

A generalization of Kruskal's theorem

What happens when you replace the Kruskal ranks with standard ranks in Kruskal's theorem?

Benjamin Lovitz* and Fedor Petrov**

*Institute for Quantum Computing, University of Waterloo

**St. Petersburg State University; St. Petersburg Department of Steklov Mathematical Institute
of Russian Academy of Sciences

AMS Spring Central Sectional Meeting

Special Session on Nonlinear Algebra with Applications to Statistics III

March 27, 2022

[arXiv:2103.15633](https://arxiv.org/abs/2103.15633)

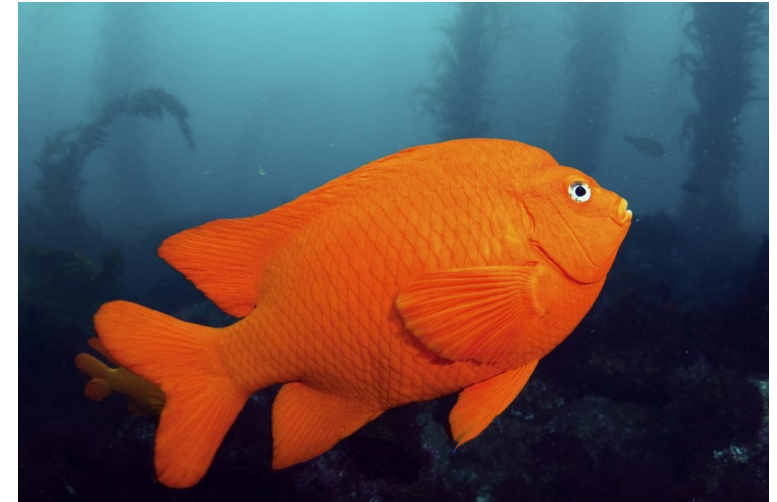


UNIVERSITY OF
WATERLOO

IQC Institute for
Quantum
Computing

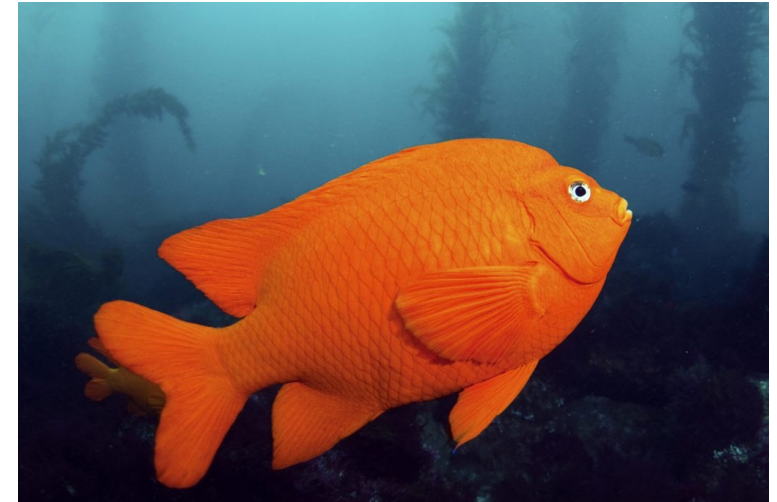
Outline

- Uniqueness of tensor decompositions
- Kruskal's theorem
- Splitting theorem
- A generalization of Kruskal's theorem



Outline

- Uniqueness of tensor decompositions
- Kruskal's theorem
- Splitting theorem
- A generalization of Kruskal's theorem



Tensor decompositions

Definition: Let $n \in \mathbb{N}$ and $[n] := \{1, \dots, n\}$.

Let \mathbb{F} be a field, and let X, Y, Z be \mathbb{F} -vector spaces of dimension at least 2.

For $T \in X \otimes Y \otimes Z$, an expression $T = \sum_{a \in [n]} x_a \otimes y_a \otimes z_a \in X \otimes Y \otimes Z$

is called a **decomposition** of T into product tensors

$\text{rank}(T) := \min\{n: \text{there exists a decomposition of } T \text{ into } n \text{ product tensors}\}$

Uniqueness of tensor decompositions

Definition: Let $n \in \mathbb{N}$ and $[n] := \{1, \dots, n\}$.

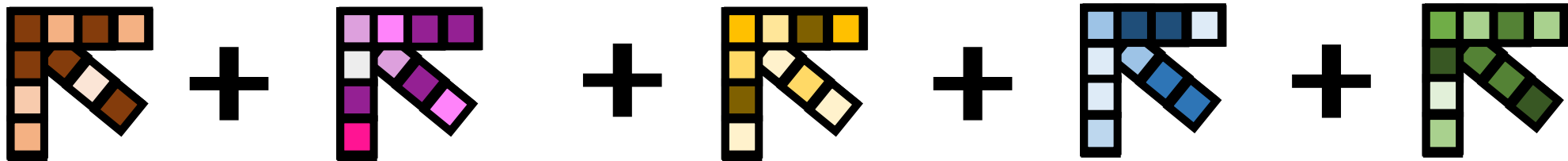
A rank decomposition

$$T = \sum_{a \in [n]} x_a \otimes y_a \otimes z_a \in X \otimes Y \otimes Z$$

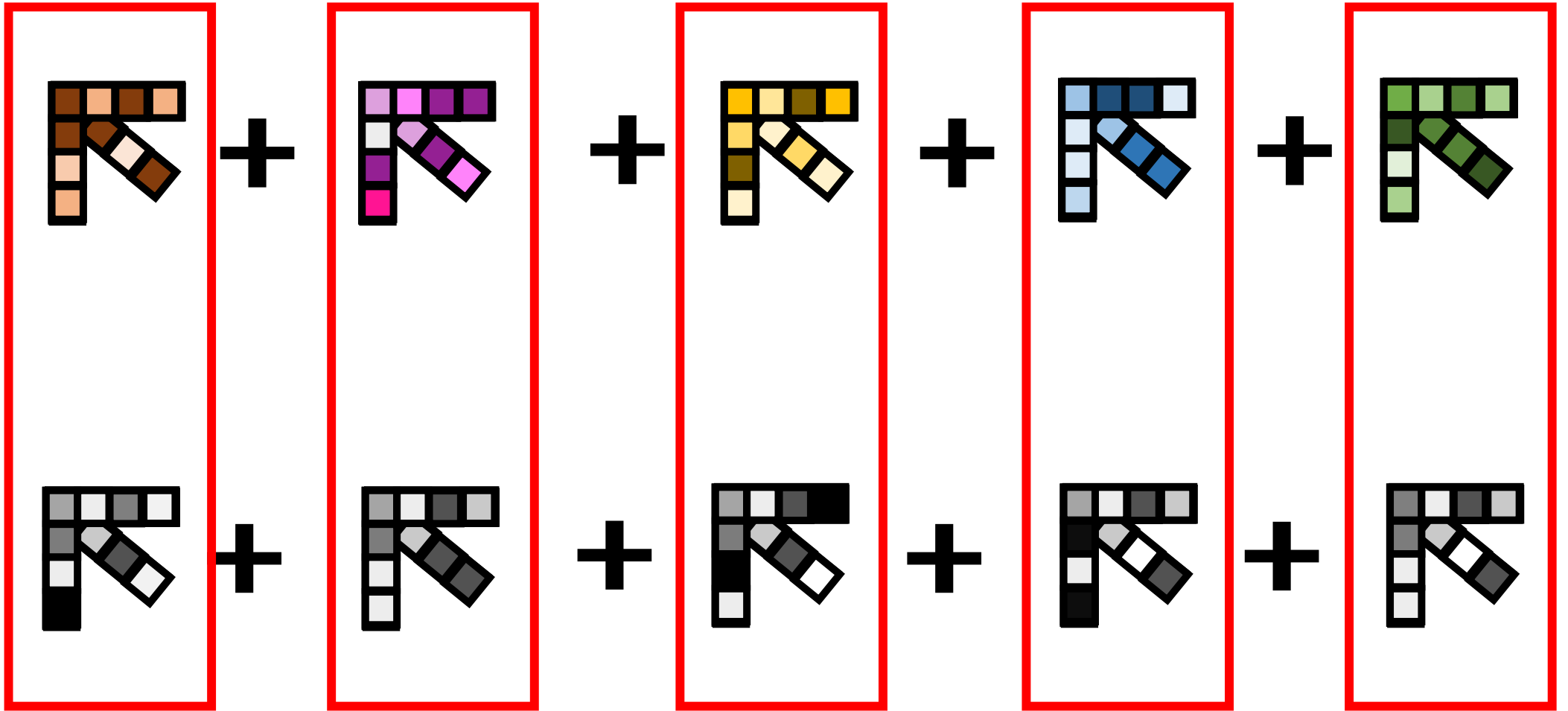
is called the **unique (rank) decomposition** of T if for any other decomposition

$$T = \sum_{a \in [n]} x'_a \otimes y'_a \otimes z'_a \in X \otimes Y \otimes Z$$

there is a permutation $\sigma \in S_n$ such that $x_a \otimes y_a \otimes z_a = x'_{\sigma(a)} \otimes y'_{\sigma(a)} \otimes z'_{\sigma(a)}$ for all $a \in [n]$.



