Sandia National Laboratories

Exceptional service in the national interest

ADVANCING TENSOR DATA ANALYSIS: THE GENERALIZED MULTILINEAR MODEL (GMLM)

Carlos Llosa,

Danny Dunlavy, Jeremy Myers, Rich Lehoucq, Tian Ma

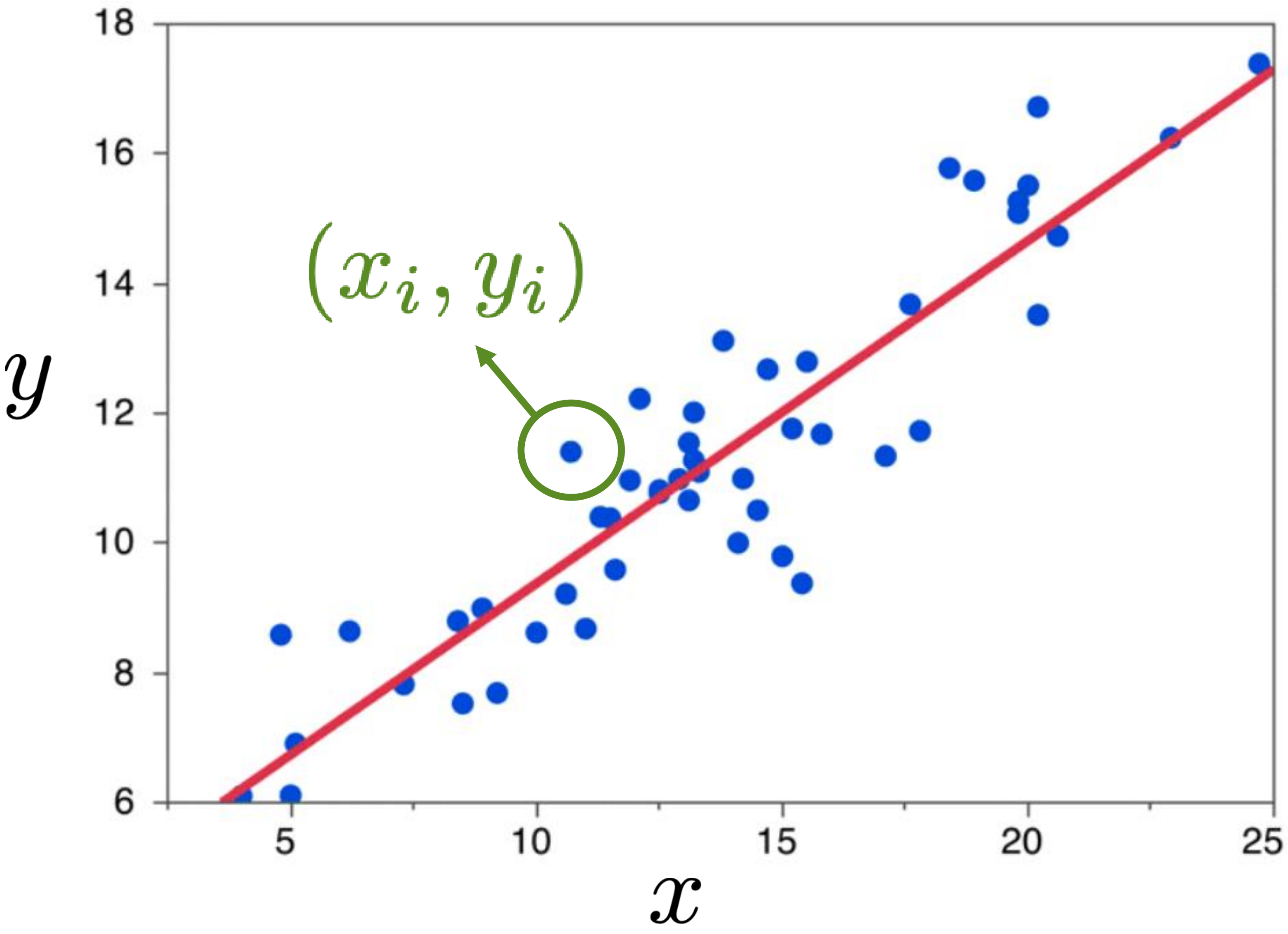
Sandia National Laboratories, USA

U.S. DEPARTMENT of ENERGY NNSA National Nuclear Security Administration

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

SAND2025-12497PE

LINEAR REGRESSION



Simple linear regression:

$$y_i = \beta x_i + \epsilon_i$$

Multiple linear regression:

$$y_i = \boldsymbol{\beta}^T \mathbf{x}_i + \epsilon_i$$

responses

covariates

TENSOR-ON-TENSOR REGRESSION

$$\langle S|T\rangle = \begin{cases} \text{inner product} & dim(S) = dim(T) \\ \text{contraction} & dim(S) \neq dim(T) \end{cases}$$

