NeurIPS 2018
Dec. 2–8, 2018

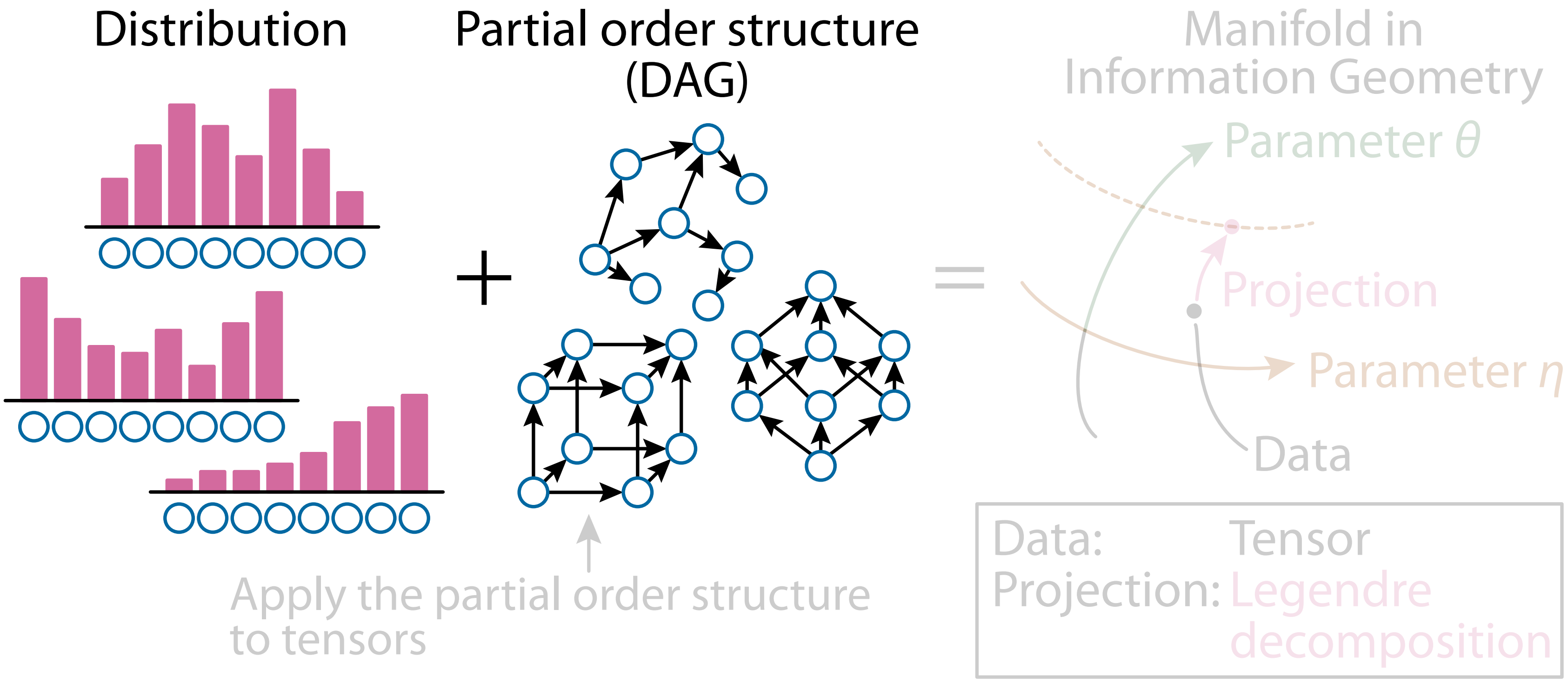
Inter-University Research Institute Corporation / Research Organization of Information and Systems
National Institute of Informatics

PRESTO
SAKIGAKE

Legendre Decomposition for Tensors

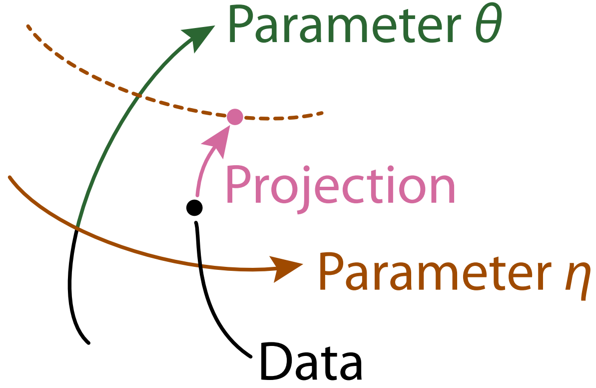
Mahito Sugiyama (National Institute of Informatics, JST PRESTO)
Hiroyuki Nakahara (RIKEN CBS)
Koji Tsuda (The University of Tokyo, NIMS, RIKEN AIP)