=



Application: Latent parameter learning

 *L is for latent*

- Let A, B, C, L be finite random variables such that A, B, C are conditionally independent, i.e.

$$\Pr(a, b, c|l) = \Pr(a|l) \Pr(b|l) \Pr(c|l) \quad \text{for all } a, b, c, l.$$

- Goal: Given the probability vector $\Pr(A, B, C)$, determine $\Pr(A, B, C, L)$.
- Method:

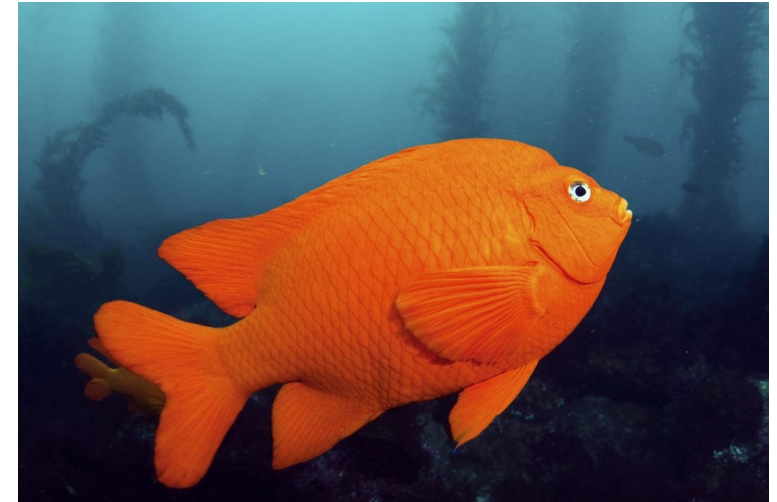
$$\Pr(A, B, C) = \sum_l \Pr(l) \Pr(A, B, C|l) = \sum_l \underbrace{\Pr(l) \Pr(A|l) \otimes \Pr(B|l) \otimes \Pr(C|l)}$$

... If $\Pr(A, B, C)$ has a unique decomposition, then we can recover $\Pr(A, B, C, l)$,

- Applications: Learning mixtures of spherical gaussians, phylogenetic tree reconstruction, hidden Markov models, orbit retrieval, blind signal separation, document topic models, ...

Outline

- Uniqueness of tensor decompositions
- **Kruskal's theorem**
- Splitting theorem
- A generalization of Kruskal's theorem



Kruskal's theorem

$$T = \sum_{a \in [n]} x_a \otimes y_a \otimes z_a \in X \otimes Y \otimes Z \quad (1)$$

General flavor of these results:

We are handed a rank decomposition (1), and we want to know if it is the unique rank decomposition of T .

Kruskal's theorem

$$T = \sum_{a \in [n]} x_a \otimes y_a \otimes z_a \in X \otimes Y \otimes Z \quad (1)$$

Definition: The **Kruskal rank** of $\{x_1, \dots, x_n\} \in X$ is the largest integer k_x such every subset $S \subseteq \{x_1, \dots, x_n\}$ of size $|S| = k_x$ is linearly independent.

$$T = e_1^{\otimes 3} + e_2^{\otimes 3} + e_3^{\otimes 3} + e_4^{\otimes 3} + (e_1 + e_2) \otimes (e_2 + e_3) \otimes (e_1 + e_4)$$

$$x_1 \otimes y_1 \otimes z_1 \quad \dots \quad x_5 \quad \otimes \quad y_5 \quad \otimes \quad z_5$$

$$\{x_1, \dots, x_5\} = \{e_1, e_2, e_3, e_4, e_1 + e_2\}, \quad k_x = 2.$$

Kruskal's theorem

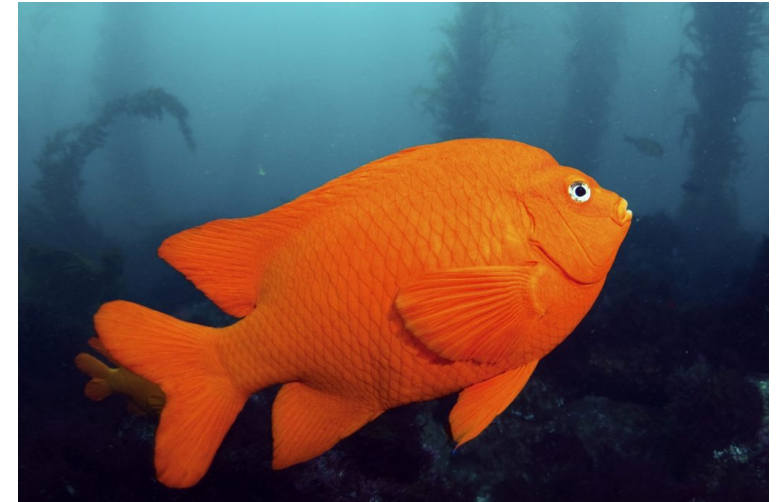
$$T = \sum_{a \in [n]} x_a \otimes y_a \otimes z_a \in X \otimes Y \otimes Z \quad (1)$$

Definition: The **Kruskal rank** of $\{x_1, \dots, x_n\} \in X$ is the largest integer k_x such every subset $S \subseteq \{x_1, \dots, x_n\}$ of size $|S| = k_x$ is linearly independent.

Kruskal's theorem: If $2n \leq k_x + k_y + k_z - 2$, then (1) is the unique rank decomposition of T.

Outline

- Uniqueness of tensor decompositions
- Kruskal's theorem
- **Splitting theorem**
- A generalization of Kruskal's theorem

