# Accelerated Dempster-Shafer using Tensor Decompositions

Cassandra Armstrong, Duc P. Truong, Erik Skau, Kari Sentz

ISIPTA 2023
July 11-14, 2023

## Dempster-Shafer Motivation

Dempster-Shafer Theory [1,2] is a mathematical framework that allows for the preservation of imprecision and uncertainty in decision making. Traditional Dempster-Shafer algorithms work on mass functions defined on subsets of a frame of discernment.
A *frame of discernment*, $\Omega$, holds a collection of propositions one can choose from.

We have an object that can be one of three colors, $\Omega = \{red, green, blue\}$.

A *Basic Probability Assignment* (BPA) assigns masses to subsets of $\Omega$ based on the degree of belief in that subset, the masses are nonnegative and sum to 1.

Two observers detect the object:
Observer 1 is red-green colorblind and assigns masses

$$m_1(\{blue\}) = 0.1 \qquad m_1(\{red, green\}) = 0.9 \tag{1}$$

Observer 2 is green-blue colorblind and assigns masses

$$m_2(\{green, blue\}) = 0.9 \qquad m_2(\{red, green, blue\}) = 0.1 \tag{2}$$

The *conjunctive join* is a bilinear rule of combination that combines mass functions into a new mass function, $m_{1,2} = m_1 \oplus m_2$, where

$$m_{1,2}(z) = \sum_{x \cap y = z} m_1(x) \cdot m_2(y) \tag{3}$$

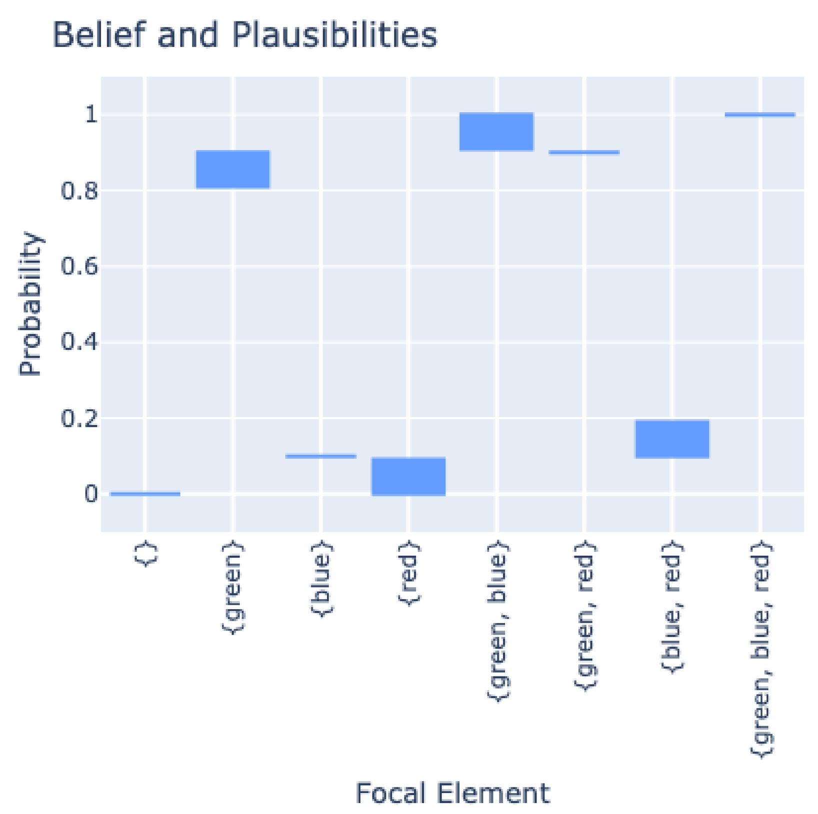
$$m_{1,2}(\{green\}) = 0.81 \qquad m_{1,2}(\{blue\}) = 0.1 \qquad m_{1,2}(\{red, green\}) = 0.09 \tag{4}$$

*Belief* and *plausibility* are linear operators that provide lower and upper bounds, respectively, for the degree of belief in any subset of the frame of discernment.

$$\text{belief}_m(y) = \sum_{x \subseteq y} m(x)$$

$$\text{plausibility}_m(y) = \sum_{x \cap y \neq \emptyset} m(x)$$

In contrast, Bayesian approaches assign probabilities to the individual components, $\{red\}$, $\{green\}$, and $\{blue\}$, and do not allow the observers to indicate the known ignorance in reporting their beliefs.