Our Approach



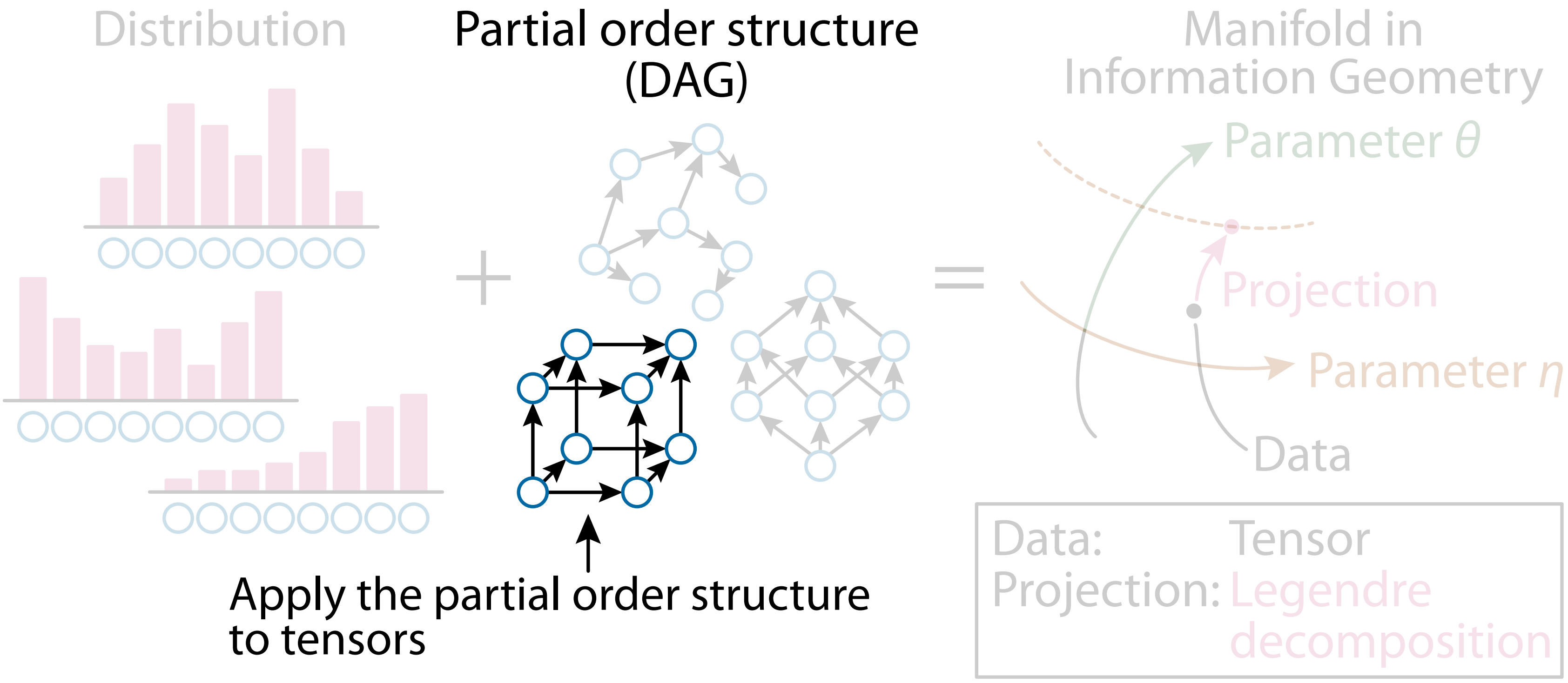
Our Approach

Manifold in
Information Geometry