Linear regression: $$y_i = \langle \boldsymbol{x}_i | \boldsymbol{\beta} \rangle + \epsilon_i$$

Tensor regression: $$y_i = \langle \mathcal{X}_i | \mathcal{B} \rangle + \epsilon_i$$

Multivariate linear regression: $$\boldsymbol{y}_i = \langle \boldsymbol{x}_i | B \rangle + \boldsymbol{\epsilon}_i$$

Tensor-on-Tensor regression (ToTR): $$\mathcal{Y}_i = \langle \mathcal{X}_i | \mathcal{B} \rangle + \mathcal{E}_i$$

Lock (2018); Llosa & Maitra (2022)

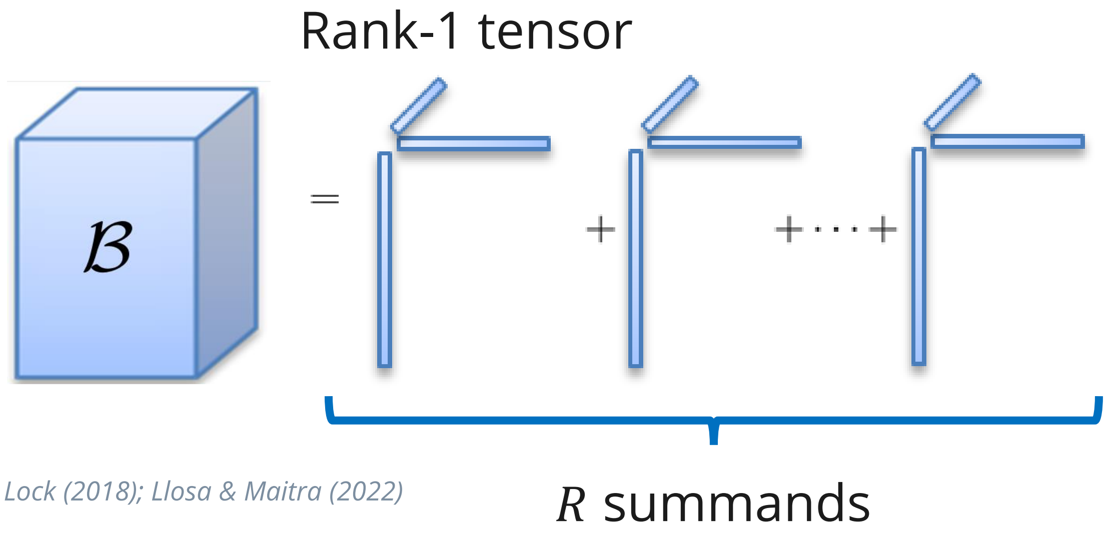
TOTR: HOW MUCH DATA DO WE NEED TO FIT THE MODEL

Tensor-on-Tensor regression (ToTR):

$$\underbrace{\mathcal{Y}_i}^{M_1 \times \cdots \times M_P} = \langle \overbrace{\mathcal{X}_i}^{N_1 \times \cdots \times N_Q} | \underbrace{\mathcal{B}}_{N_1 \times \cdots \times N_Q \times M_1 \times \cdots \times M_P} \rangle + \mathcal{E}_i$$

$i = 1, 2, \ldots, n \longrightarrow$ **so many samples needed!**

Low-rank Canonical Polyadic (CP) Model:



Lock (2018); Llosa & Maitra (2022)

$$\mathcal{B} = [\![\overbrace{V_{N_1}}^{N_1 \times R}, \ldots, V_{N_Q}, \overbrace{U_{M_1}}^{M_1 \times R}, \ldots, U_{M_P}]\!]$$