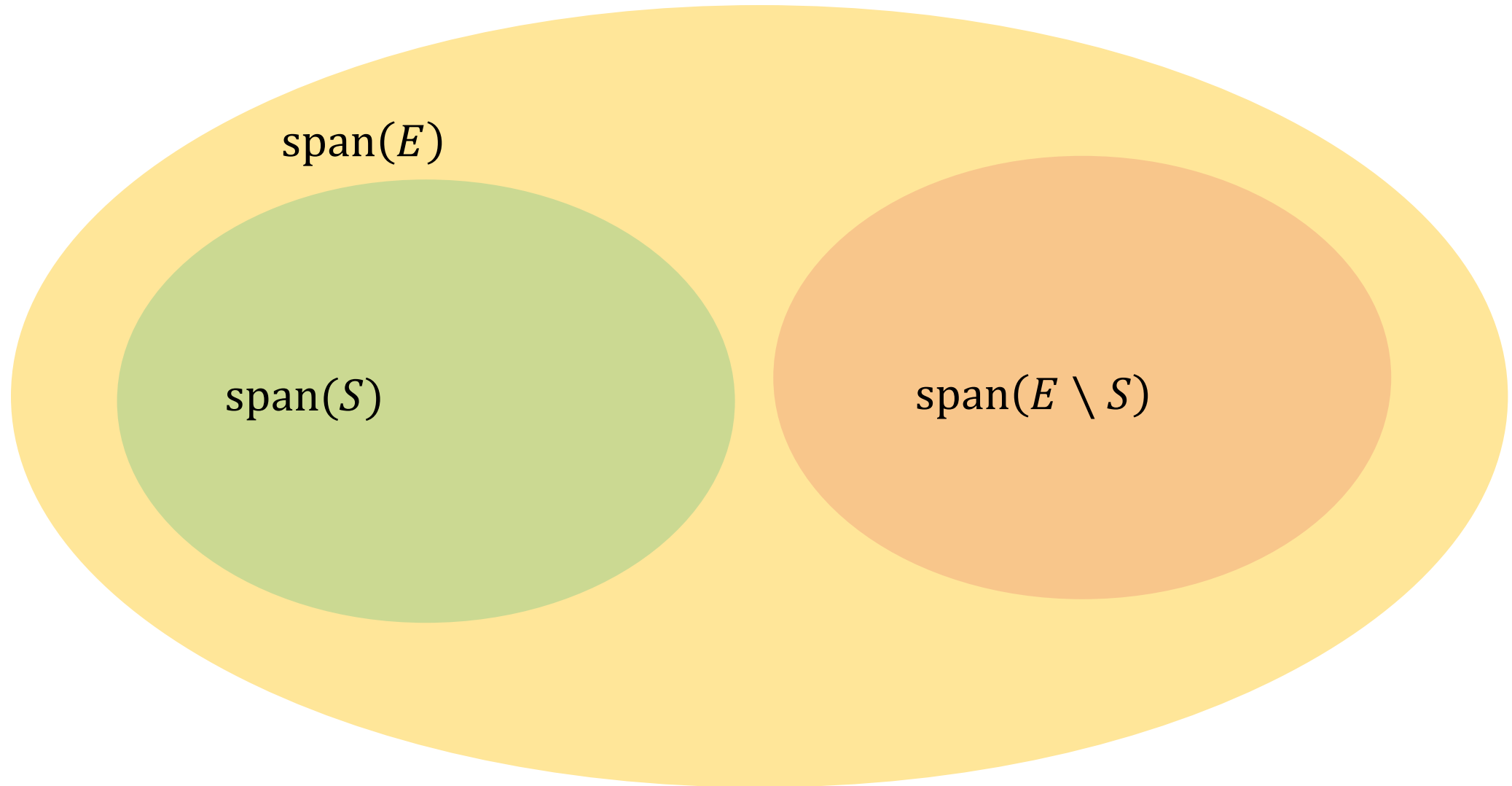


Splitting

Definition: A set of vectors $E = \{v_1, \dots, v_n\}$ **splits** if there exists a non-trivial subset $S \subseteq E$ such that

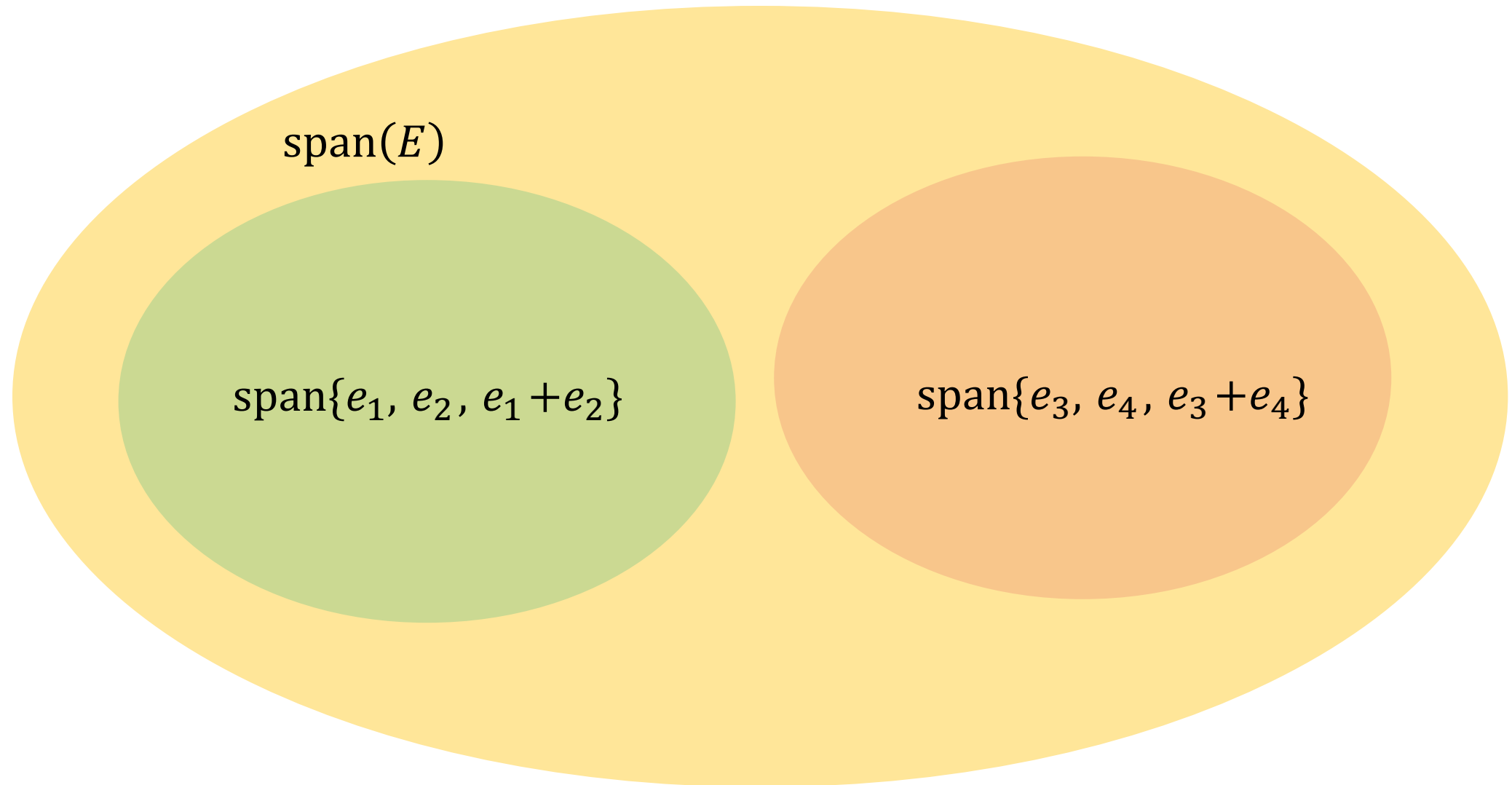
$$\text{span}(S) \cap \text{span}(E \setminus S) = \{0\} \quad (2)$$

E **splits** if there exists $S \subseteq E$ such that $\text{span}(S) \cap \text{span}(E \setminus S) = \{0\}$