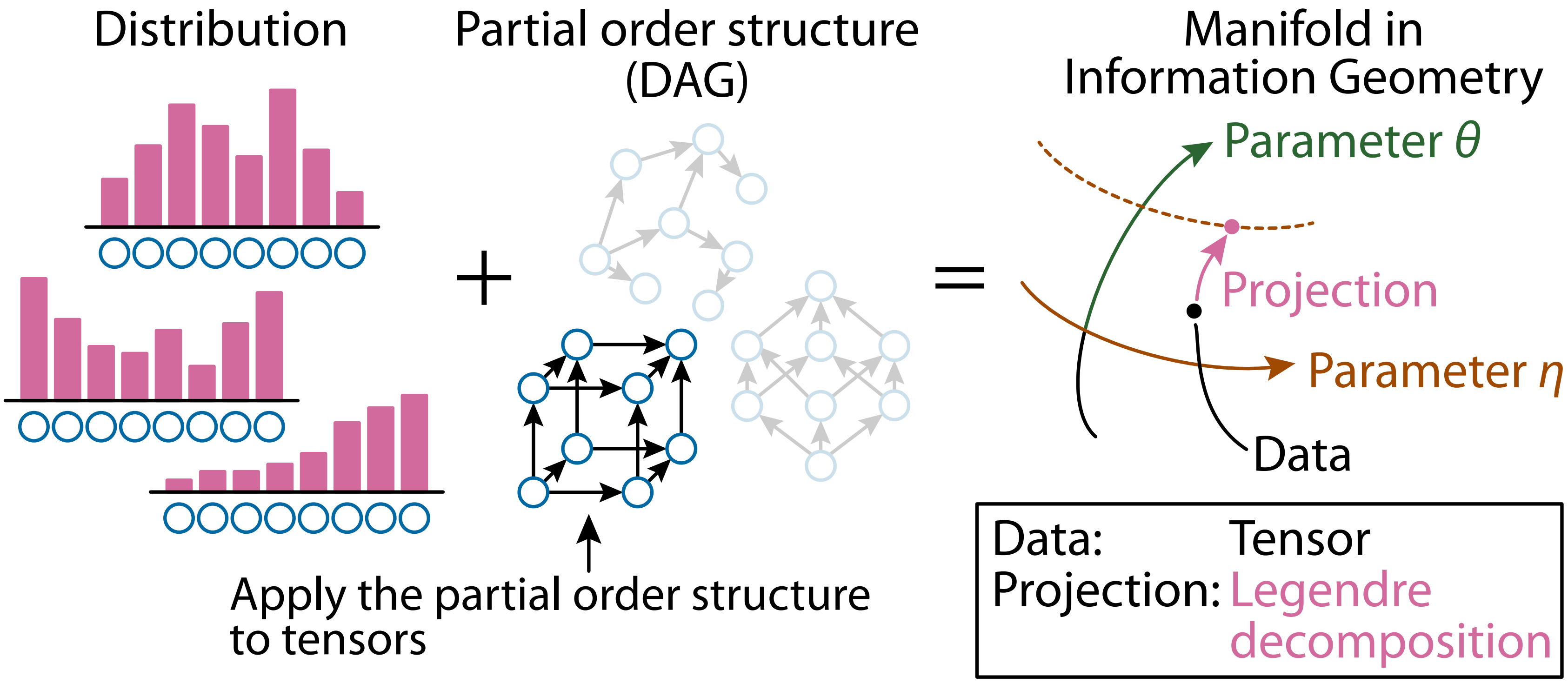


Data: Tensor
Projection: Legendre decomposition