Fewer parameters $\longrightarrow$ fewer samples needed

$$\prod_p M_p \prod_q N_q \to R\Big(\sum_p M_p + \sum_q N_q\Big)$$

GENERALIZING TENSOR-ON-TENSOR REGRESSION

Tensor-on-Tensor regression (ToTR):

$$\underbrace{\mathcal{Y}_i}_{N_1 \times \cdots \times N_Q} = \langle \underbrace{\mathcal{X}_i}_{M_1 \times \cdots \times M_P} | \underbrace{\mathcal{B}}_{N_1 \times \cdots \times N_Q \times M_1 \times \cdots \times M_P} \rangle + \mathcal{E}_i$$

ToTR

Lock (2018); Llosa & Maitra (2022)

$$\mathcal{Y}_i = \langle \mathcal{X}_i | \mathcal{B} \rangle + \mathcal{E}_i$$

$$\mathcal{Y}_i \sim N\left(\langle \mathcal{X}_i | \mathcal{B} \rangle, \Sigma_1, \ldots, \Sigma_P\right)$$

CP, Tucker, Tensor Train, Tensor Network, …

Poisson ToTR (PToTR)

Llosa & Dunlavy (2025)

$$\mathcal{Y}_i \sim Poisson\left(\langle \mathcal{X}_i | \mathcal{B} \rangle\right)$$

CP

GMLM

Llosa, Dunlavy, Myers, Lehoucq & Ma (2025)

$$g\left(\mathbb{E}_{f_{\mathcal{Y}}}\left(\mathcal{Y}_i\right)\right) = \langle \mathcal{X}_i | \mathcal{B} \rangle$$

CP

PTOTR: PARAMETER INFERENCE VIA MAXIMUM LIKELIHOOD ESTIMATION

Model

$$\mathcal{Y}_i \overset{ind.}{\sim} Poisson(\langle \mathcal{X}_i | \mathcal{B} \rangle)$$

Loglikelihood

$$\ell(\mathcal{B}) = \sum_{i=1}^{n} \sum_{\boldsymbol{j}} [\mathcal{Y}_{i\boldsymbol{j}} \log(\langle \mathcal{X}_i | \mathcal{B} \rangle_{\boldsymbol{j}}) - \langle \mathcal{X}_i | \mathcal{B} \rangle_{\boldsymbol{j}}]$$

Constraints

$$\mathcal{B} = [\![\boldsymbol{\lambda}; U_1, \ldots, U_l, V_1, \ldots, V_p]\!] > 0$$

New multiplicative update rules extend those in the CP-alternating Poisson regression (CP-APR) algorithm [1]

	Estimation of V_k	**Estimation of U_k**
Alternative expression for Loglikelihood:	$\sum_{i=1}^{n} \mathbf{1}' \left[V_k^* G_{ik} - \mathcal{Y}_{i(k)} * \log(V_k^* G_{ik}) \right] \mathbf{1}$	$\sum_{i=1}^{n} [(\text{ vec } U_k^*)' H_{ik} - (\text{ vec } \mathcal{Y}_i)' * \log ((\text{ vec } U_k^*)' H_{ik})] \mathbf{1}$
Multiplicative update: (non-decreasing loglikelihood)	$\widehat{V}_k^* \leftarrow \widehat{V}_k^* * \left\{ \sum_{i=1}^{n} \left[\left(\mathcal{Y}_{i(k)} \oslash (\widehat{V}_k^* G_{ik}) \right) G_{ik}' \right] \right\} \oslash \left\{ \mathbf{1} (\sum_{i=1}^{n} \boldsymbol{w}_i)' \right\}$	$(\text{ vec } \widehat{U}_k^*) \leftarrow \left\{ \sum_{i=1}^{n} \left[H_{ik} \left(\text{vec}(\mathcal{Y}_i) \oslash \left(H_{ik}' (\text{ vec } \widehat{U}_k^*) \right) \right) \right] \right\} * \text{ vec} \left(\widehat{U}_k^* \oslash \sum_{i=1}^{n} W_i \right).$

[1] Chi & Kolda (2012)

Takeaway: Alternating method possible for more challenging PDF

THE GENERALIZED MULTILINEAR MODEL (GMLM)

Same as in ToTR

$$g\left(\mathbb{E}_{f_{\mathcal{Y}}}\left(\mathcal{Y}_i\right)\right) = \langle \mathcal{X}_i | \mathcal{B} \rangle$$

Link Function

- $g : \mathbb{R}^{\times_{p=1}^{P} M_p} \to \mathbb{R}^{\times_{p=1}^{P} M_p}$
- Applied entry-wise
- e.g., identity, logit, log, inverse,…