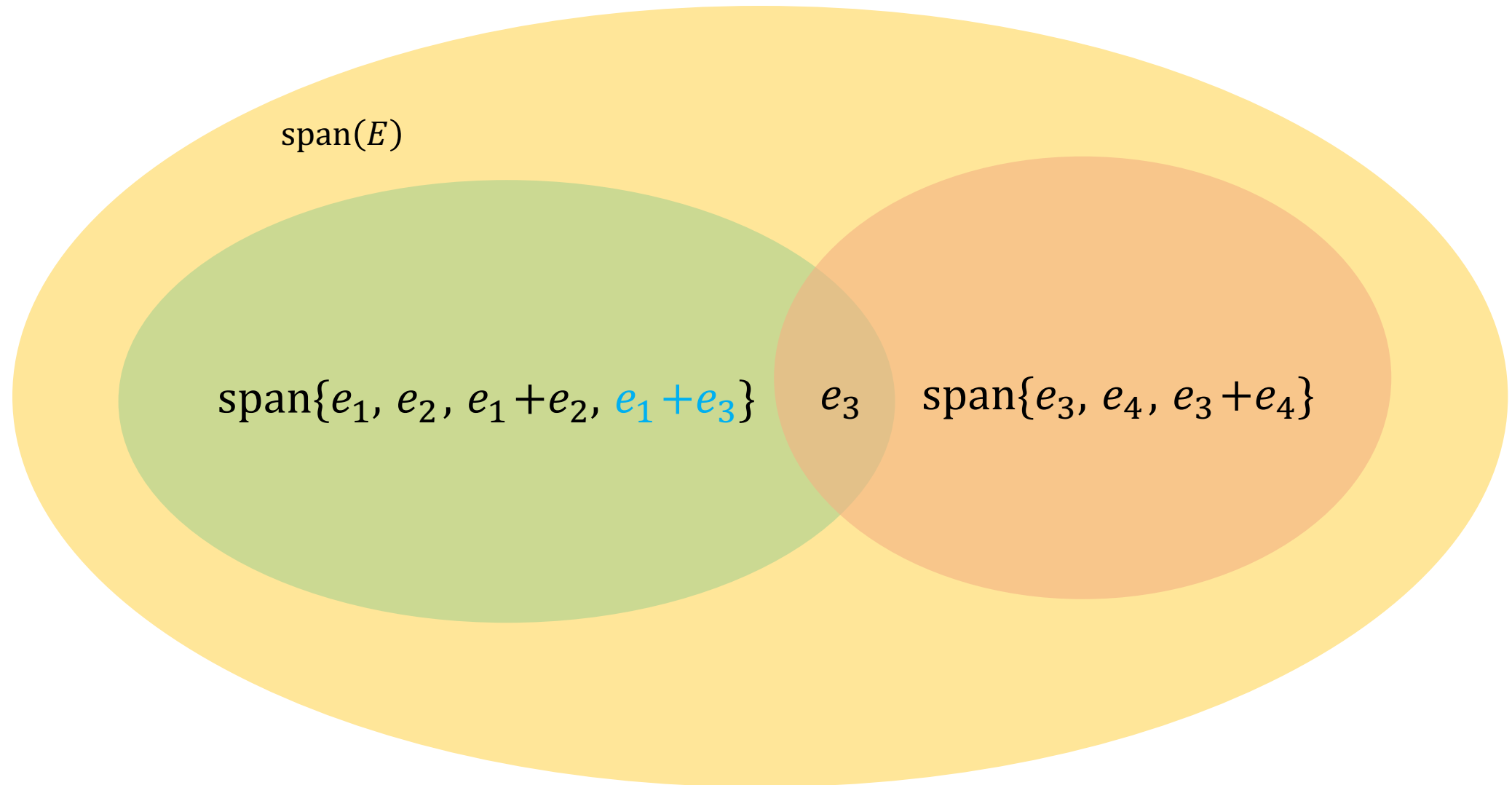
$$E = \{e_1, e_2, e_1 + e_2, e_3, e_4, e_3 + e_4\}$$

E **splits** if there exists $S \subseteq E$ such that $\text{span}(S) \cap \text{span}(E \setminus S) = \{0\}$



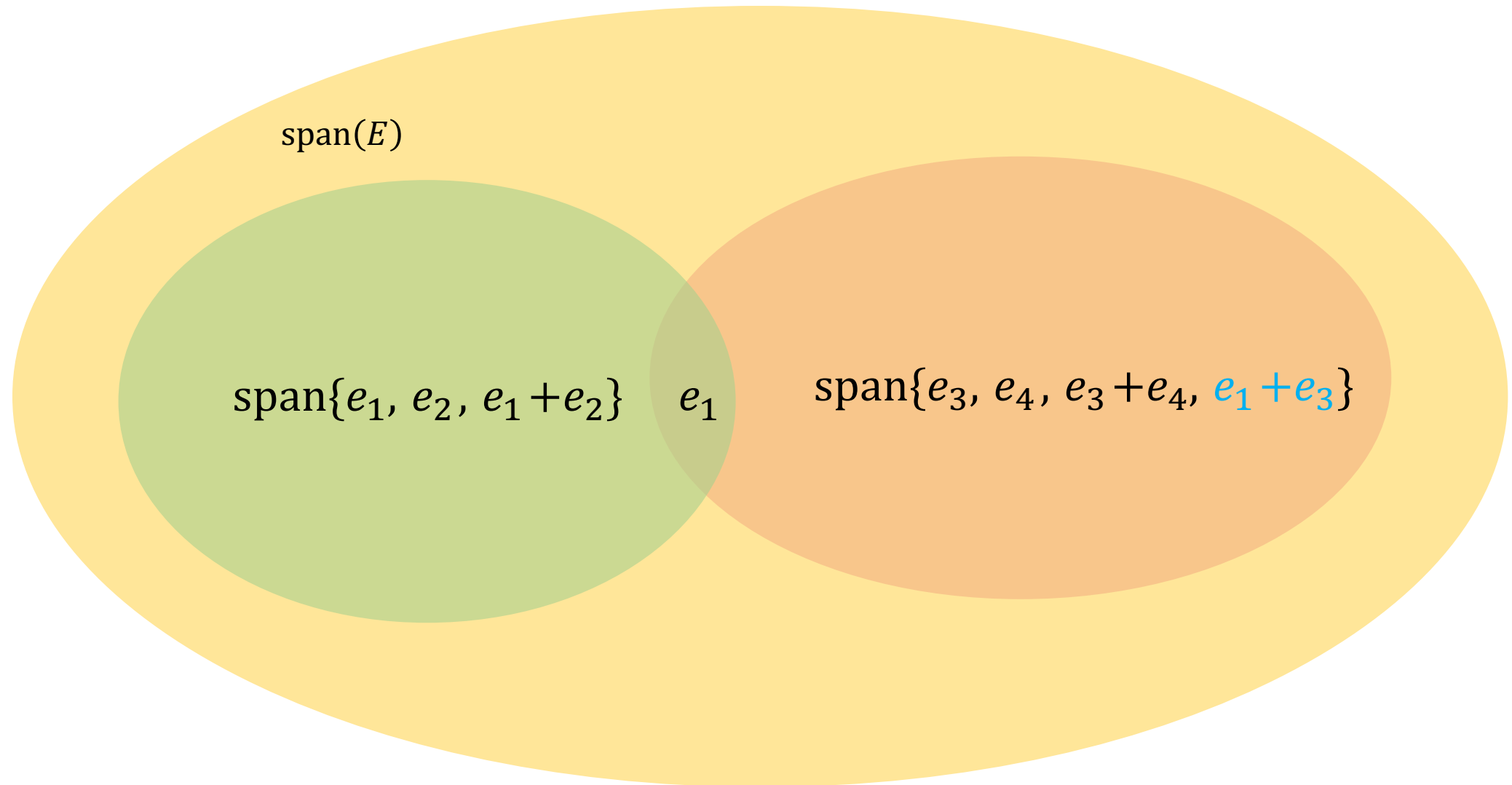
E **splits** if there exists $S \subseteq E$ such that

$$E = \{e_1, e_2, e_1 + e_2, e_3, e_4, e_3 + e_4\} \cup \{e_1 + e_3\} \quad \text{span}(S) \cap \text{span}(E \setminus S) = \{0\}$$



E **splits** if there exists $S \subseteq E$ such that

$$E = \{e_1, e_2, e_1 + e_2, e_3, e_4, e_3 + e_4\} \cup \{e_1 + e_3\} \quad \text{span}(S) \cap \text{span}(E \setminus S) = \{0\}$$



Splitting theorem

E **splits** if there exists $S \subseteq E$ such that $\text{span}(S) \cap \text{span}(E \setminus S) = \{0\}$

Splitting theorem [L-Petrov]: Let $E = \{x_a \otimes y_a : a \in [n]\}$.

If

$$\text{dimspan}(E) \leq d_x^{[n]} + d_y^{[n]} - 2$$

then E splits.


$$d_x^{[n]} = \text{dimspan}\{x_1, \dots, x_n\}$$

Splitting theorem

E **splits** if there exists $S \subseteq E$ such that $\text{span}(S) \cap \text{span}(E \setminus S) = \{0\}$

Splitting theorem [L-Petrov]: Let $E = \{x_a \otimes y_a : a \in [n]\}$.

If

$$\text{dimspan}(E) \leq d_x^{[n]} + d_y^{[n]} - 2$$

then E splits.

 $d_x^{[n]} = \text{dimspan}\{x_1, \dots, x_n\}$

(Our generalization of Kruskal's theorem is a corollary to this)

Splitting theorem

E **splits** if there exists $S \subseteq E$ such that $\text{span}(S) \cap \text{span}(E \setminus S) = \{0\}$

Splitting theorem [L-Petrov]: Let $E = \{x_a \otimes y_a \otimes z_a : a \in [n]\}$.