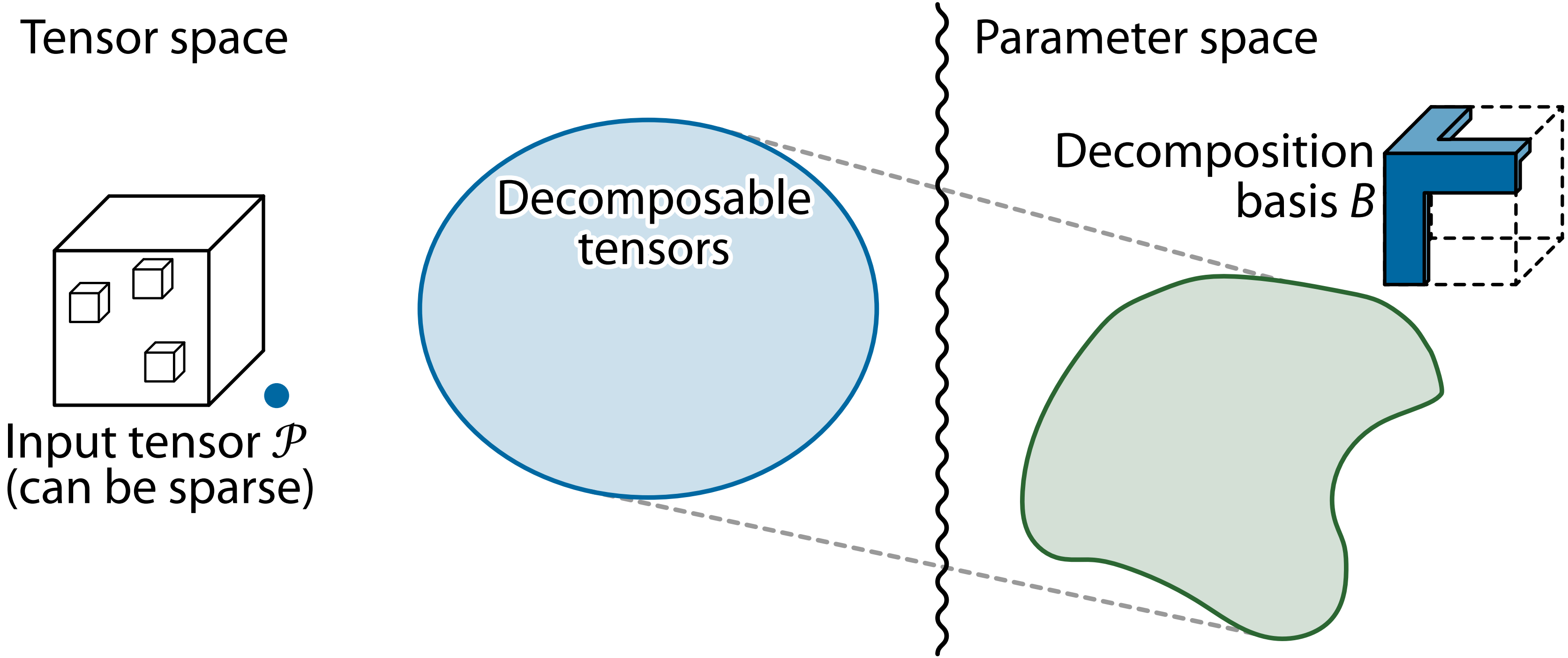
Our Approach



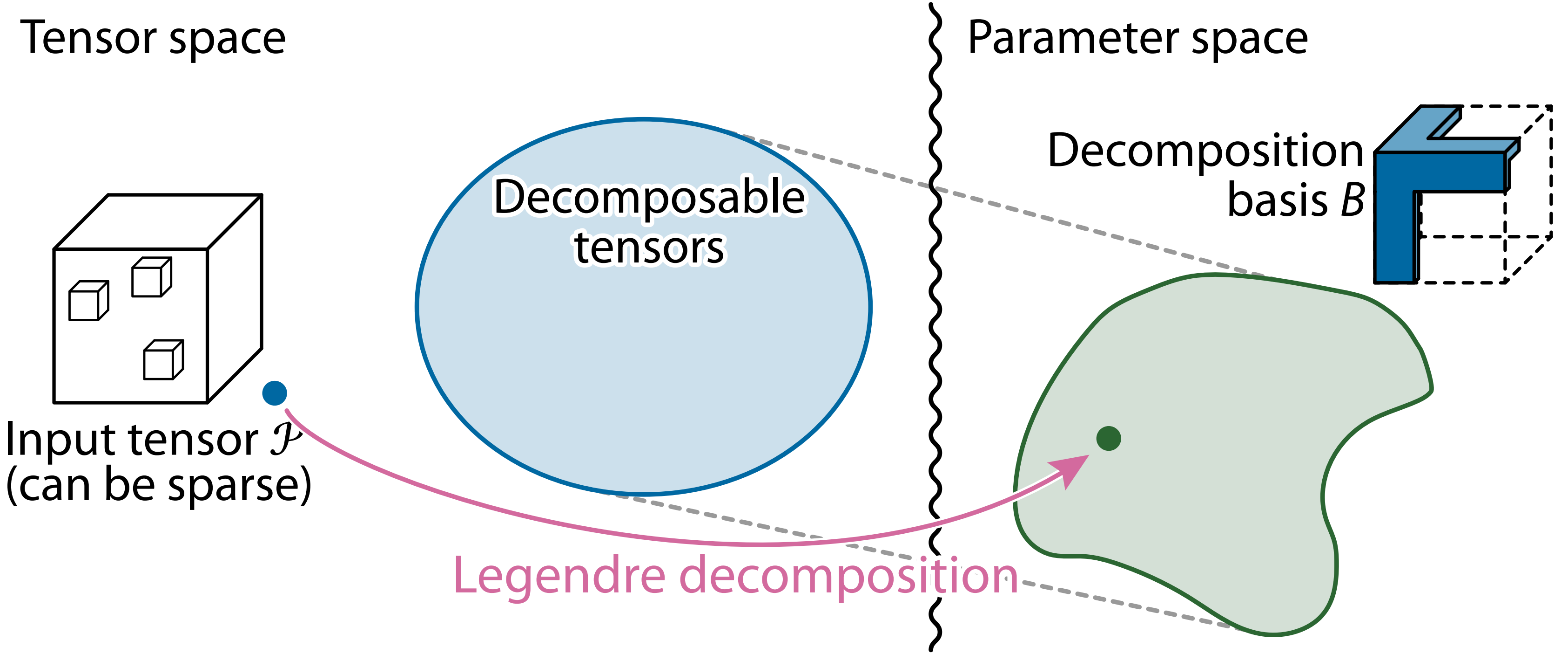
Our Approach



