Sandia National Laboratories

Exceptional service in the national interest

ADVANCING TENSOR DATA ANALYSIS: GENERALIZED TENSOR-ON-TENSOR REGRESSION MODELS (GTOTRS)

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.

SAND2026-17879PE

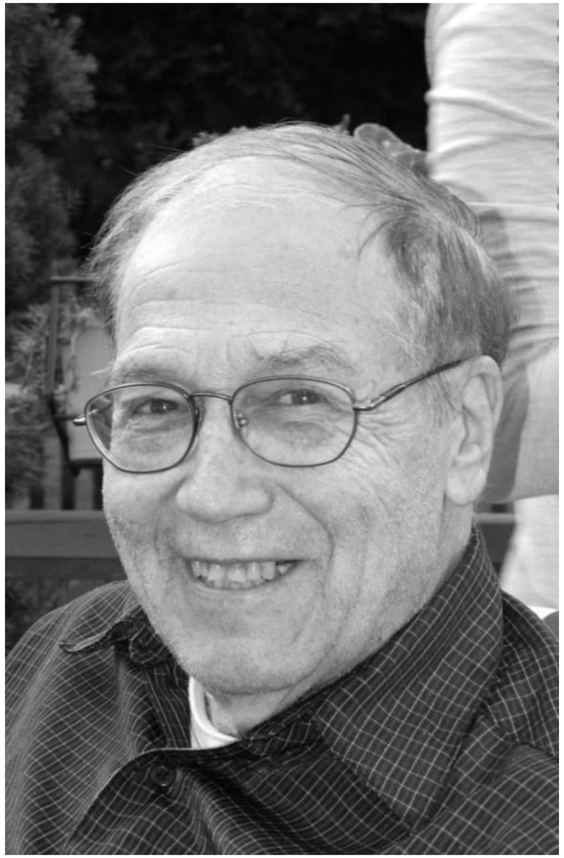
PERSONAL CONNECTION TO PURDUE

My grandfather, Dr. Carlos Llosa, earned his PhD in Agronomy from Purdue University in 1970.

71-9430

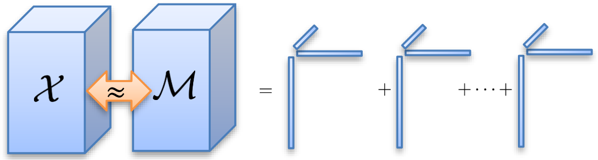
LLOSA BALUARTE, Carlos, 1929-
GENETIC STUDIES OF RATE OF EMERGENCE AND PERCENT EMERGENCE UNDER CONTROLLED TEMPERATURE AND MOISTURE LEVELS IN CROSSES OF WINTER WHEAT (<u>TRITICUM AESTIVUM</u> L.).

Purdue University, Ph.D., 1970
Agronomy

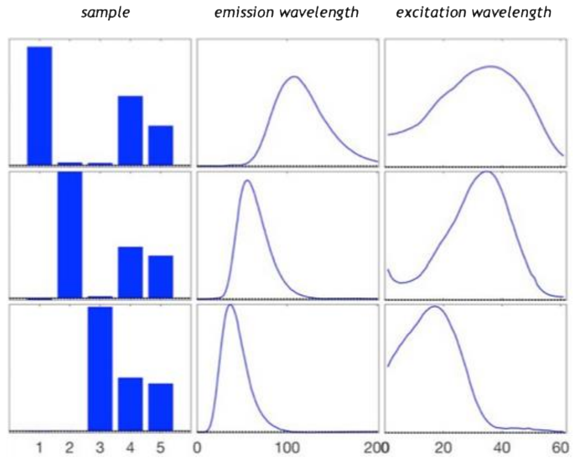


OVERVIEW

Low-Rank Tensor Decompositions



Example: Amino acid analysis

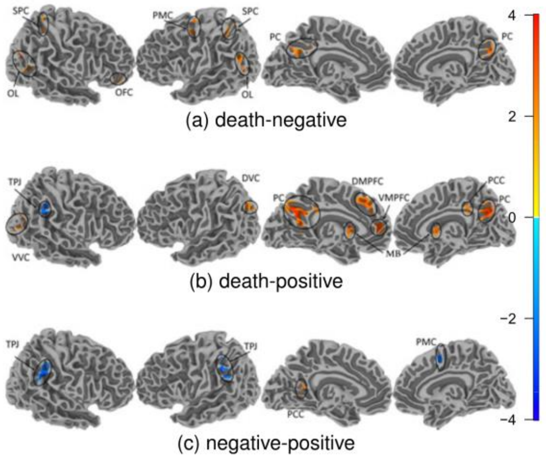


Andersen & Bro (2003)

Tensor-on-Tensor Regression Models

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

Example: Brain activity/suicide risk

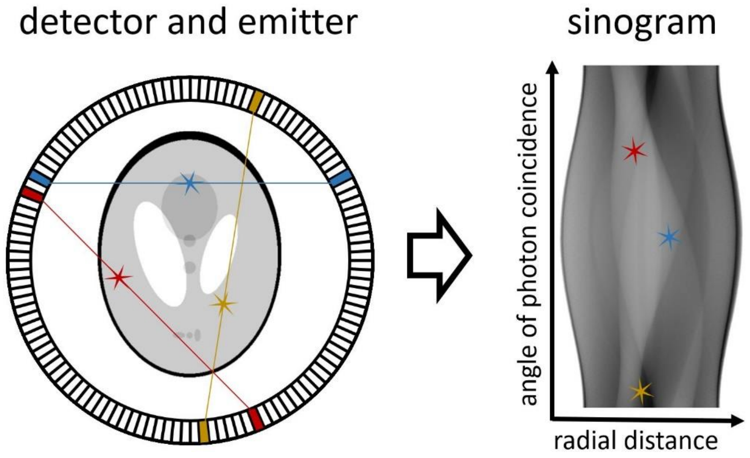


Lock (2018); Llosa & Maitra (2022)

Generalized Tensor-on-Tensor Regression Models (GToTRs)

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

Example: PET image reconstruction

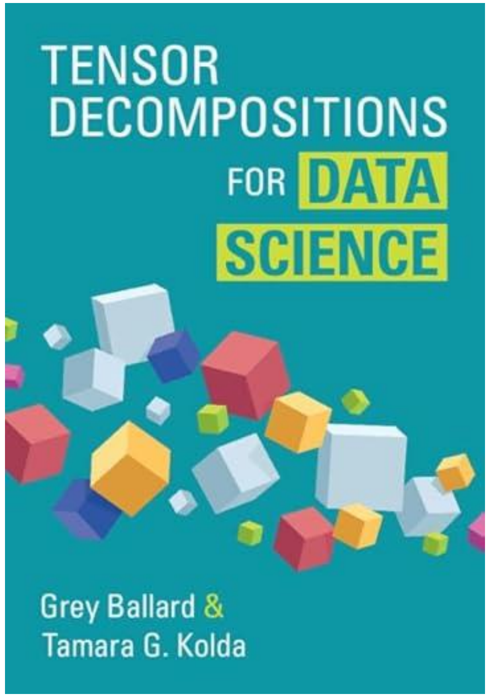


Llosa, Dunlavy, Myers, Lehoucq & Ma (2025)