Random component

- Exponential dispersion family:

$$\log f_{\mathcal{Y}}(\psi) = \sum_{\boldsymbol{j}} \left(\frac{\mathcal{Y}_{\boldsymbol{j}} \psi_{\boldsymbol{j}} - b(\psi_{\boldsymbol{j}})}{a(\delta_{\boldsymbol{j}})} + C \right)$$

- e.g., normal, binomial, Poisson, …

$\boxed{\textbf{Extends the Generalized Linear Model (GLM) to tensors}}$

Nelder & Wedderburn (1972)

FROM GLM TO GMLM

$$\mathcal{Y}_i \sim P_{f_\mathcal{Y}}(\psi_i), \quad \mathbb{E}(\mathcal{Y}_i) = \mu_i, \quad \mu_i = g^{-1}(\eta_i), \quad \eta_i = \langle \mathcal{X}_i | \mathcal{B} \rangle$$

where $\mathcal{Y}_i, \psi_i, \mu_i, \eta_i$ are all tensors of size $M_1 \times \cdots \times M_P$

Low-rank constraint: $\mathcal{B} = [\![V_1, \ldots, V_{N_Q}, U_1, \ldots, U_{M_P}]\!]$

Loglikelihood: $\ell(\boldsymbol{\theta}) = \sum_{i,\boldsymbol{j}} \log f_\mathcal{Y}(\psi_{i,\boldsymbol{j}})$

Gradient: $\frac{\partial}{\partial \boldsymbol{\theta}} \ell(\boldsymbol{\theta}) = \left[\left(\frac{\partial}{\partial \operatorname{vec}(V_1)} \ell(\boldsymbol{\theta}) \right)^\top \ldots \left(\frac{\partial}{\partial \operatorname{vec}(U_P)} \ell(\boldsymbol{\theta}) \right)^\top \right]^\top$

where $\boldsymbol{\theta} = [\ \operatorname{vec}(V_1)^\top \ldots\ \operatorname{vec}(V_Q)^\top\ \operatorname{vec}(U_1)^\top \ldots\ \operatorname{vec}(U_P)^\top]^\top$

GMLM: MAXIMUM LIKELIHOOD ESTIMATION

Stack all $\mathcal{Y}_i, \mathcal{X}_i$ into $\mathcal{Y} \in \mathbb{R}^{(\times_{p=1}^{P} M_p)\times n}, \mathcal{X} \in \mathbb{R}^{(\times_{q=1}^{Q} N_q)\times n}$. Use $\mathrm{GLM}(X, \boldsymbol{y})$ as follows:

- **Row-based inference** for U_p, where $W = \mathcal{X}_{(Q+1)}(\odot_q V_q)$: **Khatri-Rao Product**

$$U_p[j_p, :] \leftarrow \mathrm{GLM}\Big(X = \mathrm{reshape}\Big[(\bigodot_{k\neq p} U_k) \odot W\Big],\ \boldsymbol{y} = (Y_{(p)}[j_p, :])^\top\Big)$$

- **Factor-based inference** for V_q, where $W_q = \mathcal{X}_{(Q+1,q)}(\odot_{k\neq q} V_k)$:

$$\mathrm{vec}(V_q) \leftarrow \mathrm{GLM}\Big(X = \mathrm{reshape}\Big[(\bigodot_p U_p) \odot W_q\Big],\ \boldsymbol{y} = \ \mathrm{vec}(\mathcal{Y})\Big)$$

1 outer iteration

NOTE: Leverage iteratively reweighted least squares (IRLS) to solve each GLM.

EXAMPLE APPLICATIONS

PARAMETER INFERENCE VALIDATION: SIMULATED DATA

$$g\left(\mathbb{E}_{f_{\mathcal{Y}}}\left(\mathcal{Y}_i\right)\right) = \langle \mathcal{X}_i | \mathcal{B} \rangle$$

Gaussian distribution, identity link:

- 4 different camelid images
- 150 noisy samples of each

- $\mathcal{Y}_i$: 87 x 106 (*pixel height* x *width*)
- $\mathcal{X}_i$: 4 x 3 (*camelid* x *RGB channel*)
- $\mathcal{B}$: 4 x 3 x 87 x 106

Can recover original images from noisy samples using GMLM with sufficiently high rank of parameter tensor.