If

$$\text{dimspan}(E) \leq d_x^{[n]} + d_y^{[n]} + d_z^{[n]} - 3$$

then E splits.


$$d_x^{[n]} = \text{dimspan}\{x_1, \dots, x_n\}$$

Splitting theorem

E **splits** if there exists $S \subseteq E$ such that $\text{span}(S) \cap \text{span}(E \setminus S) = \{0\}$

Splitting theorem [L-Petrov]: Let $E = \{x_a \otimes y_a \otimes z_a : a \in [n]\}$.

If

$$\text{dimspan}(E) \leq d_x^{[n]} + d_y^{[n]} + d_z^{[n]} - 3$$

then E splits.


$$d_x^{[n]} = \text{dimspan}\{x_1, \dots, x_n\}$$

Corollary: If

$$n \leq d_x^{[n]} + d_y^{[n]} + d_z^{[n]} - 2,$$

then E splits.

Splitting theorem

E **splits** if there exists $S \subseteq E$ such that $\text{span}(S) \cap \text{span}(E \setminus S) = \{0\}$

Splitting theorem [L-Petrov]: Let $E = \{x_a \otimes y_a \otimes z_a : a \in [n]\}$.

If

$$\dim \text{span}(E) \leq d_x^{[n]} + d_y^{[n]} + d_z^{[n]} - 3$$

then E splits.

$$d_x^{[n]} = \dim \text{span}\{x_1, \dots, x_n\}$$

Corollary: If

$$n \leq d_x^{[n]} + d_y^{[n]} + d_z^{[n]} - 2,$$

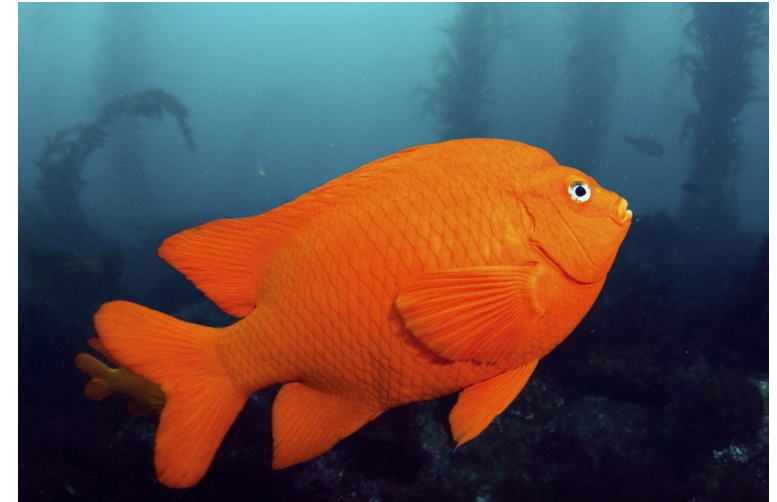
then E splits.

Replaces Kruskal ranks with standard ranks

Corollary: If $2n \leq d_x^{[n]} + d_y^{[n]} + d_z^{[n]} - 2$, then for any other set of product tensors $E' = \{x'_a \otimes y'_a \otimes z'_a : a \in [n]\}$, $E \cup E'$ splits.

Outline

- Uniqueness of tensor decompositions
- Kruskal's theorem
- Splitting theorem
- A generalization of Kruskal's theorem



Recall the tensor decomposition setup...

We are handed a rank decomposition

$$T = \sum_{a \in [n]} x_a \otimes y_a \otimes z_a$$

... and wish to control other rank decompositions

$$T = \sum_{a \in [n]} x'_a \otimes y'_a \otimes z'_a.$$

Recall the tensor decomposition setup...

We are handed a rank decomposition

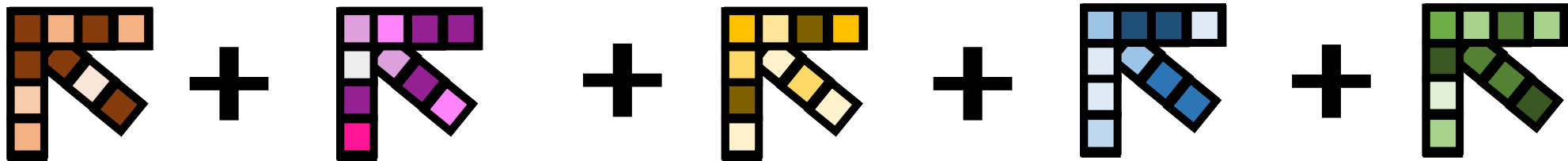
$$T = \sum_{a \in [n]} x_a \otimes y_a \otimes z_a$$

... and wish to control other rank decompositions

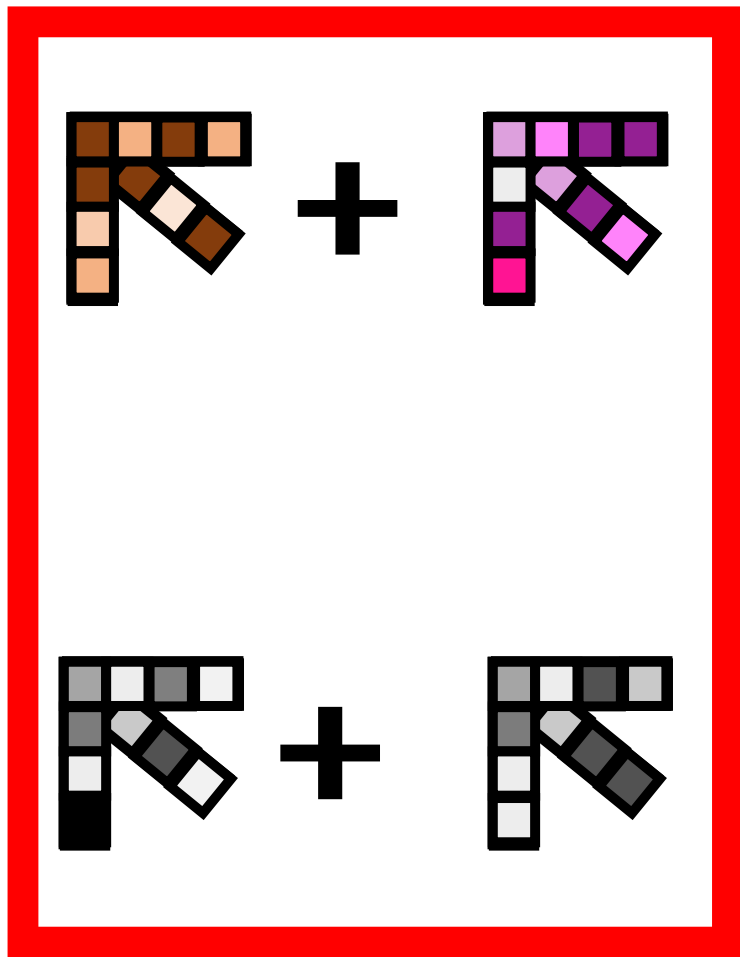
$$T = \sum_{a \in [n]} x'_a \otimes y'_a \otimes z'_a.$$