OUTLINE

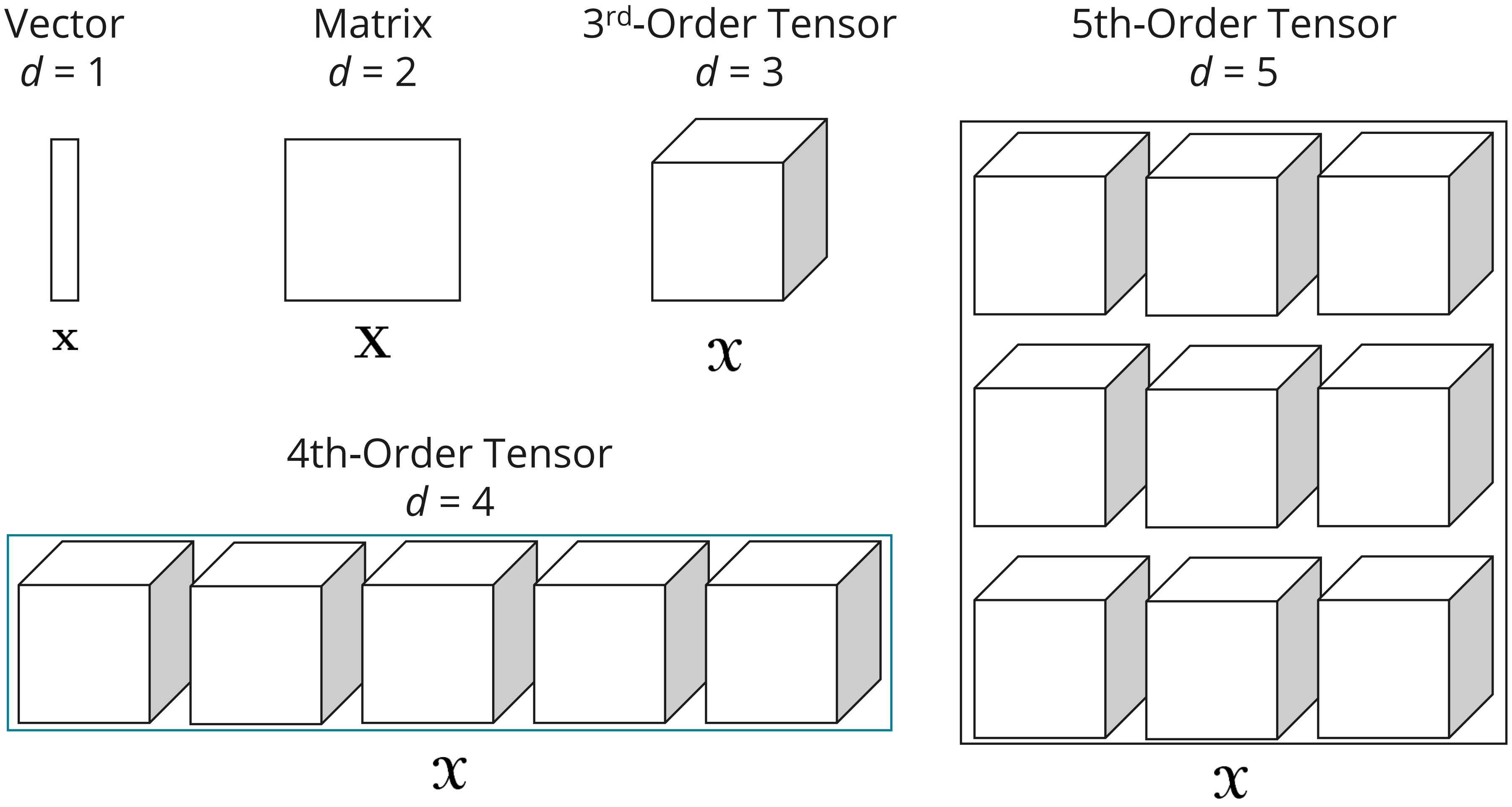
- Low-Rank Canonical Polyadic (CP) Tensor Model
- Tensor-on-Tensor Regression (ToTR)
- Generalized Linear Model (GLM)
- Generalized Tensor-on-Tensor Regression Models (GToTRs)
- Example Applications

New Book!



https://doi.org/10.1017/9781009471664

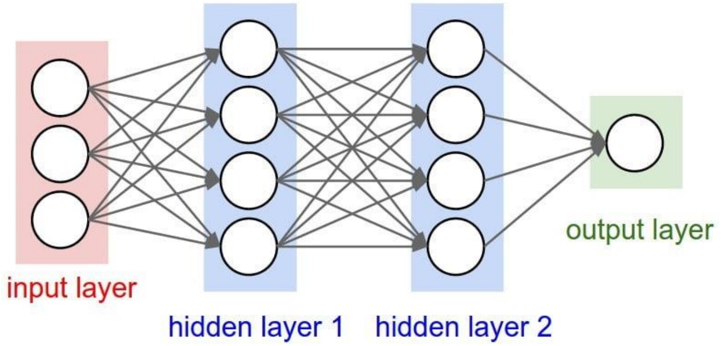
TENSORS: *D*-WAY DATA ARRAYS



We refer to data arrays with 3 or more ways as ***tensors***.

TENSOR DATA AT SANDIA

Machine Learning and Deep Learning

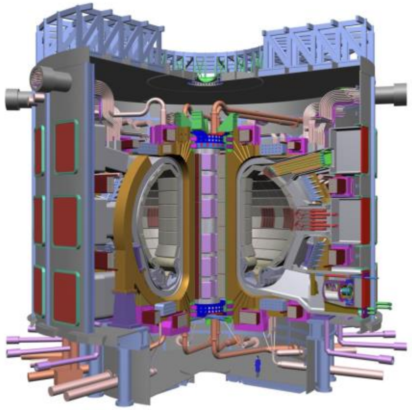


https://www.sandia.gov/machine-and-deep-learning-workshop/

Cybersecurity

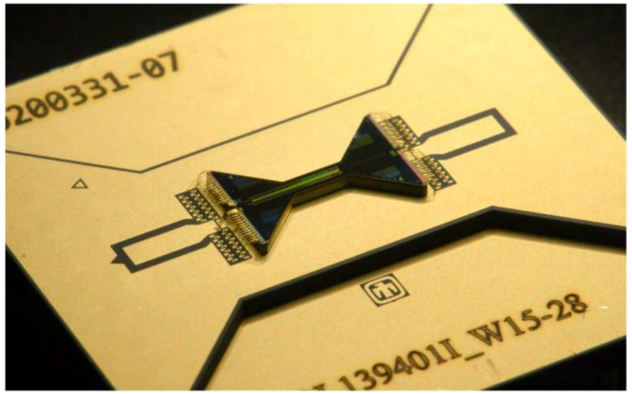
https://www.sandia.gov/cyber/research/

Advanced Simulation and Computing



https://www.sandia.gov/asc/

Quantum Computing

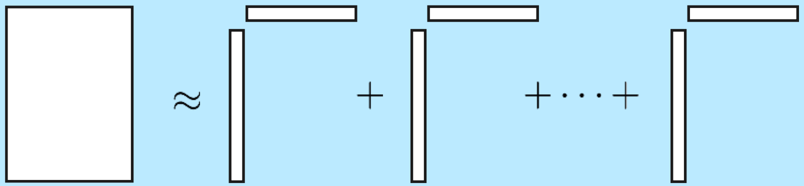


https://www.sandia.gov/quantum/quantum-computing/

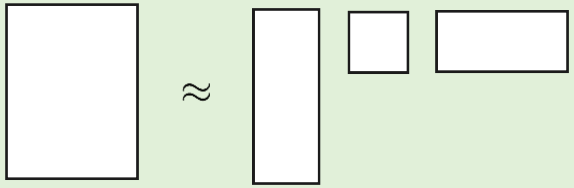
LOW-RANK MATRIX AND TENSOR DECOMPOSITIONS

Low-Rank Matrix Decompositions

Viewpoint 1: Sum of vector outer products, useful for interpretation



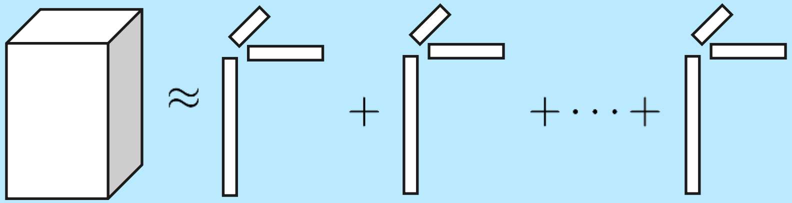
Viewpoint 2: High-variance subspaces, useful for compression



Singular value decomposition (SVD), principal component analysis (PCA), eigenvalue decomposition (EVD), etc.

