CISUC
CENTRO DE INFORMÁTICA E SISTEMAS DA UNIVERSIDADE DE COIMBRA

bcam
basque center for applied mathematics

On the Identifiability of Tensor Ranks via Prior Predictive Matching

Eliezer da Silva · eliezer.silva@uc.pt
Arto Klami · arto.klami@helsinki.fi
Diego Mesquita · diego.mesquita@fgv.br
Iñigo Urteaga · iurteaga@bcamath.org

1. Why this problem matters

Tensor factorizations are powerful, but choosing latent dimensions is still often handled by grid search or heuristics.

Question: Can low-order moments of observed tensor data reveal the ranks directly?

Answer: Yes for **CP**, **Tensor Train**, and **Tensor Ring**; no for standard **Tucker** under our prior-predictive moment-matching framework.

- Build moment equations from the prior predictive distribution.

$$C_S := \mathbb{E}\left[\mathrm{Cov}(Y_{\mathbf{i}}, Y_{\mathbf{j}} \mid \mathbf{i}, \mathbf{j} \text{ share exactly } S)\right],$$

$$C_S = \sum_{\emptyset \neq S' \subseteq S} v_{S'}, \qquad v_S = \sum_{S' \subseteq S} (-1)^{|S|-|S'|} C_{S'}.$$

- Turn them into a **log-linear system**.
- Eliminate nuisance prior moments.
- Check whether the reduced system uniquely determines the ranks.

Big picture pipeline

Observed tensor $\mathbf{Y}$
↓
Estimate shared-index moments $\widehat{C}_S$
↓
Möbius inversion $\Rightarrow \hat{v}_S$
↓
Reduced log-linear system
↓
Closed-form rank estimate

2. Main theoretical statement

Write unknowns as logarithms of ranks and prior moments:

$$\mathbf{x} = (\log r_p,\ \log \mu_p^2,\ \log \sigma_p^2), \qquad \mathbf{A}\mathbf{x} = \mathbf{b}.$$

The ranks are identifiable iff nuisance terms can be eliminated, leaving a reduced system in log-ranks only:

$$\mathbf{A}_{\mathrm{red}}\mathbf{x}_r = \mathbf{b}_{\mathrm{red}}, \qquad \mathrm{rank}(\mathbf{A}_{\mathrm{red}}) = \dim(\mathbf{x}_r).$$

Interpretation: The model's latent interaction topology determines whether moment matching isolates the ranks.

Model-by-model verdict

Model	Result	Reason
Tucker	No	Symmetry leaves the reduced system underdetermined.
CP / PARAFAC	Yes	The global component rank separates in moment ratios.
Tensor Train	Yes	Local chain neighborhoods yield bond-rank identities.
Tensor Ring	Yes	Cyclic equations form an invertible circulant system.

Worked Example: CP Rank from Pure Moments

For an order-M CP/PARAFAC model,

$$\eta_{i_1,\ldots,i_M} = \sum_{s=1}^{r} \prod_{k=1}^{M} \theta^{(k)}_{i_k,s}, \qquad \mathbb{E}[\theta^{(k)}] = \mu_k, \quad \mathrm{Var}(\theta^{(k)}) = \sigma_k^2.$$

The squared prior predictive mean is

$$(\mathbb{E}[Y])^2 = r^2 \prod_{k=1}^{M} \mu_k^2.$$

The pure interaction terms obtained from shared-index covariances satisfy

$$v_{\{p\}} = r\sigma_p^2 \prod_{k \neq p} \mu_k^2,$$

and, for $p \neq q$,

$$v_{\{p,q\}} = r\sigma_p^2\sigma_q^2 \prod_{k \neq p,q} \mu_k^2.$$

Taking logs gives one row per monomial:

$$\log v_{\{p\}} = \log r + \log \sigma_p^2 + \sum_{k \neq p} \log \mu_k^2,$$

$$\log v_{\{p,q\}} = \log r + \log \sigma_p^2 + \log \sigma_q^2 + \sum_{k \neq p,q} \log \mu_k^2.$$

Now combine the observable moments:

$$\frac{v_{\{p,q\}}(\mathbb{E}[Y])^2}{v_{\{p\}}v_{\{q\}}} = \frac{\left(r\sigma_p^2\sigma_q^2 \prod_{k \neq p,q} \mu_k^2\right)\left(r^2 \prod_{k=1}^{M} \mu_k^2\right)}{\left(r\sigma_p^2 \prod_{k \neq p} \mu_k^2\right)\left(r\sigma_q^2 \prod_{k \neq q} \mu_k^2\right)} = r.$$

Therefore,

$$\boxed{r = \frac{v_{\{p,q\}}(\mathbb{E}[Y])^2}{v_{\{p\}}v_{\{q\}}}}$$

and all nuisance prior moments cancel.