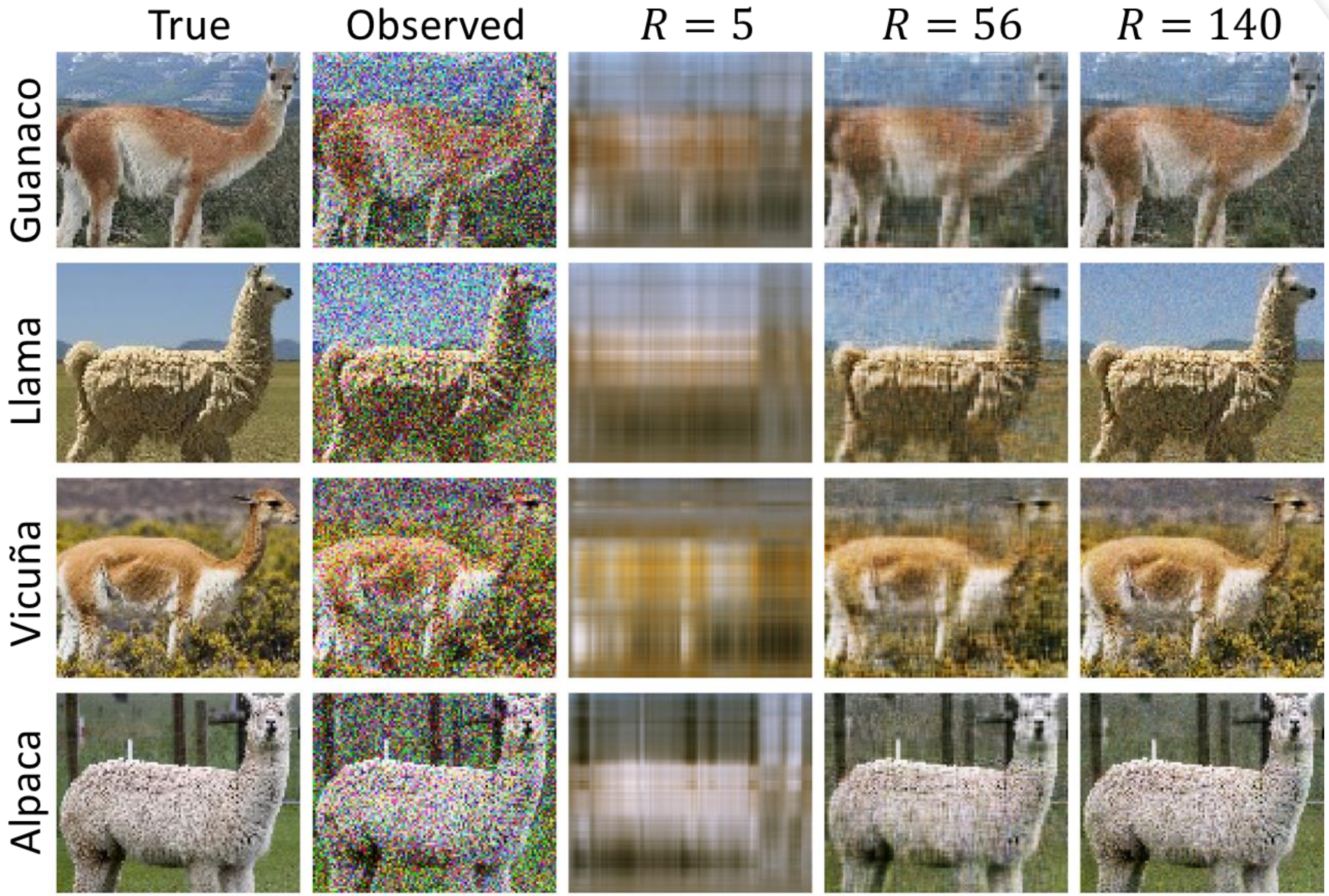


IMAGE CLASSIFICATION (PREDICTION): SIMULATED DATA

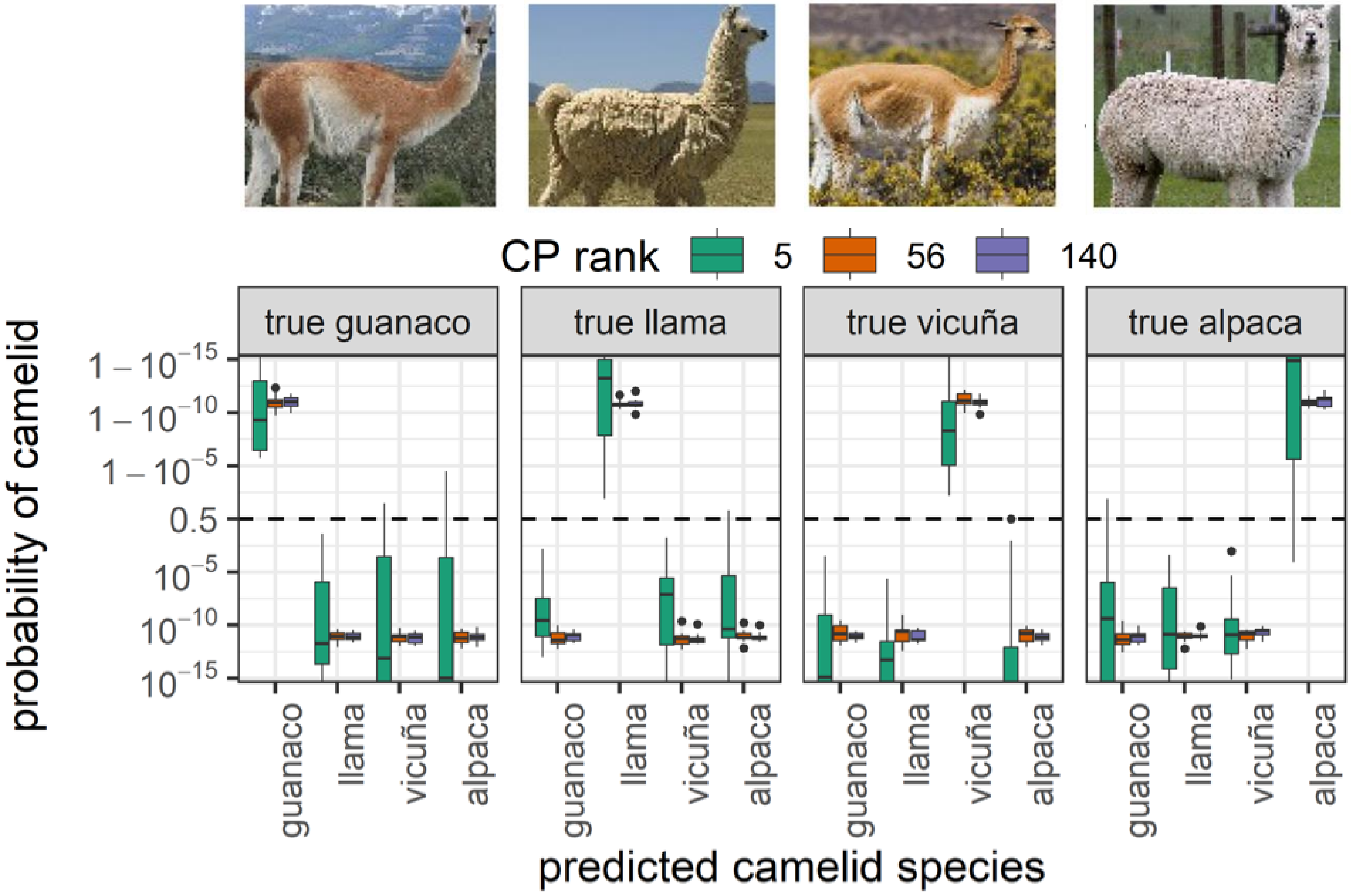
$$g\left(\mathbb{E}_{f_{\mathcal{X}}}\left(\mathcal{X}_i\right)\right) = \langle \mathcal{Y}_i | \mathcal{B} \rangle$$

Binomial distribution, logit link:

- 4 different camelid images
- 150 noisy samples of each

- $\mathcal{Y}_i$: 87 x 106 (*pixel height* x *width*)
- $\mathcal{X}_i$: 4 x 1 (*camelid* x *RGB channel*)
- $\mathcal{B}$: 87 x 106 x 4 x 1

Can accurately predict image classifications from noisy samples using GMLM with sufficiently high rank of parameter tensor.