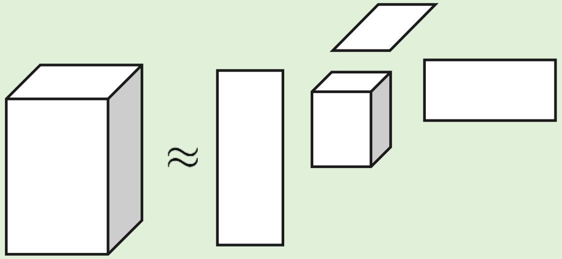
Low-Rank Tensor Decompositions

Sum of d-way vector outer products, useful for interpretation



Canonical Polyadic (CP) Decomposition

Project onto high-variance subspaces to reduce dimensionality

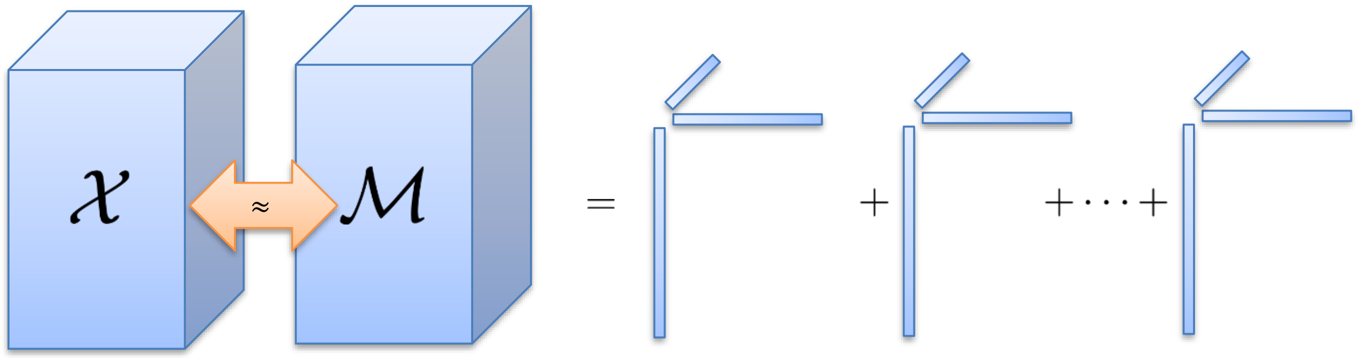


Tucker Decomposition

Other models for compression include hierarchical Tucker, tensor train, tensor ring, tensor network, etc.

Kolda & Ballard (2025)

GENERALIZING CP DECOMPOSITIONS FOR D-WAY TENSORS



$$= [\![A^{(1)}, A^{(2)}, \ldots, A^{(d)}]\!] = [\![\{A^{(k)}\}_1^d]\!]$$

Multi-index for *d*-way tensors: $\boldsymbol{j} = (j_1, j_2, \ldots, j_d) \rightarrow x_{\boldsymbol{j}} = \mathcal{X}(j_1, j_2, \ldots, j_d)$

CP-ALS, CP-OPT

Carroll & Chang (1970); Harshman (1970); Acar, Dunlavy & Kolda (2011)

$$\min_{\{A^{(k)}\}_1^d} \sum_{\boldsymbol{j}} (x_{\boldsymbol{j}} - m_{\boldsymbol{j}})^2$$

$$x_{\boldsymbol{j}} \overset{indep.}{\sim} N(m_{\boldsymbol{j}}, \sigma^2)$$

CP-APR, CP-POPT

Chi & Kolda (2012); Ranadive & Baskaran (2021)

$$\min_{\{A^{(k)}\}_1^d} \sum_{\boldsymbol{j}} m_{\boldsymbol{j}} - x_{\boldsymbol{j}} \log(m_{\boldsymbol{j}})$$

$$x_{\boldsymbol{j}} \overset{indep.}{\sim} Poisson(m_i)$$

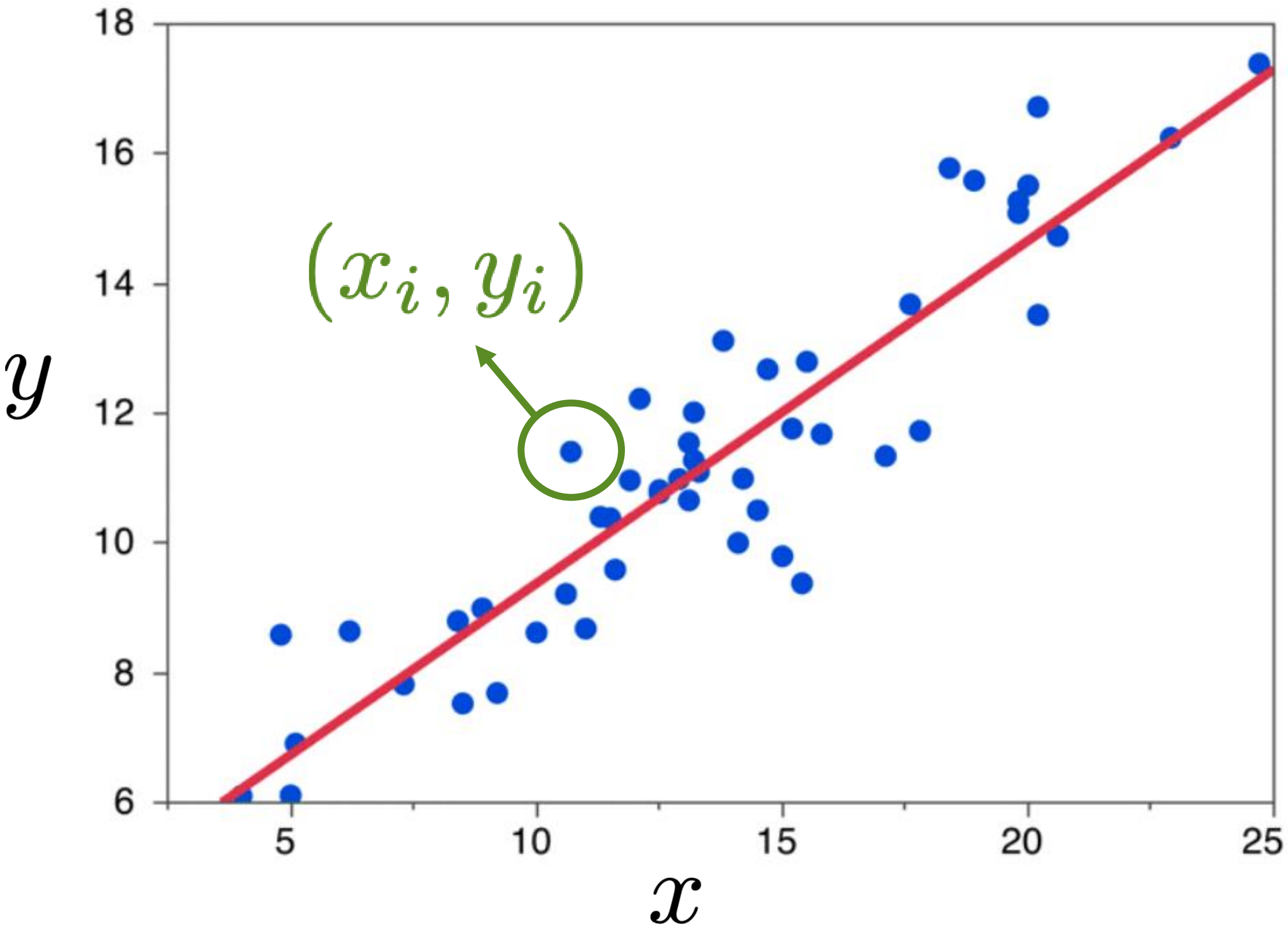
GCP

Hong, Kolda & Duersch (2020); Kolda & Hong (2020)

$$\min_{\{A^{(k)}\}_1^d} \sum_{\boldsymbol{j}} f(x_{\boldsymbol{j}}, m_{\boldsymbol{j}})$$

$$x_{\boldsymbol{j}} \overset{indep.}{\sim} P_f(m_i)$$

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

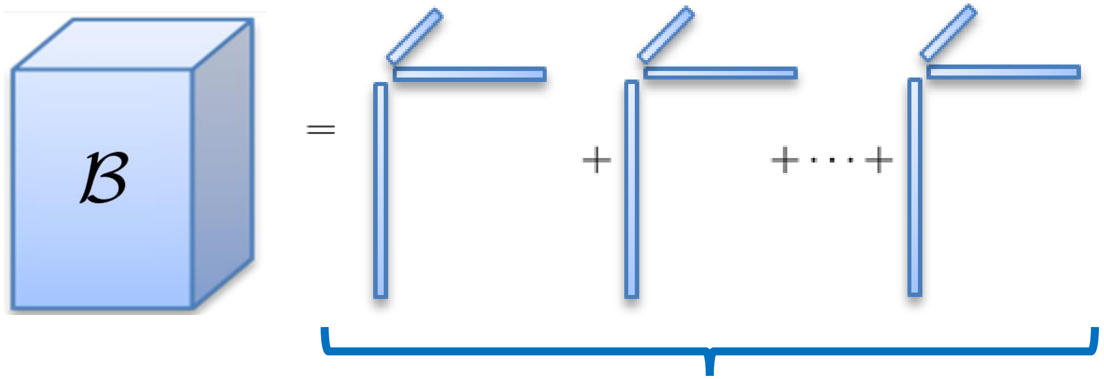
Tensor-on-Tensor regression (ToTR):

