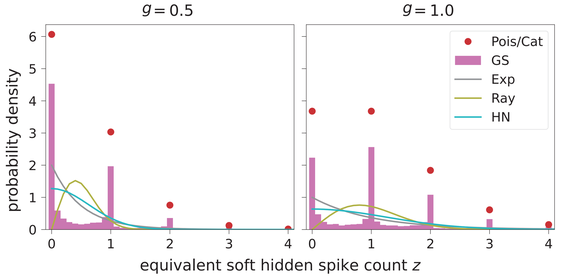
Relaxation

Poisson $\mathrm{Pois}(g)$ $\xrightarrow[\text{upperbound } M]{\text{truncate at an}}$ categorical $\mathrm{Cat}(\boldsymbol{\pi}(g))$ $\xrightarrow[\text{approximation}]{\text{reparameterization}}$ Gumbel-Softmax $\mathrm{GS}(\boldsymbol{\pi}(g);\tau)$

References: [1] Pillow & Latham, NeurIPS, 2007. [2] Jimenez Rezende & Gerstner, Front. Comp. Neur., 2014. [3] Linderman et al., NeurIPS, 2017. [4] Jang et al., arXiv, 2016. [5] Maddison et al., arXiv, 2016. [6] Kajino, ICML, 2021. [7] Li et al., ICLR, 2024.

Candidate distributions

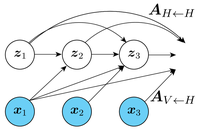
Replace the distribution $\mathbb{P}[z;g]$ in the generative and variational model with the following distributions governed by $\mathbb{E}[z;g] = g$.



distribution	can pathwise
$\mathrm{Pois}(g)$	✗
$\mathrm{Cat}(\boldsymbol{\pi}(g))$	✗
$\mathrm{GS}(\boldsymbol{\pi}(g);\tau)$	✓
$\mathrm{Exp}\left(\frac{1}{g}\right)$	✓
$\mathrm{Ray}\left(\sqrt{\frac{2}{\pi}}g\right)$	✓
$\mathrm{HN}\left(\sqrt{\frac{\pi}{2}}g\right)$	✓

VI: Sampling schemes

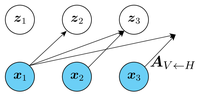
Three message passing designs, where $\boldsymbol{c}$ mimicing $\boldsymbol{b}$ and $\boldsymbol{A}$ mimicing $\boldsymbol{W}$.



Forward-self (FS)

$$\boldsymbol{g}_t = \sigma\left(\boldsymbol{c}_H + \boldsymbol{A}_{H\leftarrow V}\left(\sum_{l=1}^{L}\psi_l\,\boldsymbol{x}_{t-l}\right) + \boldsymbol{A}_{H\leftarrow H}\left(\sum_{l=1}^{L}\psi_l\,\boldsymbol{z}_{t-l}\right)\right)$$

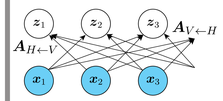
$$\phi = \left\{\boldsymbol{c}_H \in \mathbb{R}^H, \boldsymbol{A}_{H\leftarrow V} \in \mathbb{R}^{H\times V}, \boldsymbol{A}_{H\leftarrow H} \in \mathbb{R}^{H\times H}\right\}$$



Forward (F)

$$\boldsymbol{g}_t = \sigma\left(\boldsymbol{c}_H + \boldsymbol{A}_{H\leftarrow V}\left(\sum_{l=1}^{L}\psi_l\,\boldsymbol{x}_{t-l}\right)\right)$$

$$\phi = \left\{\boldsymbol{c}_H \in \mathbb{R}^H, \boldsymbol{A}_{H\leftarrow V} \in \mathbb{R}^{H\times V}\right\}$$



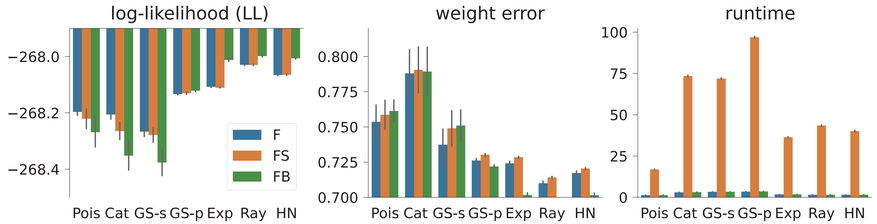
Forward-backward (FB)