## Problem Description

Traditional algorithms suffer from the curse of dimensionality and can render high dimensional problems intractable with exponential complexities. If there are $N = |\Omega|$ elements in the frame of discernment then the naive complexities are:

- Mass function storage complexity: $\mathcal{O}(2^N)$
- Conjunctive join computational complexity: $\mathcal{O}(2^{3N})$
- Belief/Plausibility computational complexity: $\mathcal{O}(2^{2N})$.

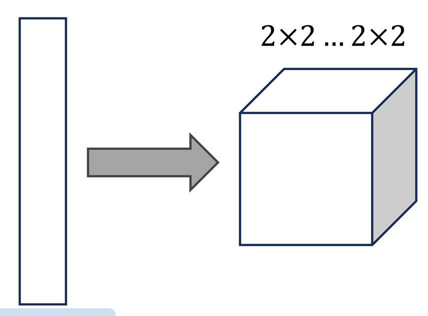
## Solution Approach

We investigate the use of tensor decompositions [3] to overcome the curse of dimensionality and accelerate the traditional computations. If the Dempster-Shafer operators are viewed as tensor operators, rather than (bi)linear, they have low rank decompositions. This observation enables two improvements in the complexities by:

- representing mass functions with tensors rather than vectors,
- approximating mass functions with tensor networks.

## Low Rank Structure of Operators

Representing mass functions with tensors rather than vectors allows us to exploit a low rank structure in the operators.




Similar to decomposing a linear operator $A$, and applying it to data $X$, to obtain a solution $B$.

$$AX = B, \qquad A = uv^\top \to u(v^\top X) = B. \tag{5}$$

By considering mass functions as multidimensional arrays rather than vectors, the complexities can be reduced:

- Conjunctive join computational complexity: $\mathcal{O}(N2^{2N})$
- Belief/Plausibility computational complexity: $\mathcal{O}(N2^N)$.

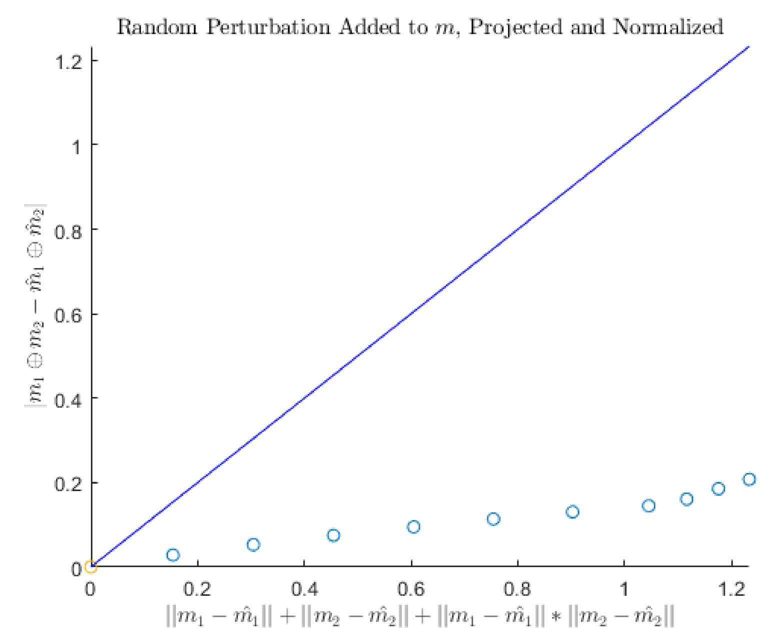
## BPA Approximation Error Analysis

Given an approximation, $\hat{m}$, of a mass function, $m$, with the properties

$$\|m - \hat{m}\|_1 \leq \epsilon, \qquad \sum \hat{m} = 1, \qquad \hat{m} \geq 0, \tag{6}$$

then we can ensure that:

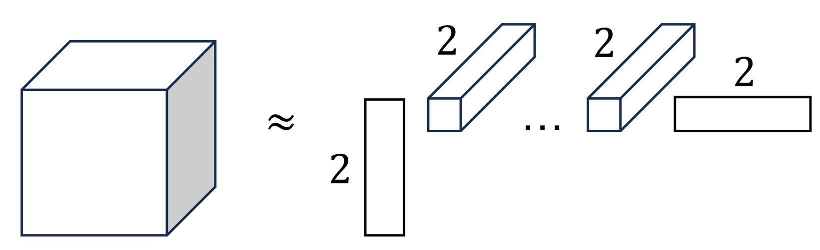
- $|\text{belief}_m(x) - \text{belief}_{\hat{m}}(x)| \leq \epsilon$
- $|\text{plausibility}_m(x) - \text{plausibility}_{\hat{m}}(x)| \leq \epsilon$
- $|(m_1 \oplus m_2)(x) - (\hat{m}_1 \oplus \hat{m}_2)(x)| \leq \epsilon_1 + \epsilon_2 + \epsilon_1 \cdot \epsilon_2$




## Low Rank Structure of Data

We consider decompositions of the form,

$$m \approx \hat{m} = G_1^{2\times k_1} G_2^{k_1 \times 2 \times k_2} \cdots \times G_{N-1}^{k_{N-2} \times 2 \times k_{N-1}} \times G_N^{k_{N-1} \times 2} \tag{7}$$




Similar to decomposing the data $X$, for the application of a linear operator $A$, to obtain a solution $B$.

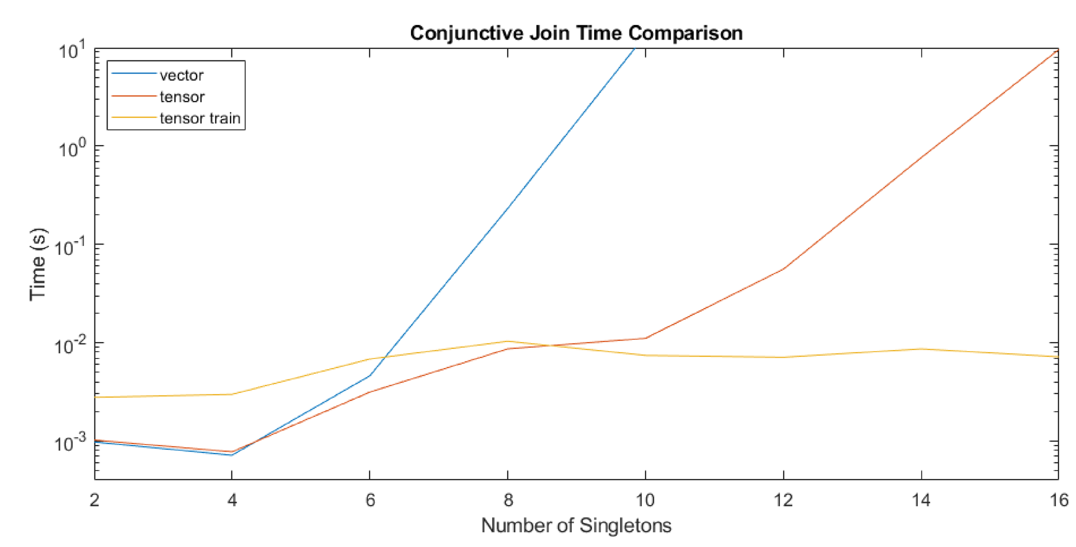
$$AX = B, \quad A = uv^\top, \qquad X \approx WH \to u(v^\top W)H \approx B \tag{8}$$

By approximating mass functions with tensor decompositions with the desired properties, complexities can be reduced:

- Mass function storage complexity: $\mathcal{O}(NK^2)$
- Conjunctive join computational complexity: $\mathcal{O}(NK_1^2K_2^2)$
- Belief/Plausibility computational complexity: $\mathcal{O}(NK^2)$.

## Preliminary Results

Timings of conjunctive join computations on synthetic data for $K_1 = K_2 = 4$ and various $N = |\Omega|$.




- Conjunctive join of vectors is the slowest.
- Conjunctive join of tensors (reshaped vectors) is faster.
- Conjunctive join of tensor trains (decomposed tensors) is fastest.

## Acknowledgements

Research presented in this poster was supported by the Laboratory Directed Research and Development program of Los Alamos National Laboratory under project number 20220799DI.

1. Dempster, Arthur P. "Upper and lower probabilities induced by a multivalued mapping." Classic works of the Dempster-Shafer theory of belief functions 219.2 (2008): 57-72.
2. Shafer, Glenn. A mathematical theory of evidence. Vol. 42. Princeton university press, 1976.
3. Kolda, Tamara G., and Brett W. Bader. "Tensor decompositions and applications." SIAM review 51.3 (2009): 455-500.



Los Alamos
NATIONAL LABORATORY