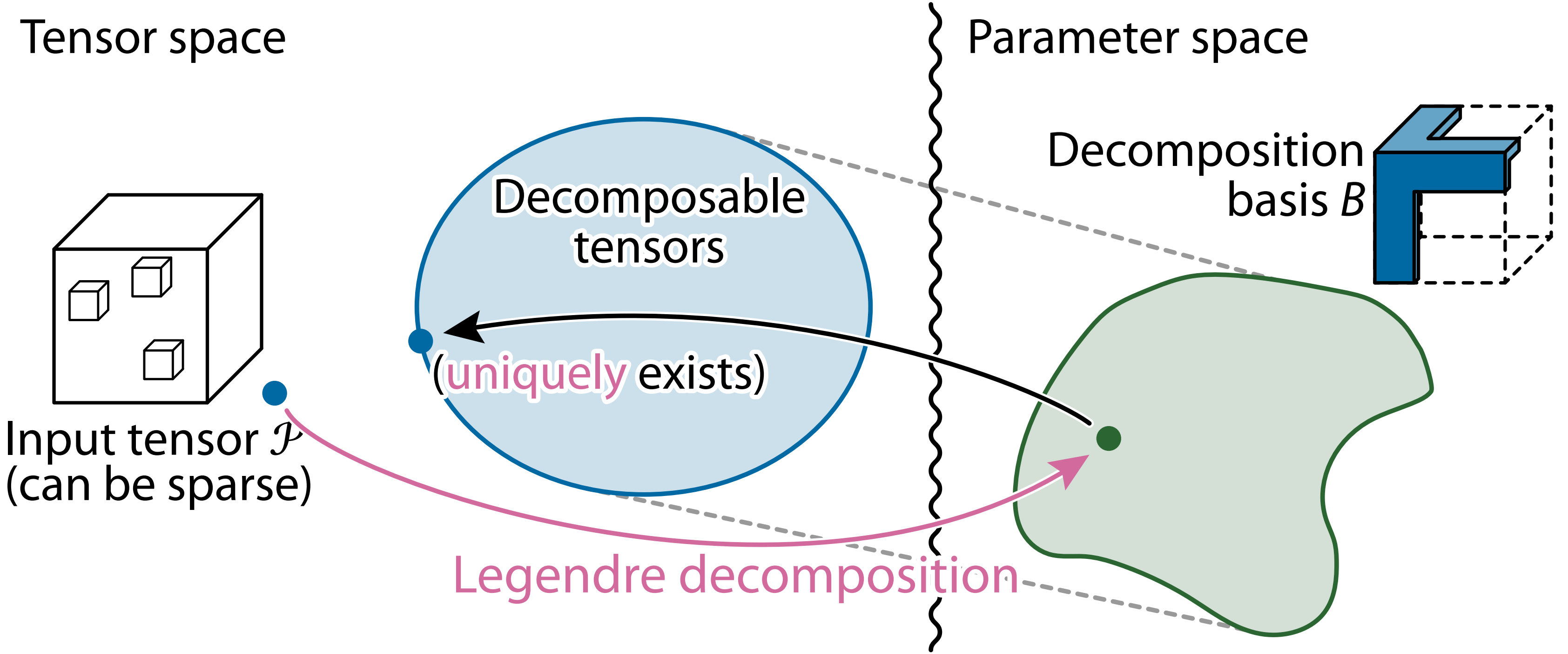
Legendre Decomposition



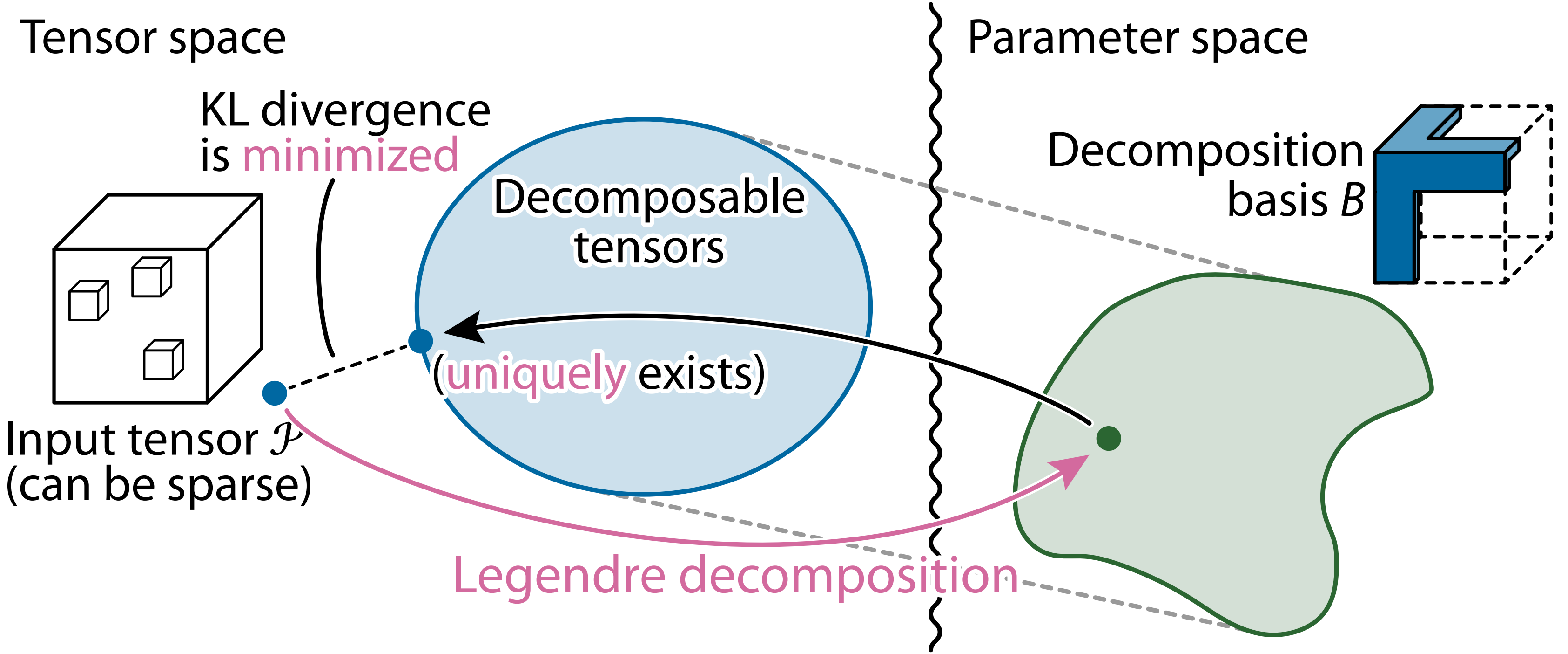
Legendre Decomposition