Replaces Kruskal ranks with standard ranks

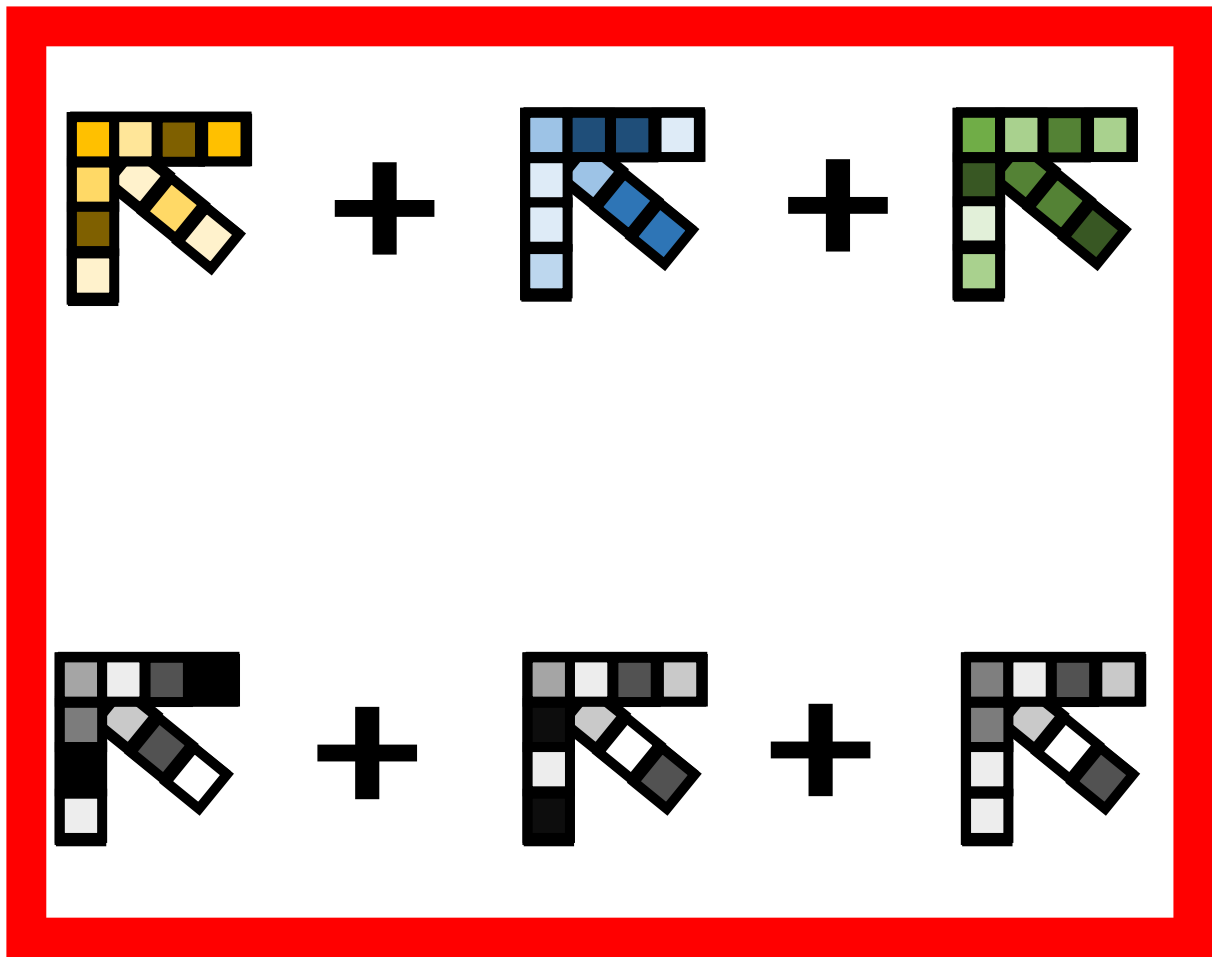
Corollary: If $2n \leq d_x^{[n]} + d_y^{[n]} + d_z^{[n]} - 2$, then for any other set of product tensors $E' = \{x'_a \otimes y'_a \otimes z'_a : a \in [n]\}$, $E \cup E'$ splits.



=



+



Our generalization

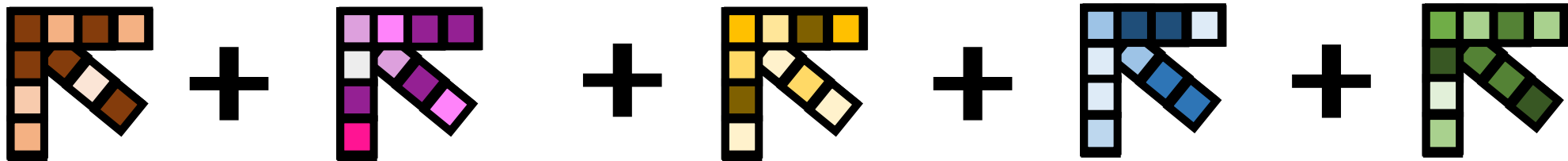
$$T = \sum_{a \in [n]} x_a \otimes y_a \otimes z_a \in X \otimes Y \otimes Z \quad (1)$$

Theorem [L-Petrov]: If for every subset $S \subseteq [n]$ of size $|S| \geq 2$, it holds that

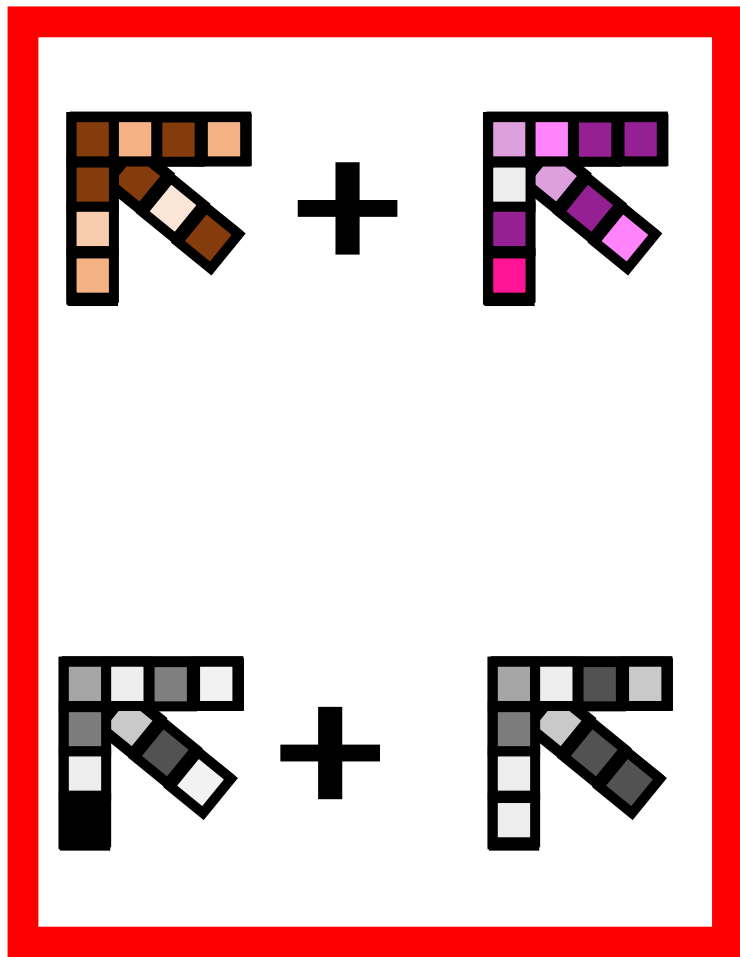
$$2|S| \leq d_x^S + d_y^S + d_z^S - 2,$$