POSITRON EMISSION TOMOGRAPHY IMAGE RECONSTRUCTION

- **Element-wise regression** (ML-EM) [1]:

$$y_{i_1 i_2} \sim \text{Poisson}\left(\langle K_{i_1 i_2} | B \rangle\right)$$

- Ill-posed without constraints [2]

- **NEW: Poisson-response Tensor-on-tensor regression** (PToTR):

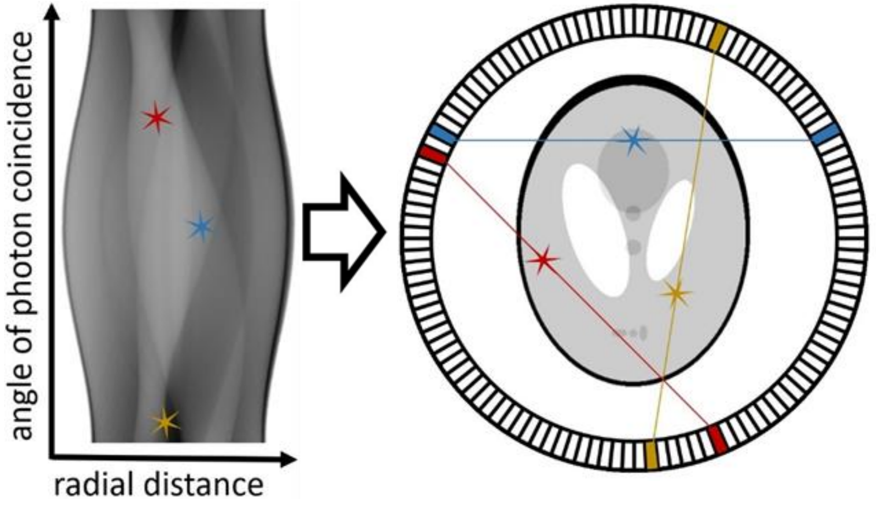
$$Y_{i_1 i_2} \sim \text{Poisson}\left(\langle K_{i_1 i_2} | \mathcal{B} \rangle\right)$$

- Recovers full image and is well-posed due to use of low-rank model for $\mathcal{B}$

$$Y \sim \text{Poisson}(\overbrace{\langle B|\mathcal{K}\rangle}^{\text{Radon transform of } B})$$