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

$\boxed{i = 1, 2, \ldots, n \longrightarrow \text{so many samples needed!}}$

Low-rank Canonical Polyadic (CP) Model:

Rank-1 tensor



R summands

Lock (2018); Llosa & Maitra (2022)

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

Fewer parameters ⟶ fewer samples needed

$$\prod_p M_p \prod_q N_q \rightarrow R\left(\sum_p M_p + \sum_q N_q\right)$$

GENERALIZING TENSOR-ON-TENSOR REGRESSION

Tensor-on-Tensor regression (ToTR):

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

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

GToTR

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

GENERALIZED LINEAR MODELS (GLMS)

$$y_i \sim P_{f_Y}(\psi_i), \quad \mathbb{E}(y_i) = \mu_i, \quad \mu_i = g^{-1}(\eta_i), \quad \eta_i = \boldsymbol{b}^\top \boldsymbol{x}_i$$

- y: response (independent)
- $\boldsymbol{x}$: covariate
- μ: mean response
- g: link function, relating random component and linear predictor
- $\boldsymbol{b}$: model parameters
- n samples $\rightarrow X = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n]\,; \boldsymbol{y} = [y_1, \ldots, y_n]$

$$\boxed{\mathrm{GLM}(X, \boldsymbol{y})}$$

Nelder & Wedderburn (1972)

GENERALIZED LINEAR MODELS (GLMS)

$$y_i \sim P_{f_Y}(\psi_i), \quad \mathbb{E}(y_i) = \mu_i, \quad \mu_i = g^{-1}(\eta_i), \quad \eta_i = \boldsymbol{b}^\top \boldsymbol{x}_i$$

Loglikelihood:

$$\ell(\boldsymbol{b}) = \sum_{i=1}^{n} \log f_Y(\psi_i)$$

We can now find the model parameters using, e.g., *maximum likelihood estimation*.

Gradient:

$$\frac{\partial}{\partial \boldsymbol{b}} \ell(\boldsymbol{b}) = \sum_{i=1}^{n} \frac{\partial}{\partial \boldsymbol{b}} \log f_Y(\psi_i)$$

$$= \sum_{i=1}^{n} \underbrace{\left[\frac{\partial}{\partial \boldsymbol{b}} \eta_i\right]}_{\boldsymbol{x}_i} \underbrace{\left[\frac{\partial}{\partial \eta_i} \mu_i\right] \left[\frac{\partial}{\partial \mu_i} \psi_i\right] \left[\frac{\partial}{\partial \psi_i} \log f_Y(\psi_i)\right]}_{\text{weight } i}$$

Nelder & Wedderburn (1972)

THE GENERALIZED TENSOR-ON-TENSOR REGRESSION MODEL (GTOTR)

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, …

Extends the Generalized Linear Model (GLM) to tensors

Nelder & Wedderburn (1972)

FROM GLM TO GTOTR

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

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

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

GTOTR: 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

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

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

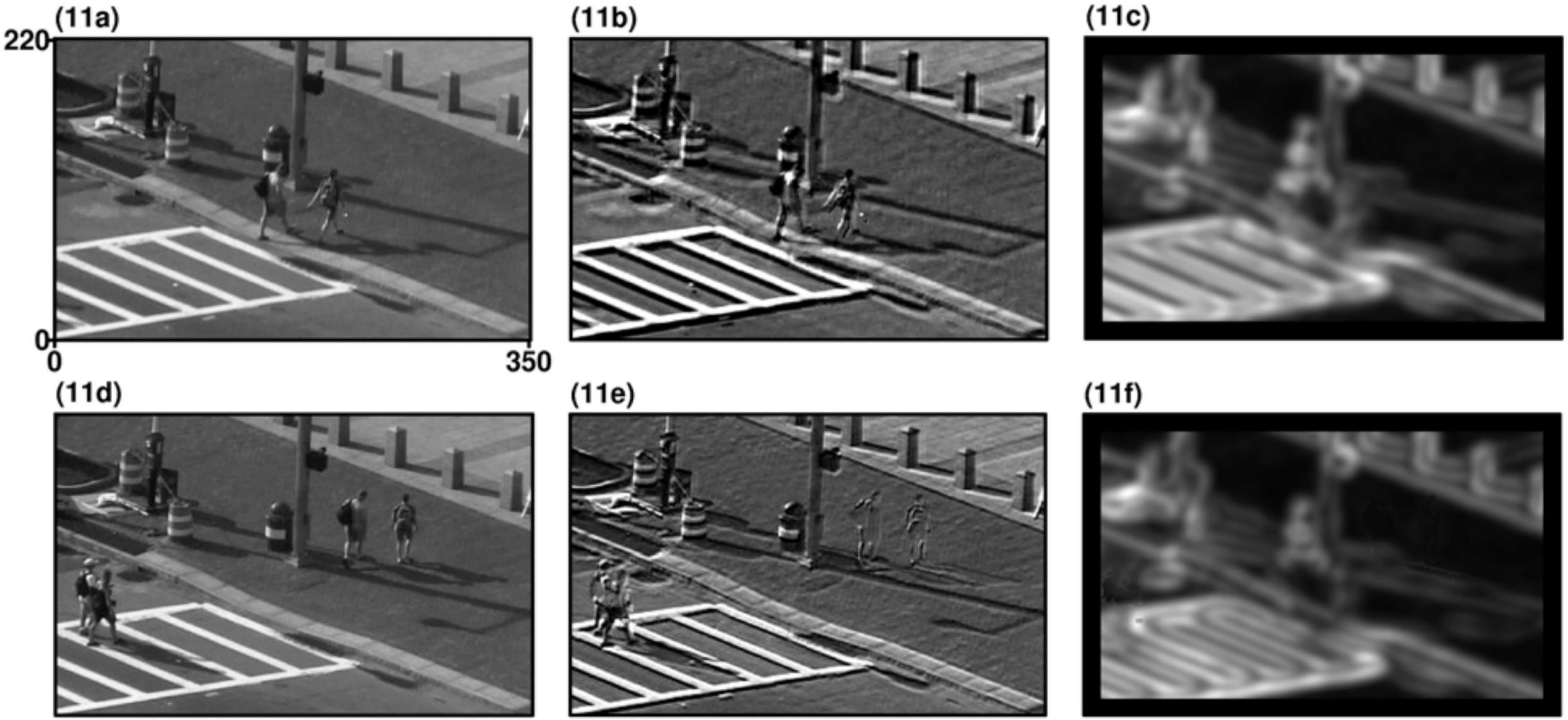
$$\mathrm{vec}(V_q) \leftarrow \mathrm{GLM}\Big(X = \mathrm{reshape}\Big[\big(\bigodot_{p} U_p\big) \odot W_q\Big],\ \boldsymbol{y} = \ \mathrm{vec}(\mathcal{Y})\Big)$$

1 outer iteration

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

EXAMPLE APPLICATIONS

MOTIVATING APPLICATION: REMOTE SENSING



Jodoin, Konrad, Saligrama, Veilleux-Gaboury (2008); Simonson & Ma (2009)

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 GToTR with sufficiently high rank of parameter tensor.