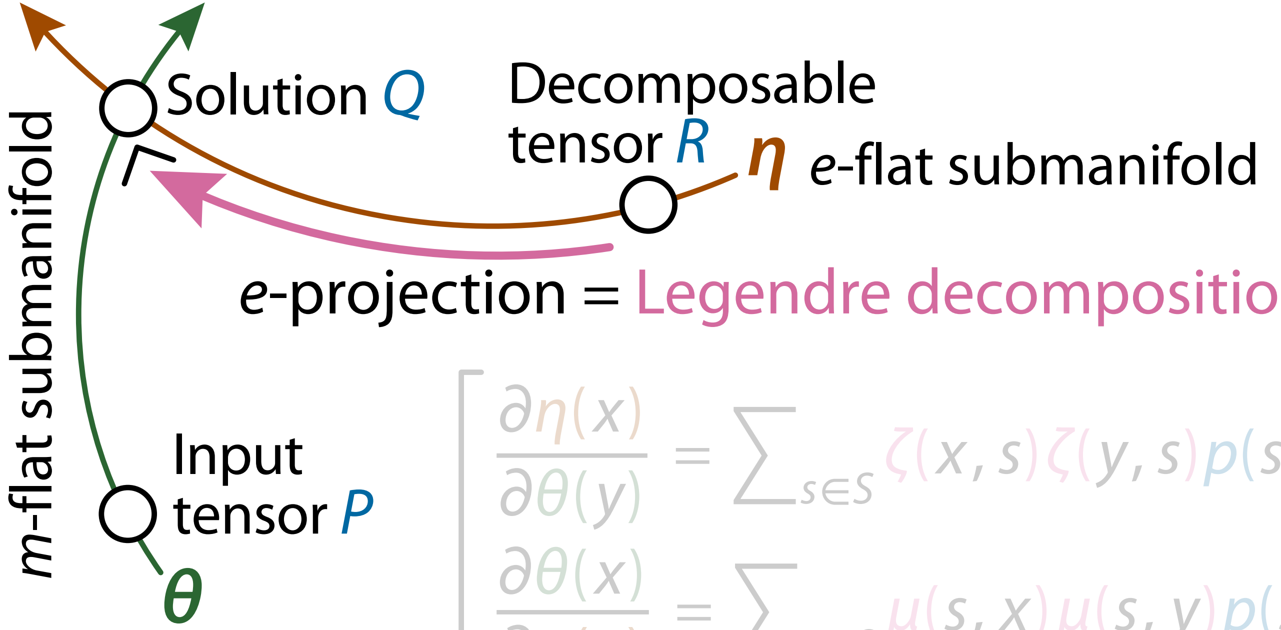
Legendre Decomposition



Legendre Decomposition