 $d_x^S = \text{dimspan}\{x_a : a \in S\}$

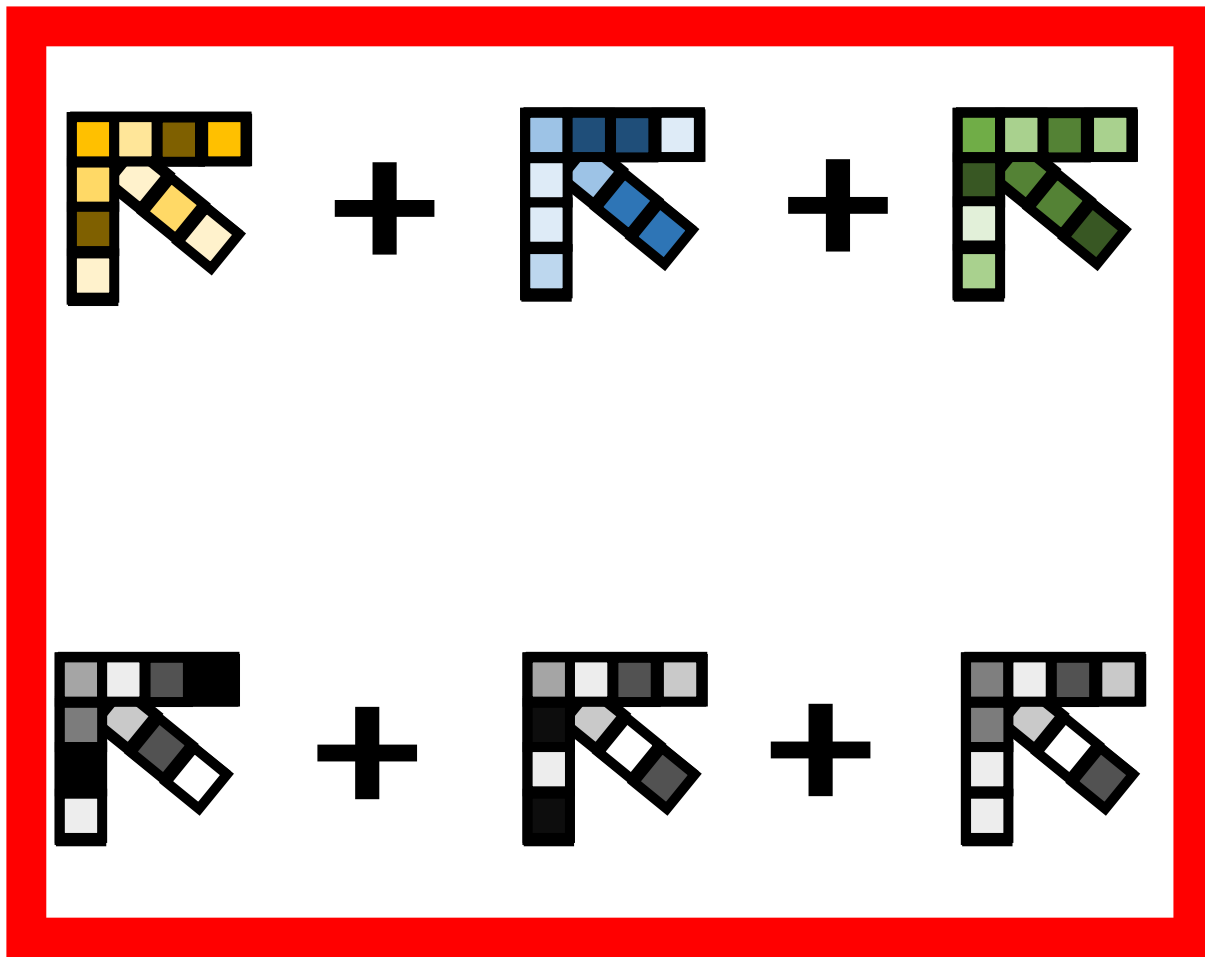
then (1) is the unique rank decomposition of T.