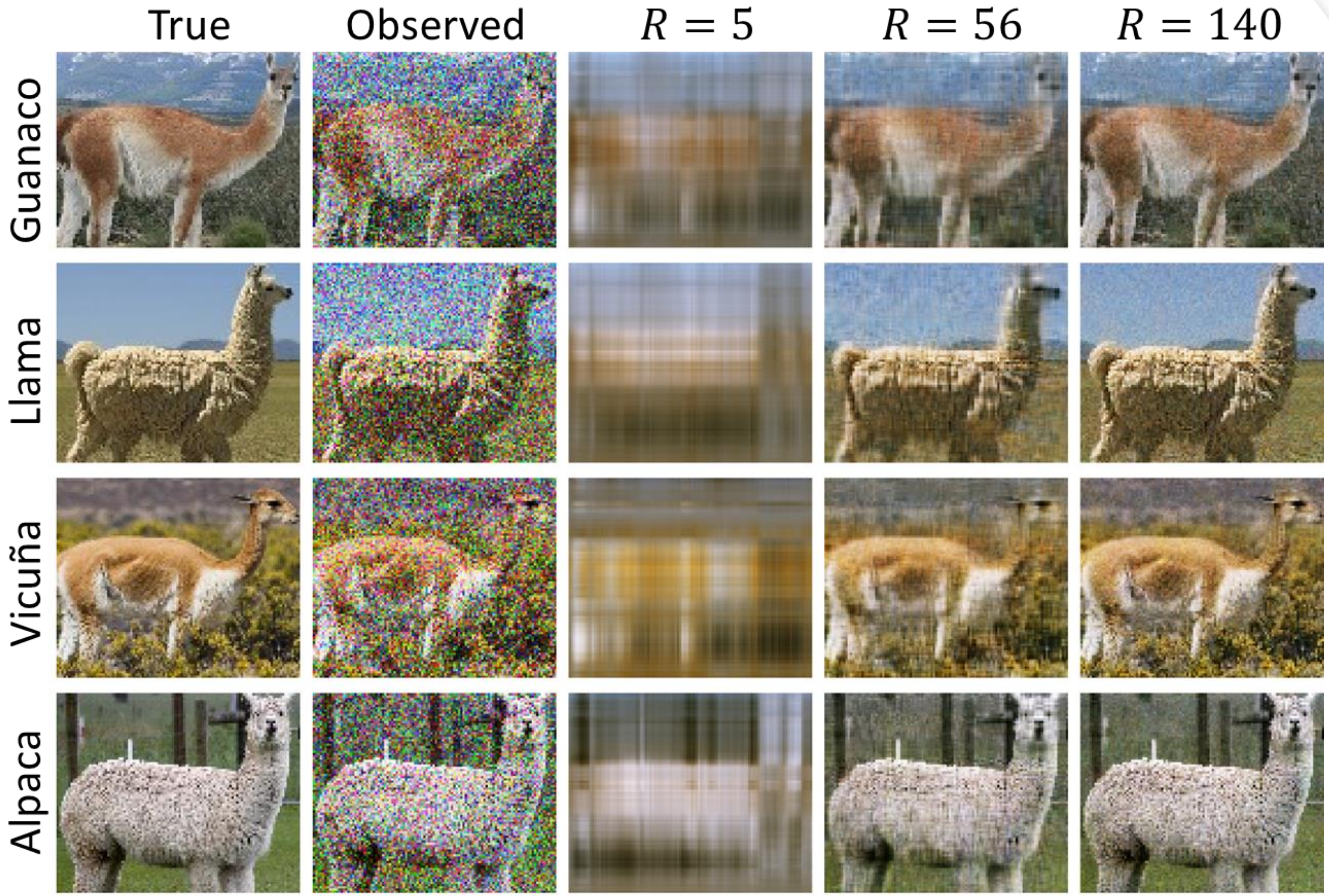


IMAGE CLASSIFICATION (PREDICTION): SIMULATED DATA

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

Binomial distribution, logit link:

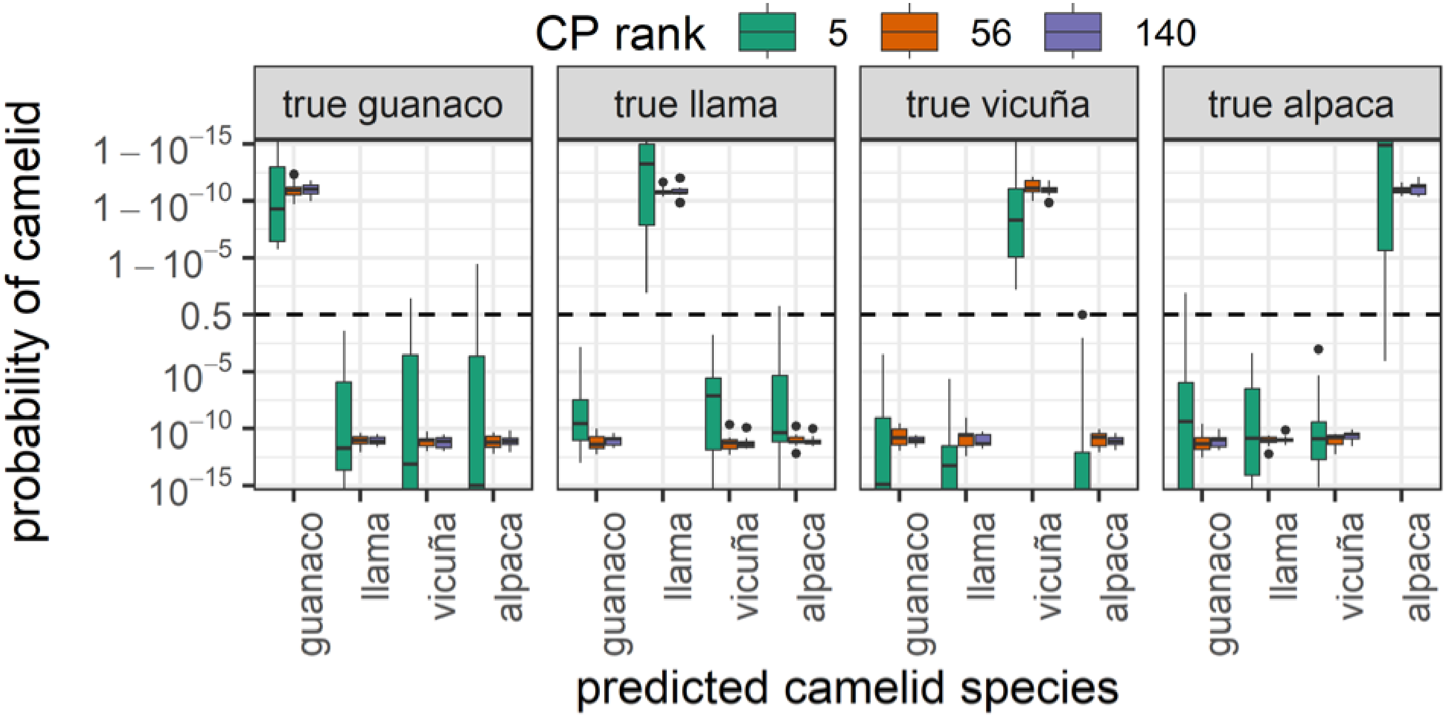
- 4 different camelid images
- 150 noisy samples of each
- $\mathcal{Y}_i$: 4 x 1

  (*camelid* x *RGB channel*)
- $\mathcal{X}_i$: 87 x 106

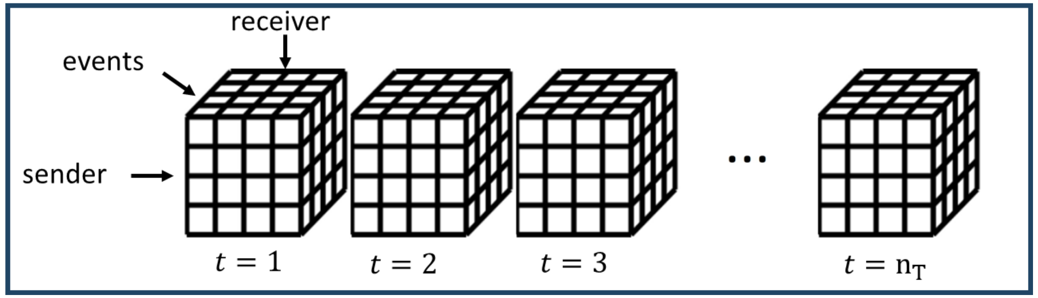
  (*pixel height* x *width*)
- $\mathcal{B}$: 87 x 106 x 4 x 1