Observed sinogram containing counts

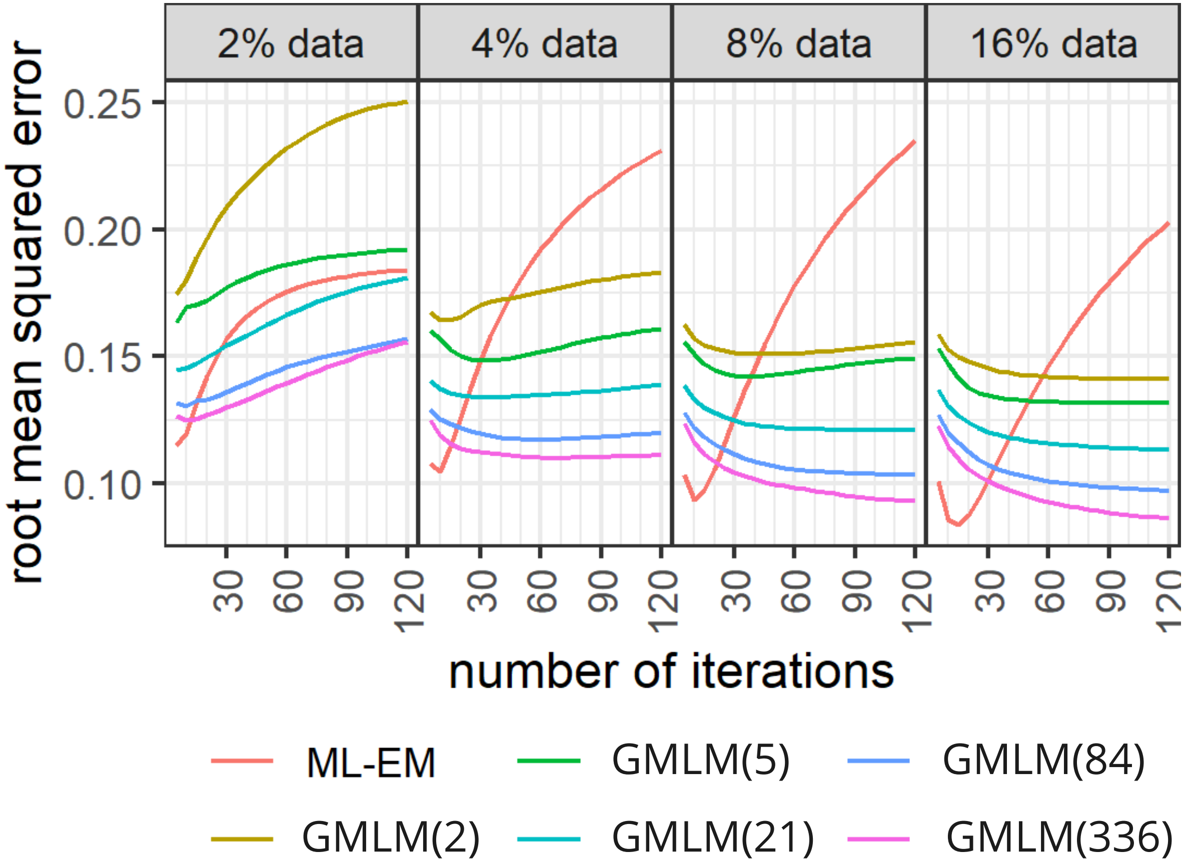
Image to reconstruct

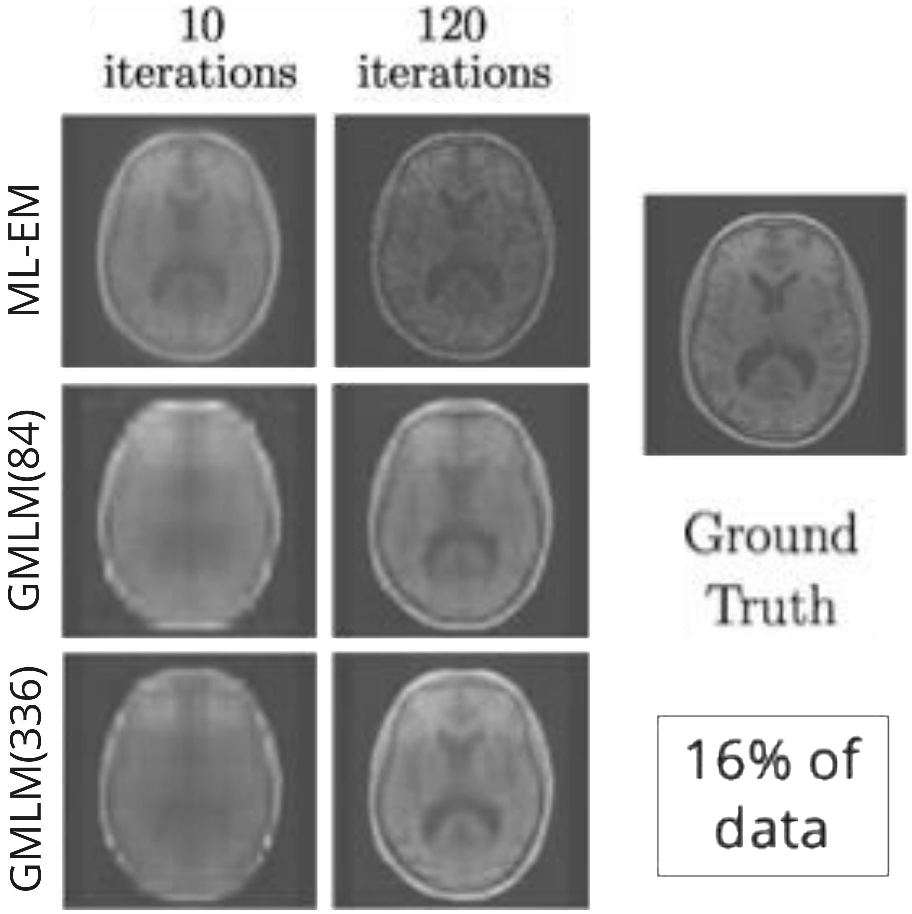


[1] Shepp & Vardi (1982); [2] Snyder, Miller, Thomas & Politte (1987)

POSITRON EMISSION TOMOGRAPHY RECONSTRUCTION

- 4-way tensor data: four MRI measurements on the same subject and scanner *Hawco, et al. (2022)*
- 256 x 256 matrix slices → 256 x 1024 sinograms
- Parameters: ML-EM (no low-rank): **~ 63 million**; GMLM: **~ 63 thousand** (rank-84)





GMLM: GENERALIZED MULTILINEAR MODEL

Carlos Llosa
cjllosa@sandia.gov