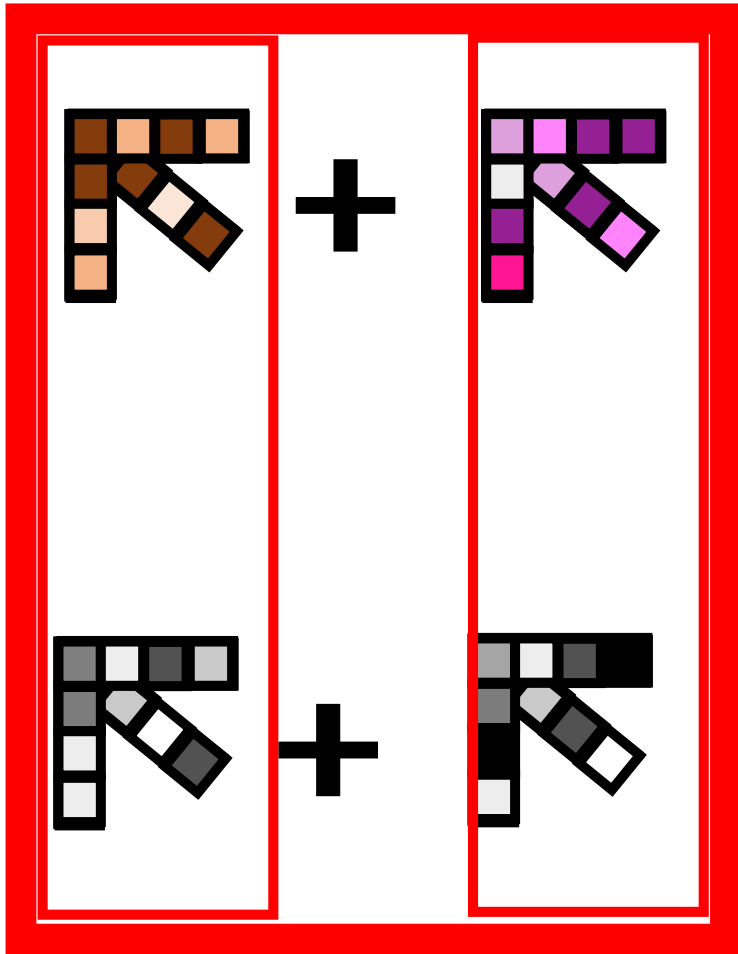
=



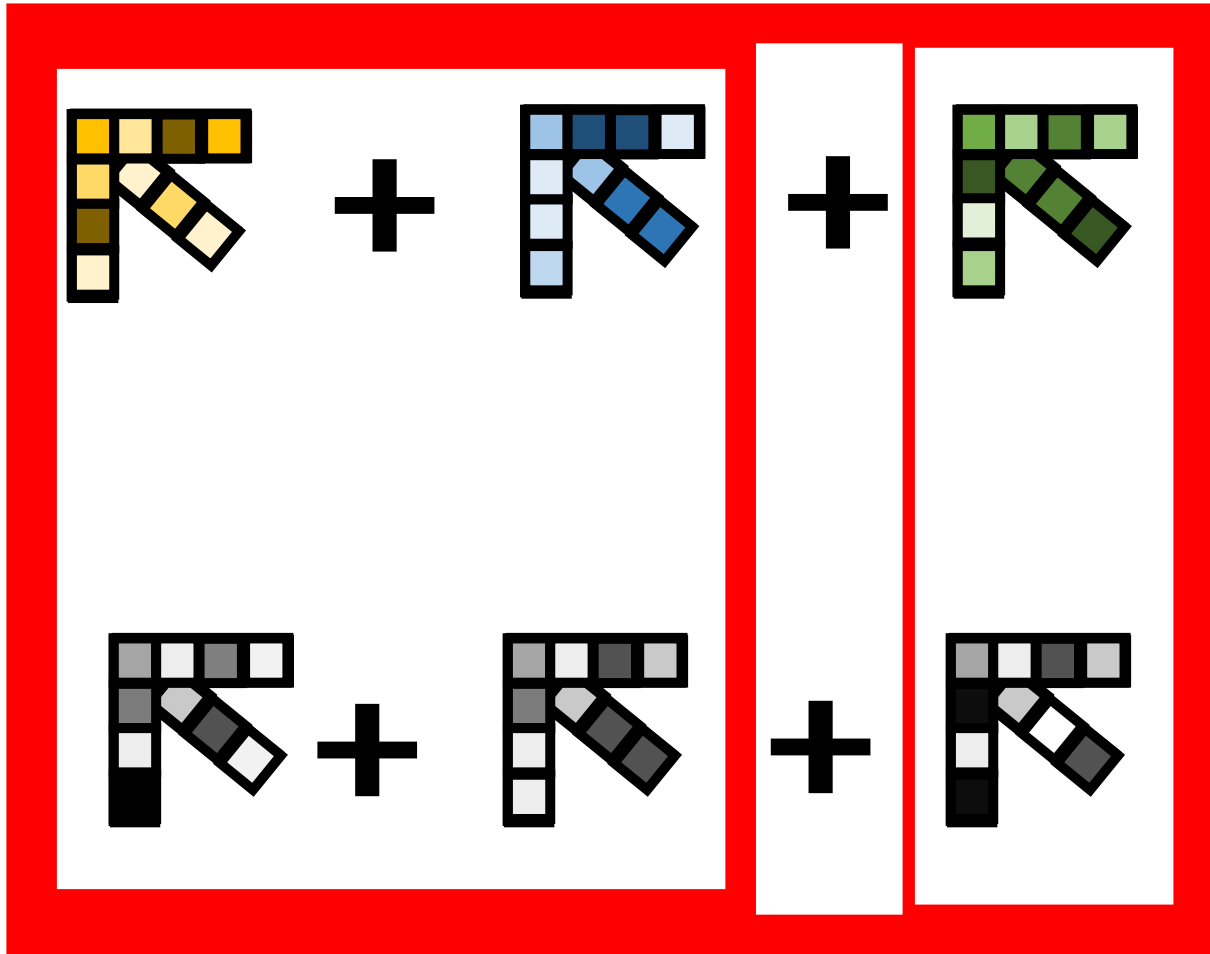
+



=



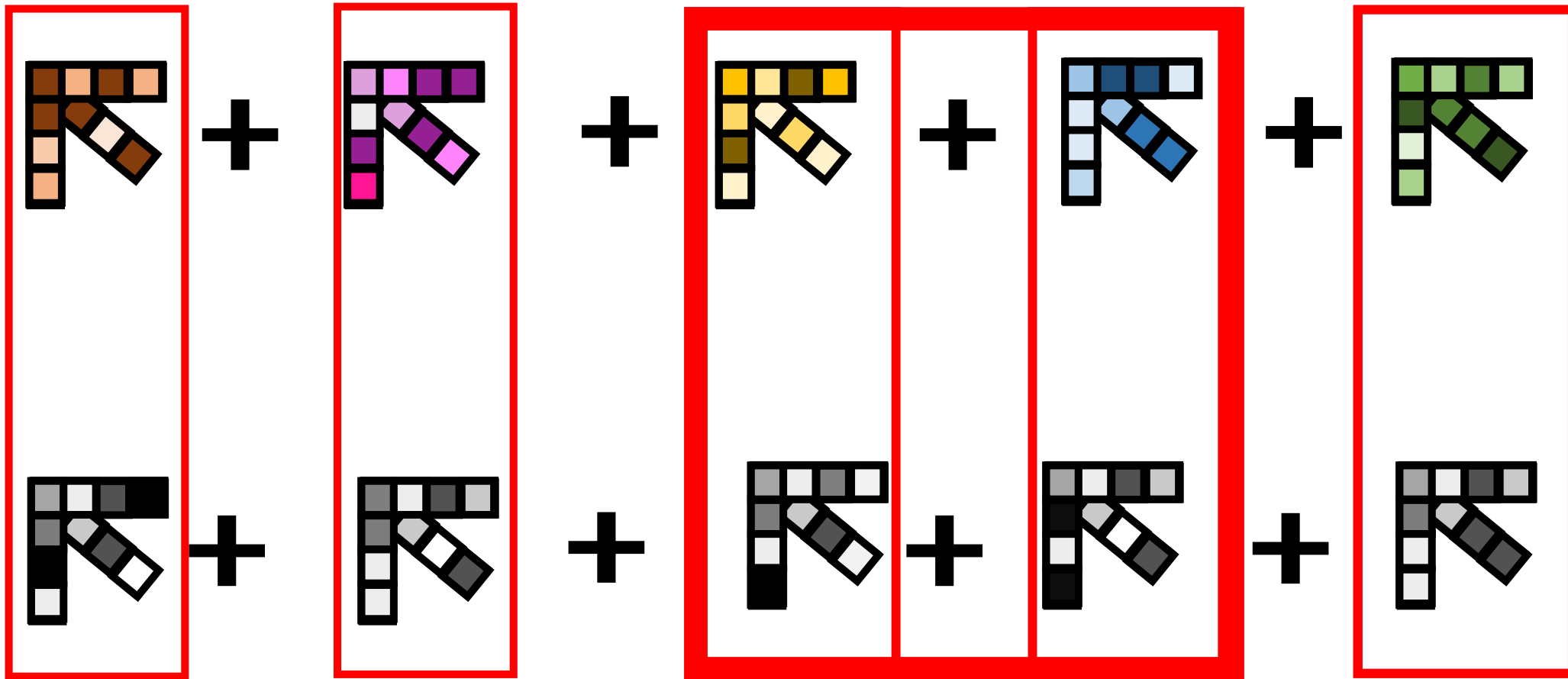
+



+

+

=



Our generalization

$$T = \sum_{a \in [n]} x_a \otimes y_a \otimes z_a \in X \otimes Y \otimes Z \quad (1)$$

Theorem [L-Petrov]: If for every subset $S \subseteq [n]$ of size $|S| \geq 2$, it holds that

$$2|S| \leq d_x^S + d_y^S + d_z^S - 2,$$

 $d_x^S = \text{dimspan}\{x_a : a \in S\}$

then (1) is the unique rank decomposition of T.

Conclusion

- Uniqueness of tensor decompositions
- Kruskal's theorem
- Splitting theorem
- A generalization of Kruskal's theorem
- More matroid theory for product tensors?



Example

$$T = \sum_{a \in [n]} x_a \otimes y_a \otimes z_a \in X \otimes Y \otimes Z \quad (1)$$

$$T = e_1^{\otimes 3} + e_2^{\otimes 3} + e_3^{\otimes 3} + e_4^{\otimes 3} + (e_1 + e_2) \otimes (e_2 + e_3) \otimes (e_1 + e_4)$$

$x_1 \otimes y_1 \otimes z_1$... $x_5 \otimes y_5 \otimes z_5$

Example

$$T = \sum_{a \in [n]} x_a \otimes y_a \otimes z_a \in X \otimes Y \otimes Z \quad (1)$$

$$T = e_1^{\otimes 3} + e_2^{\otimes 3} + e_3^{\otimes 3} + e_4^{\otimes 3} + (e_1 + e_2) \otimes (e_2 + e_3) \otimes (e_1 + e_4)$$

$x_1 \otimes y_1 \otimes z_1 \quad \dots \quad x_5 \otimes y_5 \otimes z_5$

Kruskal's theorem does not certify uniqueness

$$10 = 2n \not\leq k_x + k_y + k_z - 2 = 2 + 2 + 2 - 2 = 4$$

Example

$$T = \sum_{a \in [n]} x_a \otimes y_a \otimes z_a \in X \otimes Y \otimes Z \quad (1)$$

$$T = e_1^{\otimes 3} + e_2^{\otimes 3} + e_3^{\otimes 3} + e_4^{\otimes 3} + (e_1 + e_2) \otimes (e_2 + e_3) \otimes (e_1 + e_4)$$

$x_1 \otimes y_1 \otimes z_1$... $x_5 \otimes y_5 \otimes z_5$

Theorem [L-Petrov]: If for every subset $S \subseteq [n]$ of size $|S| \geq 2$, it holds that

$2|S| \leq d_x^S + d_y^S + d_z^S - 2$, then (1) is the unique decomposition of T.

Example

$$T = \sum_{a \in [n]} x_a \otimes y_a \otimes z_a \in X \otimes Y \otimes Z \quad (1)$$

$$T = e_1^{\otimes 3} + e_2^{\otimes 3} + e_3^{\otimes 3} + e_4^{\otimes 3} + (e_1 + e_2) \otimes (e_2 + e_3) \otimes (e_1 + e_4)$$

$x_1 \otimes y_1 \otimes z_1 \quad \dots \quad x_5 \otimes y_5 \otimes z_5$

Theorem [L-Petrov]: If for every subset $S \subseteq [n]$ of size $|S| \geq 2$, it holds that

$2|S| \leq d_x^S + d_y^S + d_z^S - 2$, then (1) is the unique decomposition of T.

For $S = \{1,2\}$, $4 = 2|S| \leq d_x^S + d_y^S + d_z^S - 2 = 2 + 2 + 2 - 2 = 4$ ✓

For $S = \{1,2,5\}$, $6 = 2|S| \leq d_x^S + d_y^S + d_z^S - 2 = 2 + 3 + 3 - 2 = 6$ ✓

For $S = \{1,2,3,5\}$, $8 = 2|S| \leq d_x^S + d_y^S + d_z^S - 2 = 3 + 3 + 4 - 2 = 8$ ✓

For $S = [5]$, $10 = 2|[n]| \leq d_x^{[n]} + d_y^{[n]} + d_z^{[n]} - 2 = 4 + 4 + 4 - 2 = 10$ ✓

Conclusion

- Uniqueness of tensor decompositions
- Kruskal's theorem
- Splitting theorem
- A generalization of Kruskal's theorem
- More matroid theory for product tensors?