Flip Xs and Ys

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



Forecasting Model for the ICEWS Database



Temporal prediction of dyadic relationships

ICEWS database [1]

- Countries as receivers and senders
- Events such as threats or aid
- **Can we predict future relations?**

[1] O'Brien, 2010.

$$\mathcal{Y}_t \overset{indep.}{\sim} \text{Poisson}(\langle \mathcal{X}_t | \mathcal{B} \rangle)$$

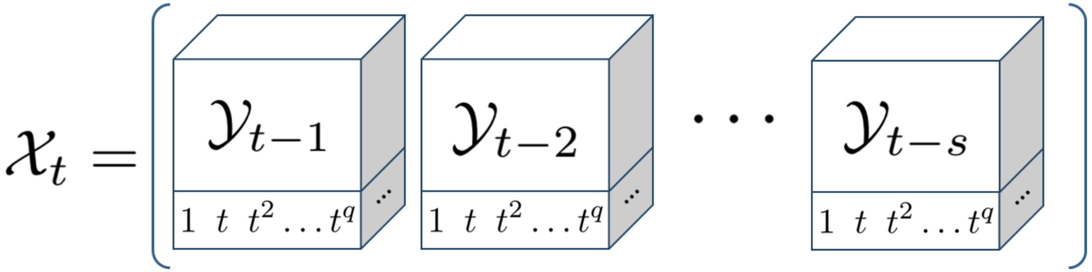
- **AR(1) with no trend** $\quad \mathcal{X}_t = \mathcal{Y}_{t-1}$

$\mathcal{Y}_t(i_1, i_2, i_3)$ = # times action i_3 was taken by country i_1 on country i_2 at week t.

$$\mathcal{Y}_t \overset{indep.}{\sim} \text{Poisson}(\langle \mathcal{Y}_{t-1} | \mathcal{B} \rangle)\ *$$

$\mathcal{B}(i_1, i_2, i_3, j_1, j_2, j_3)$ = Effect that the **previous** i_3 event from i_1 towards i_2 has on **current** j_3 event from j_1 towards j_2.

- **AR(s) with q-th order polynomial trend:**



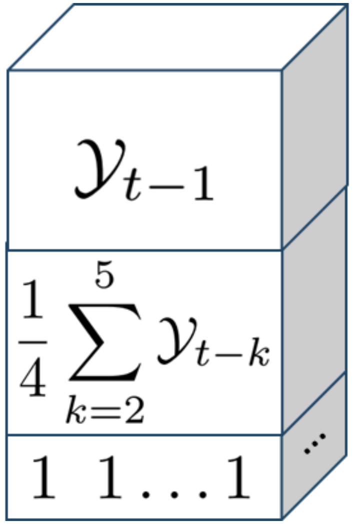
*Not well defined without accounting for trend

Autoregressive Model for the ICEWS Database

Data selected as in [1]:

- Weekly data from 2004 to mid-2014
- 25 countries
- 4 quad classes (type of events)

$$\mathcal{X}_t =$$



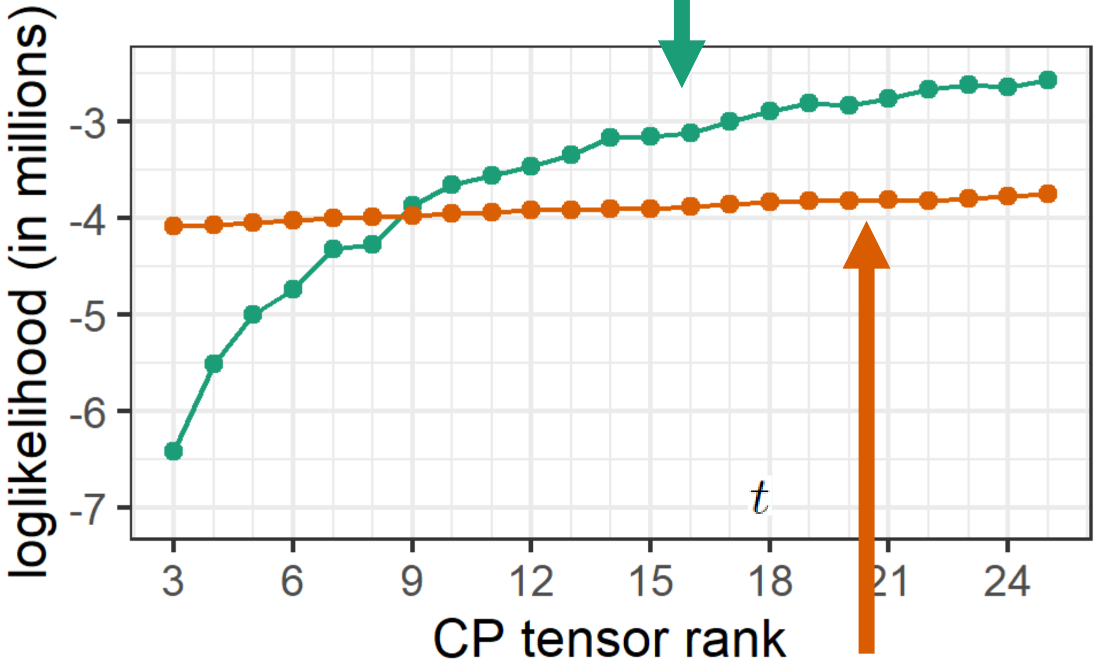
previous week

month before that

mean trend

Poisson ToTR :

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



Gaussian ToTR:

$$\mathcal{Y}_i \overset{indep.}{\sim} N(\langle \mathcal{X}_i | \mathcal{B} \rangle, \Sigma_1, \ldots, \Sigma_p)$$

Our Poisson model fits the count data better!

[1] Hoff, 2015

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):

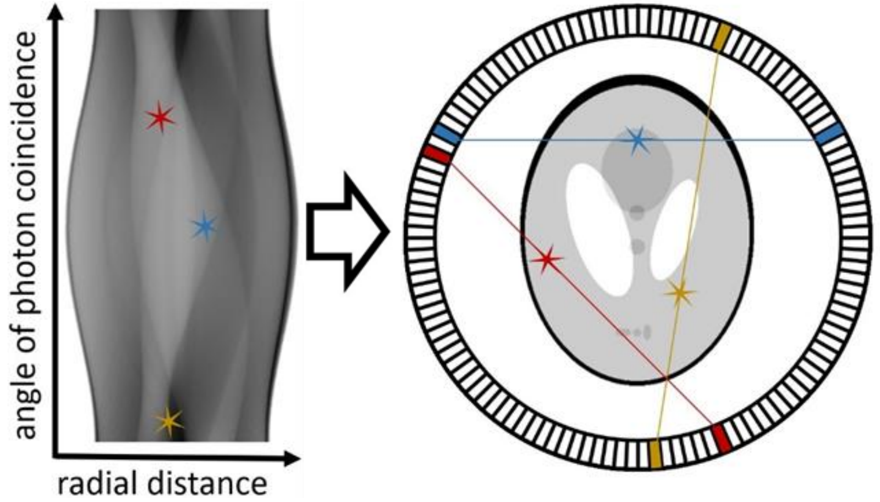
$$Y_{i_1 i_2} \sim \text{Poisson}\left(\langle K_{i_1 i_2} | \mathcal{B} \rangle\right)$$