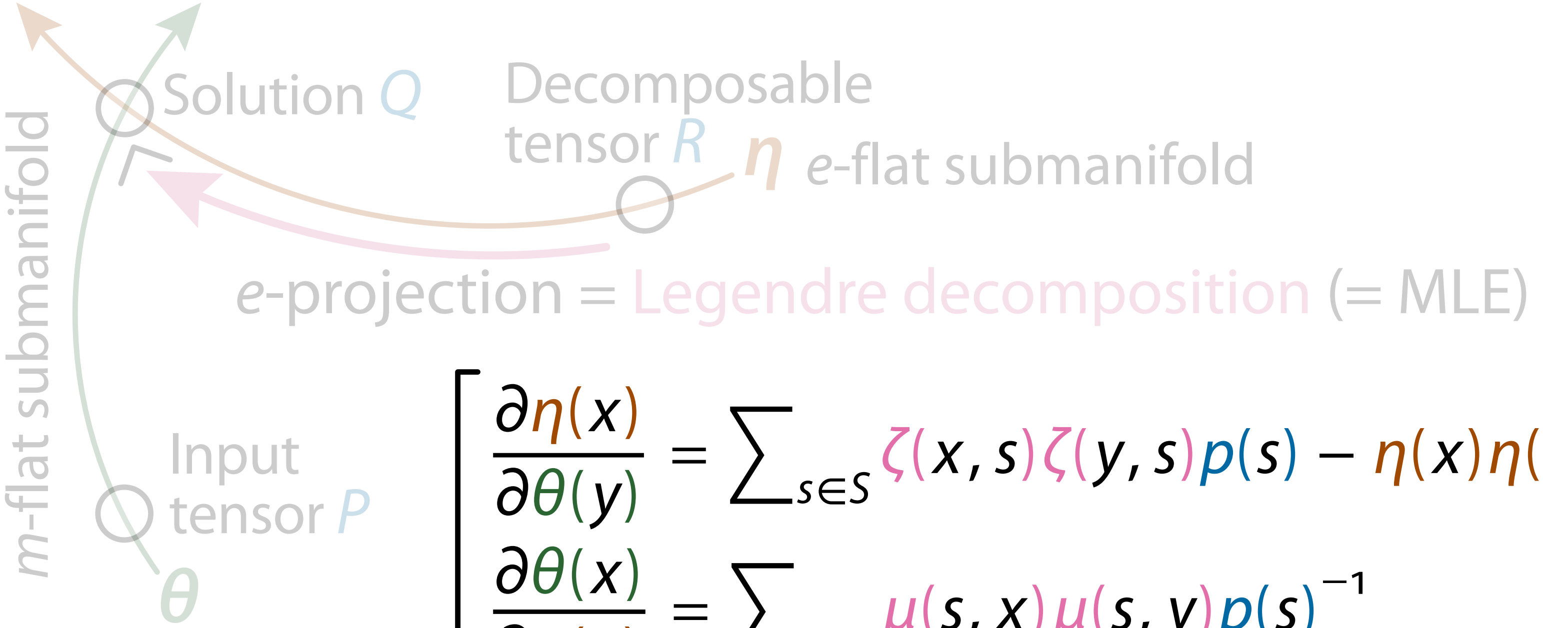
Information Geometry



e-projection = Legendre decomposition (= MLE)