3. Closed-form estimators

Pure interaction terms v_S (Möbius-inverted covariances) lead to compact rank formulas:

CP / PARAFAC

$$R = \frac{v_{\{p,q\}}\,\mathbb{E}[Y]^2}{v_{\{p\}}\,v_{\{q\}}}.$$

Tensor Train

$$r_p = \frac{v_{\{p,p+1\}}\,v_{\{p-1,p+2\}}}{v_{\{p+1\}}\,v_{\{p-1,p,p+2\}}}.$$

Tensor Ring

$$\psi_p = \log(\mathbb{E}[Y]^2) + \log v_{\{p,p+1\}} - \log v_{\{p\}} - \log v_{\{p+1\}},$$

$$\begin{bmatrix} 1 & -1 & 0 & \cdots & 0 & 1 \\ 1 & 1 & -1 & \cdots & 0 & 0 \\ 0 & 1 & 1 & \cdots & 0 & 0 \\ \vdots & & \ddots & \ddots & & \vdots \\ -1 & 0 & 0 & \cdots & 1 & 1 \end{bmatrix} \begin{bmatrix} \log r_1 \\ \vdots \\ \log r_p \\ \vdots \\ \log r_M \end{bmatrix} = \begin{bmatrix} \psi_1 \\ \vdots \\ \psi_p \\ \vdots \\ \psi_M \end{bmatrix},$$

The system given by $\mathbf{C}\log\mathbf{r} = \psi$ has a solution because matrix C is circulant with a non-zero determinant.

Practical message: Only observed first/second moments and a small family of exact-sharing covariance sets are needed.

Intuition for the estimators

- Shared-index covariances reveal which latent factors are forced to match across two entries.
- Möbius inversion removes lower-order overlaps and extracts the pure interaction term v_S.
- In CP, TT, and TR these pure terms cancel prior means and variances, leaving only rank parameters.
- In Tucker, symmetric paths survive cancellation, so the reduced system stays singular.

Finite-sample stabilization:

- Mean normalization by $\widehat{(\mathbb{E}[Y])}^2$,
- Regularization for noisy denominators,
- Bootstrap estimation over exact-sharing patterns.

4. Empirical results

We highlight the CP rank-recovery experiment (TT and TR behave similarly; see paper).

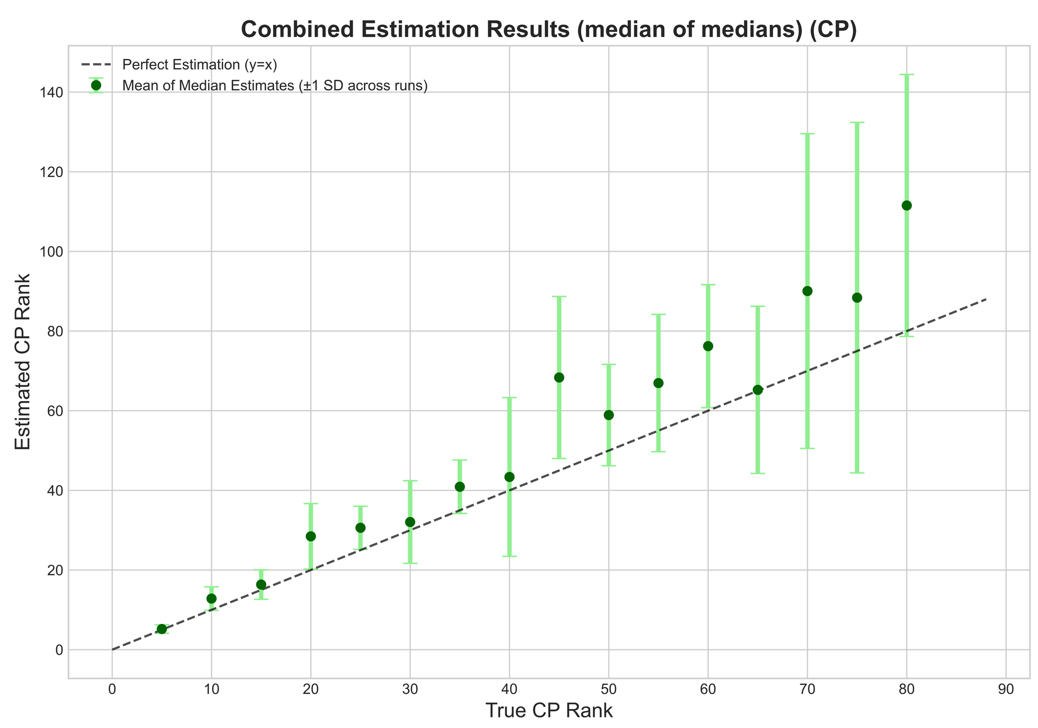

Figure 1

5. Take-home messages

- **Structure determines identifiability.**
- **Second moments are enough** for CP, TT, and TR under the paper's assumptions.
- **Tucker is a cautionary example:** symmetry can destroy rank identifiability.
- **The estimators are cheap and interpretable,** avoiding repeated model refits.

We turn prior predictive moment matching into a clean algebraic test for tensor-rank identifiability. The answer depends on topology: CP, TT, and TR admit explicit recovery from low-order moments, while Tucker does not.

Paper & code

Repository:
https://github.com/iurteagalab/tensor-ranks

Scan for paper + code

1290 UNIVERSIDADE D COIMBRA · REPÚBLICA PORTUGUESA · fct Fundação para a Ciência e a Tecnologia · LASI Intelligent Systems Associate Laboratory · HELSINGIN YLIOPISTO · ikerbasque Basque Foundation for Science