- 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

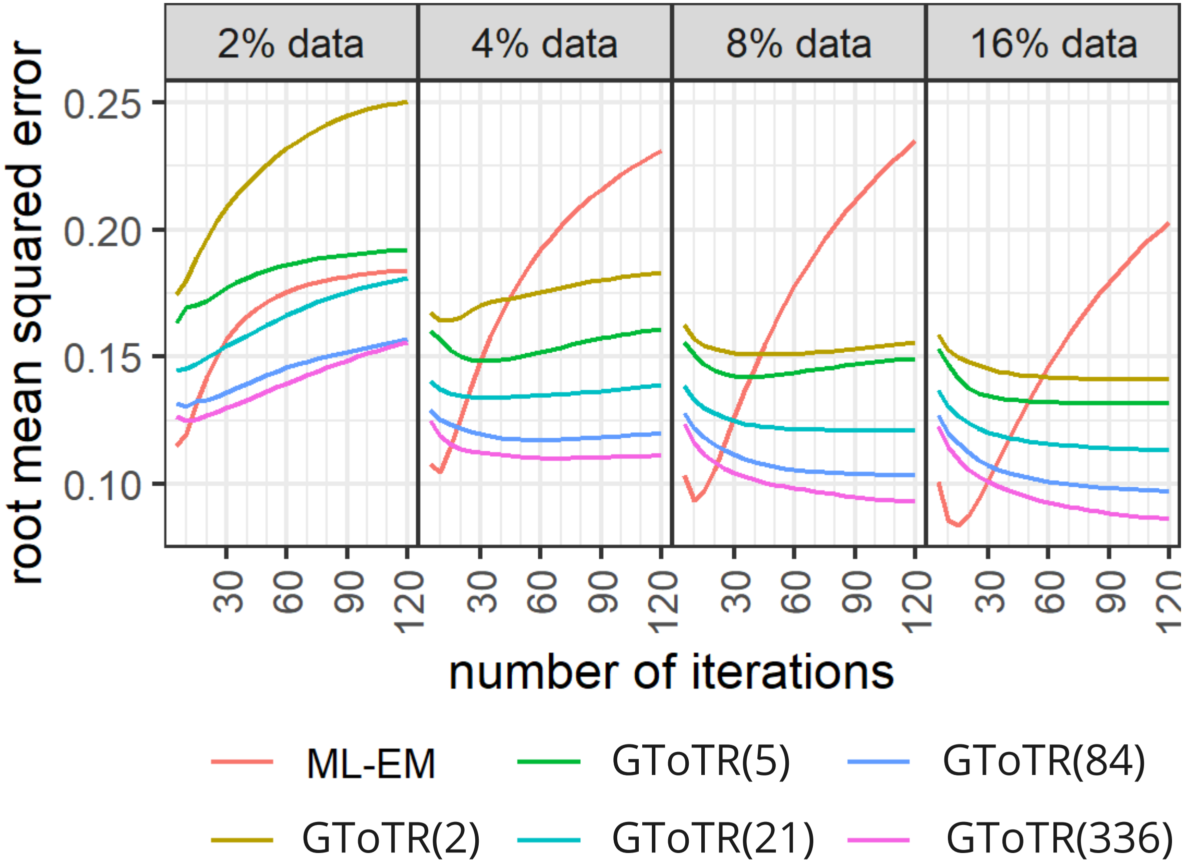
Image to reconstruct

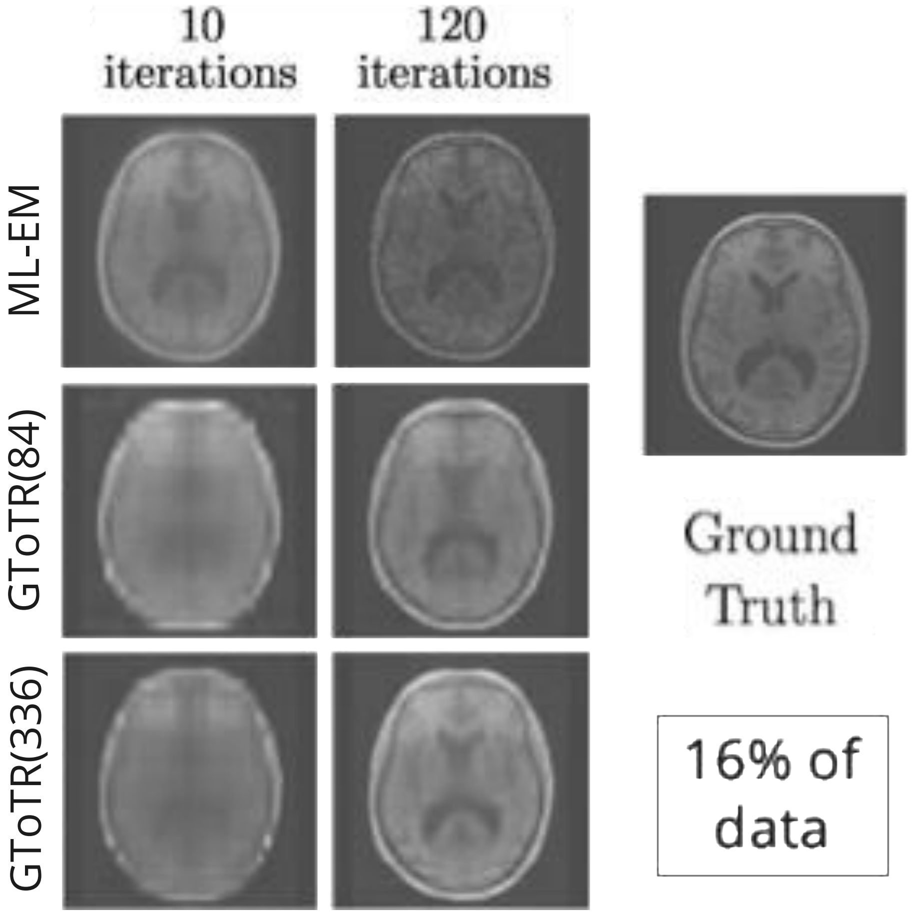


[1] Shepp & Vardi (1982); [2] Snyder, Miller, Thomas & Politte (1987)

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**; GToTR: **~ 63 thousand** (rank-84)





TENSOR-VARIATE ANALYSIS OF VARIANCE (TANOVA)

TANOVA is a special case of ToTR ($\mathcal{X}_i$ are indicator tensors) that generalizes ANOVA

General Model	Definition	Special Case (Indicator $\mathcal{X}_i$)
Linear regression (LR)	$y_i \overset{indep.}{\sim} N(\langle \boldsymbol{x}_i \vert \boldsymbol{\beta} \rangle , \sigma^2)$	**ANOVA**
Multivariate LR	$\boldsymbol{y}_i \overset{indep.}{\sim} N(\langle \boldsymbol{x}_i \vert B \rangle , \Sigma)$	**MANOVA**
Tensor regression	$y_i \overset{indep.}{\sim} N(\langle \mathcal{X}_i \vert \mathcal{B} \rangle , \sigma^2)$	**Factorial designs**
Tensor-on-tensor regression	$\mathcal{Y}_i \overset{indep.}{\sim} N(\langle \mathcal{X}_i \vert \mathcal{B} \rangle , \Sigma_1, \ldots, \Sigma_p)$	**TANOVA**

$\mathcal{Y}_i \overset{ind.}{\sim} Poisson(\langle \mathcal{X}_i \vert \mathcal{B} \rangle)$ → **Poisson TANOVA**

$g\left(\mathbb{E}_{f_{\mathcal{Y}}}\left(\mathcal{Y}_i\right)\right) = \langle \mathcal{X}_i \vert \mathcal{B} \rangle$ → **Generalized TANOVA**

Llosa & Maitra (2022)

PTANOVA AND CHANGE-POINT DETECTION

Poisson-tensor change-point detection

$\mathcal{Y}_t$ changes in mean at time τ_1

$$\mathcal{Y}_t \sim \left\{ \begin{array}{ll} \text{Poisson}(\mathcal{M}_1) & t = 1, \ldots, \tau_1 \\ \text{Poisson}(\mathcal{M}_2) & t = \tau_1 + 1, \ldots, n_T \end{array} \right\}$$

- Mean before change-point: $\mathcal{M}_1$
- Mean after change-point: $\mathcal{M}_2$
- Change-point location: τ_1