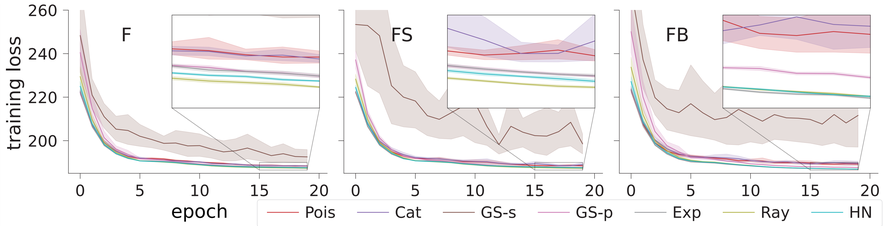
$$\boldsymbol{g}_t = \sigma\left(\boldsymbol{c}_H + \boldsymbol{A}_{H\leftarrow V}\left(\sum_{l=1}^{L}\psi_l\,\boldsymbol{x}_{t-l}\right) + \boldsymbol{A}_{V\leftarrow H}\left(\sum_{l=1}^{L}\psi_l\,\boldsymbol{x}_{t+l}\right)\right)$$

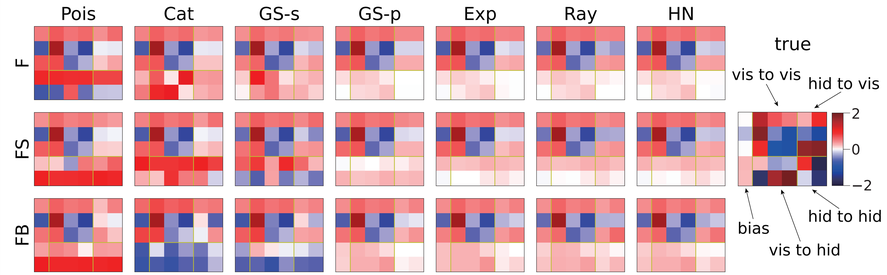
$$\phi = \left\{\boldsymbol{c}_H \in \mathbb{R}^H, \boldsymbol{A}_{H\leftarrow V} \in \mathbb{R}^{H\times V}, \boldsymbol{A}_{V\leftarrow H} \in \mathbb{R}^{V\times H}\right\}$$

Experiments

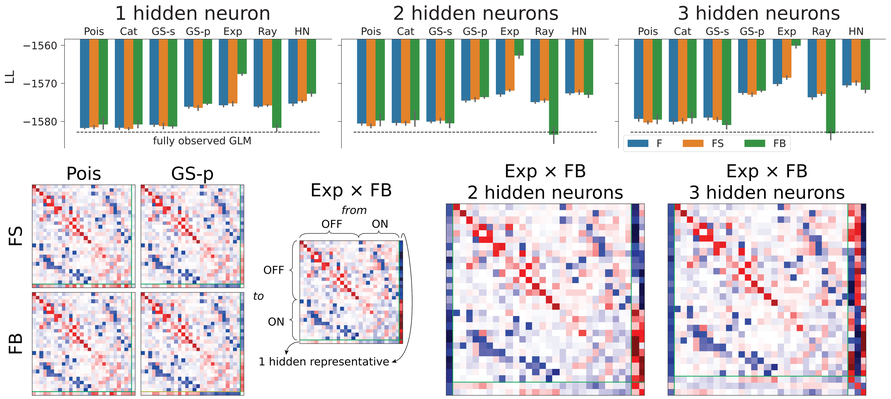
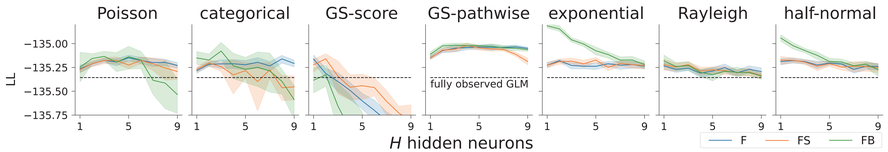
Synthetic







Retinal ganglion neurons (Pillow & Scott, 2012)



Primary visual cortex (cncrs: PVC-5)