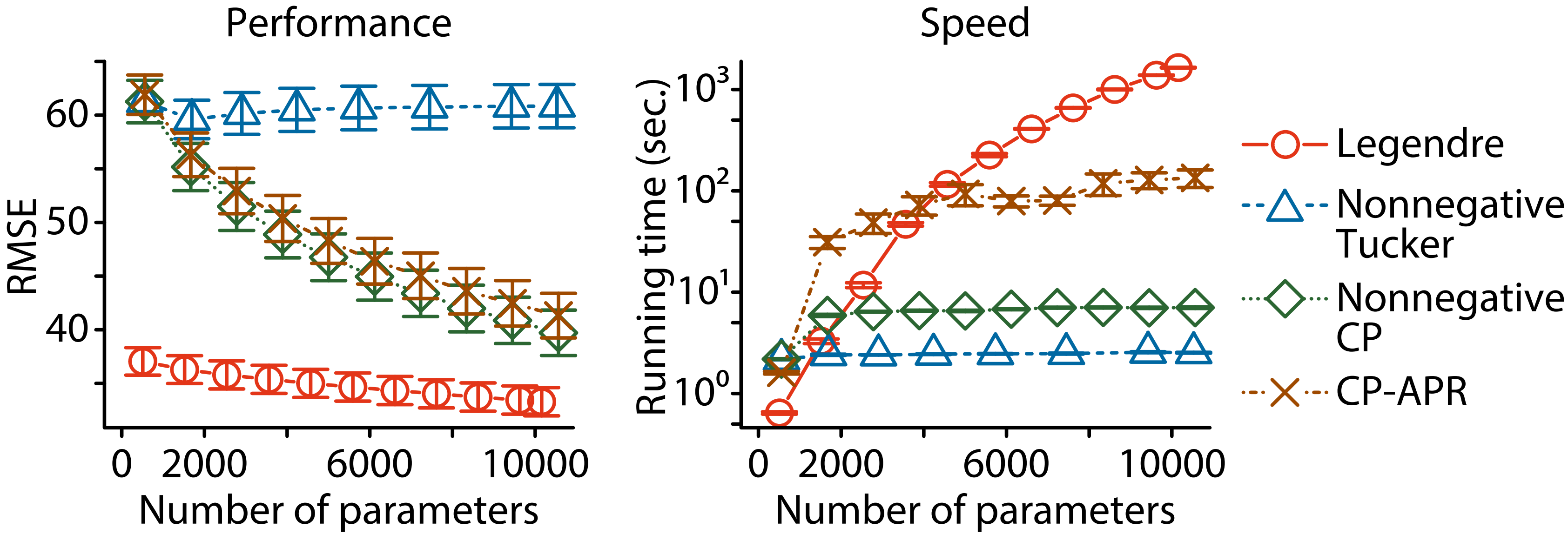
$$\left[\begin{aligned} \frac{\partial \eta(x)}{\partial \theta(y)} &= \sum_{s \in S} \zeta(x,s)\zeta(y,s)p(s) - \eta(x)\eta(y) \\ \frac{\partial \theta(x)}{\partial \eta(y)} &= \sum_{s \in S} \underbrace{\mu(s,x)\mu(s,y)}_{\text{Capture partial order structure}} p(s)^{-1} \end{aligned}\right.$$

Information Geometry



$$\left[\begin{array}{l} \dfrac{\partial \eta(x)}{\partial \theta(y)} = \sum_{s \in S} \zeta(x,s)\zeta(y,s)p(s) - \eta(x)\eta(y) \\ \dfrac{\partial \theta(x)}{\partial \eta(y)} = \sum_{s \in S} \underbrace{\mu(s,x)\mu(s,y)}_{\text{Capture partial order structure}} p(s)^{-1} \end{array}\right.$$

Experimental Results on MNIST



Summary

- We present Legendre decomposition for tensors
  - The solution always uniquely exists and minimizes the KL divergence
  - Dually flat manifold in information geometry is used
    - Parameters θ and constraints η are connected via Legendre transformation

Summary

- We present Legendre decomposition for tensors
  - The solution always uniquely exists and minimizes the KL divergence
  - Dually flat manifold in information geometry is used
    - Parameters θ and constraints η are connected via Legendre transformation