Equivalent PTANOVA formulation

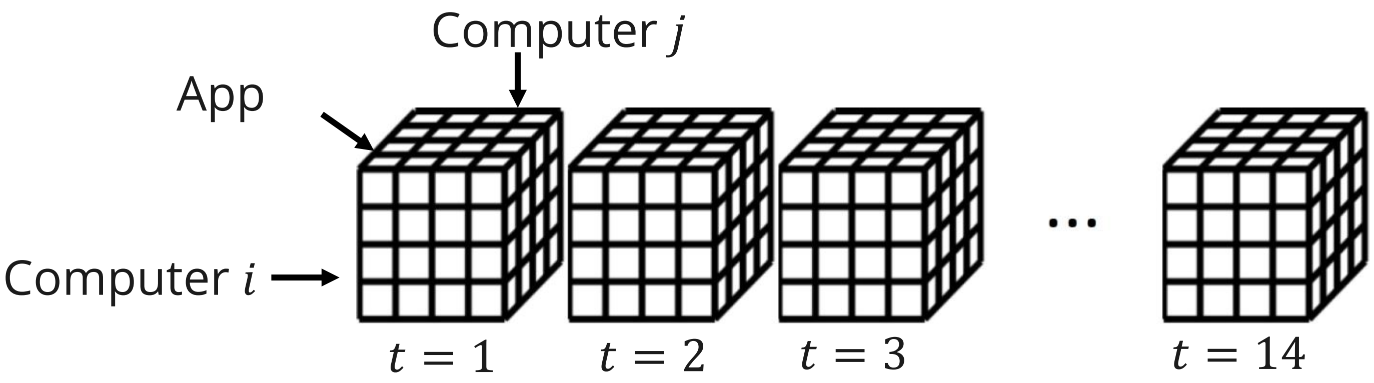
$$\mathcal{Y}_t \sim \text{Poisson}(\langle \boldsymbol{x}_t | \mathcal{B} \rangle)$$

$$\boldsymbol{x}_t = \left\{ \begin{array}{ll} (1,0)' & t = 1, \ldots, \tau_1 \\ (0,1)' & t = \tau_1 + 1, \ldots, n_T \end{array} \right\}$$

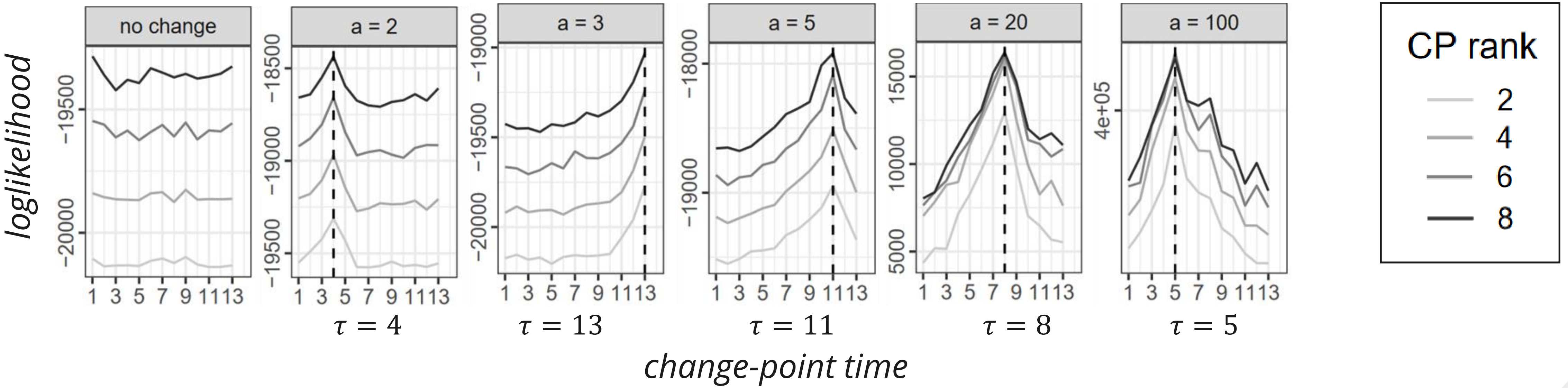
- Mean before change-point: $\langle (1,0)' | \mathcal{B} \rangle$
- Mean after change-point: $\langle (0,1)' | \mathcal{B} \rangle$
- Change-point location: τ_1

Originally solved using PToTR → Now we can perform parameter inference via GToTR

CHANGE-POINT DETECTION: SIMULATED DATA



- 10 computers communicating via 15 apps over 14 time-steps: 14 count tensors of size 10 x 10 x 15.
- Event occurs at time τ that changes communication pattern:
  - All communications generated independently from Poisson(λ), $\lambda = 5$.
  - One app usage patterns changes to Poisson($\lambda \times a$) after time τ, where $a = 1, 2, 3, 5, 20, 100$.



SUMMARY AND NEXT STEPS

- **GToTR: Generalized for Tensor-on-Tensor Regression Models**
  - *Inference efficiency:* leverages low-rank decomposition of parameter tensor
  - *Well-posed model:* low-rank parameter tensor avoids need for regularization
  - *Reasonable computation:* leverages IRLS for alternating parameter inference subproblems

- **Future Directions**
  - GMLM *vs. GCP:* GToTR reduces to tensor decomposition with $\mathcal{X}_i = 1$
  - *Accelerating computation:* randomized GLM solvers (has been done in Sandia for GCP)
  - *Other low-rank tensor formats:* Tucker, tensor trains, FCTNs, etc.
  - *Uncertainty quantification:* hypothesis testing and confidence intervals
  - *Error bounds:* Fisher Information, minimax and concentration bounds
  - *Applications:* temporal prediction, change-point detection, imaging reconstruction

REFERENCES

- Acar, Dunlavy & Kolda (2011), A Scalable Optimization Approach for Fitting Canonical Tensor Decompositions, *J. Chemometrics*, 25(2):67–86.
- Andersen & Bro (2003), Practical aspects of PARAFAC modeling of fluorescence excitation-emission data, *J. Chemometrics*, 17(4):200–215.
- Carroll & Chang (1970), Analysis of individual differences in multidimensional scaling via an n-way generalization of 'Eckart–Young' decomposition, Psychometrika, 35(3):283–319.
- Chi & Kolda (2012), On tensors, sparsity, and nonnegative factorizations, SIAM Journal on Matrix Analysis and Applications, 33(4):1272–129
- Harshman (1970), Foundations of the PARAFAC procedure: Models and conditions for an explanatory multi-modal factor analysis, *UCLA Working Papers in Phonetics*, 16(84).
- Hong, Kolda & Duersch (2020), Generalized canonical polyadic tensor decomposition, *SIAM Review*, 62(1):133–163.
- Jodoin, Konrad, Saligrama & Veilleux-Gaboury (2008), Motion Detection with an Unstable Camera, *IEEE Intl. Conference on Image Processing.*
- Kolda & Ballard (2025), *Tensor Decompositions for Data Science*, Cambridge University Press.
- Kolda & Hong (2020), Stochastic gradients for large-scale tensor decomposition, *SIAM Journal on Mathematics of Data Science*, 2(4):1066–1095.
- Llosa & Dunlavy (2025), Poisson-response Tensor-on-Tensor Regression and Applications, *in preparation*.
- Llosa, Dunlavy, Myers, Lehoucq & Ma (2025), Generalized Multilinear Models, *in preparation*.
- Llosa & Maitra (2024), Reduced-Rank Tensor-on-Tensor Regression and Tensor-Variate Analysis of Variance, IEEE Transactions on PAMI, 45(2):2282–2296.
- Lock (2018), Tensor-on-tensor regression," *Journal of Computational and Graphical Statistics*, 27(3):638–647.
- Nelder & Wedderburn (1972), Generalized Linear Models. Journal of the Royal Statistical Society: Series A (General), 135:370–384.
- Ranadive & Baskaran (2021), An All–at–Once CP Decomposition Method for Count Tensors, *IEEE High Performance Extreme Computing Conference.*
- Shepp & Vardi (1982), Maximum Likelihood Reconstruction for Emission Tomography, *IEEE Transactions on Medical Imaging*, 1(2):113–122.
- Simonson & Ma (2009), Robust Real-Time Change Detection in High Jitter, Technical Report SAND2009-5546, Sandia National Laboratories.
- Snyder, Miller, Thomas & Politte (1987), Noise and Edge Artifacts in Maximum-Likelihood Reconstructions for Emission Tomography, *IEEE Transactions on Medical Imaging*, 6(3):228–238.

GTOTR: GENERALIZED TENSOR-ON-TENSOR REGRESSION MODEL

Carlos Llosa
cjllosa@sandia.gov