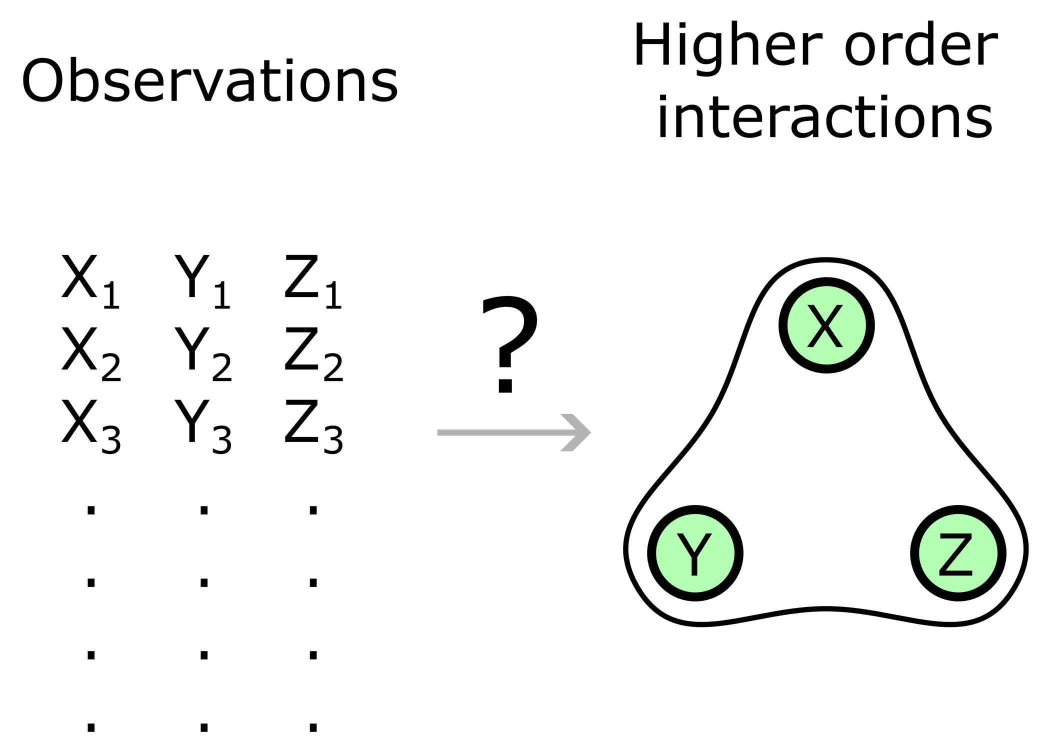
Information-Theoretic Measures on Lattices for Higher-Order Interactions

Imperial College London
Uniklinikum Würzburg

Zhaolu Liu[1], Mauricio Barahona[1] & Robert L. Peach[2]
[1]Department of Mathematics, Imperial College London
[2]Department of Neurology, University Hospital Würzburg

Motivation

1. Pairwise interactions can be insufficient
2. Relational data is not widely available
3. High-order $\neq \sum$ pairwise



Can we directly find high-order interactions from observational (non-relational) data?

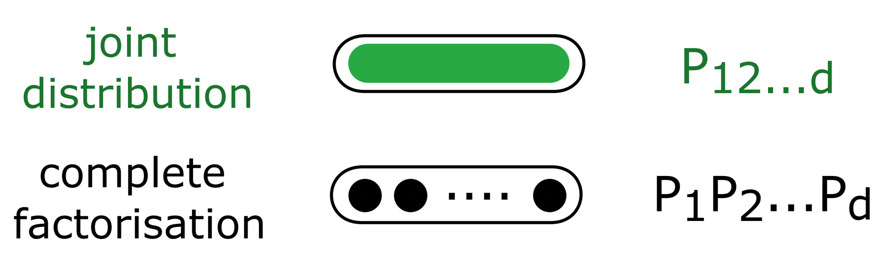
Information Measures

Dependence (2 variables) = interactions.

$$\mathrm{MI}(X_1; X_2) = \mathrm{H}(X_1) + \mathrm{H}(X_2) - \mathrm{H}(X_1, X_2)$$

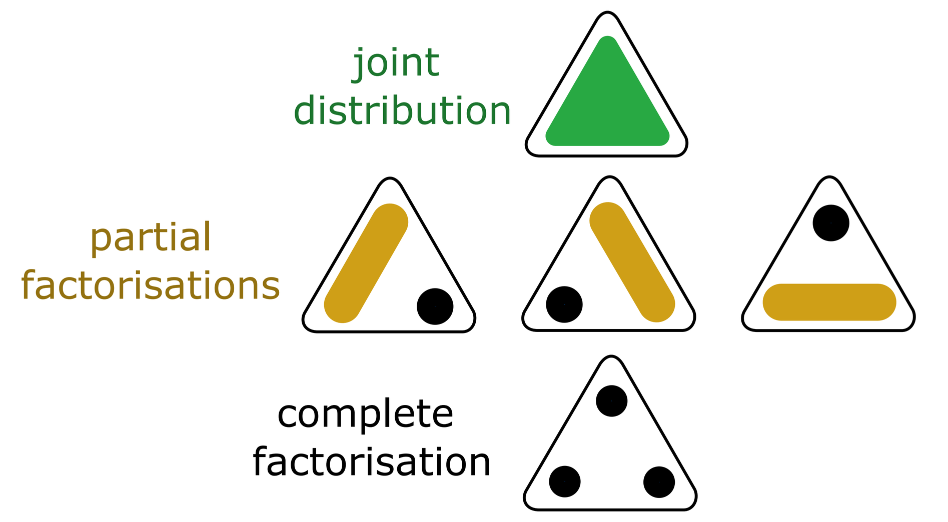
Dependence measure generalised to $d \geq 3$:

$$\mathrm{TC}(d) = \sum_{i=1}^{d} \mathrm{H}(X_i) - \mathrm{H}(X_1, \ldots, X_d)$$



Non-zero Total Correlation is less informative as d becomes large. Eg, 5 possibilities ($d = 3$):

P_{123} $P_1 P_{23}$ $P_2 P_{13}$ $P_3 P_{12}$ $P_1 P_2 P_3$



Interaction information for 3 variables:

$$\mathrm{II}(3) = \mathrm{H}(X) + \mathrm{H}(Y) + \mathrm{H}(Z) + \mathrm{H}(X,Y,Z) - \mathrm{H}(X,Y) - \mathrm{H}(X,Z) - \mathrm{H}(Y,Z)$$

Lattice

Definition 1 *Lattice ≈ set + order*

E.g., Partition lattice = partitions + refinement and its **Möbius inversion** when $d = 3$:

$$P_{123} = \Delta_{123} + \Delta_{12|3} + \Delta_{1|23} + \Delta_{2|13} + \Delta_{1|2|3}$$
$$\Delta_{123} = P_{123} - P_1 P_{23} - P_2 P_{13} - P_3 P_{12} + 2 P_1 P_2 P_3$$

Lattice formalisation is beneficial:

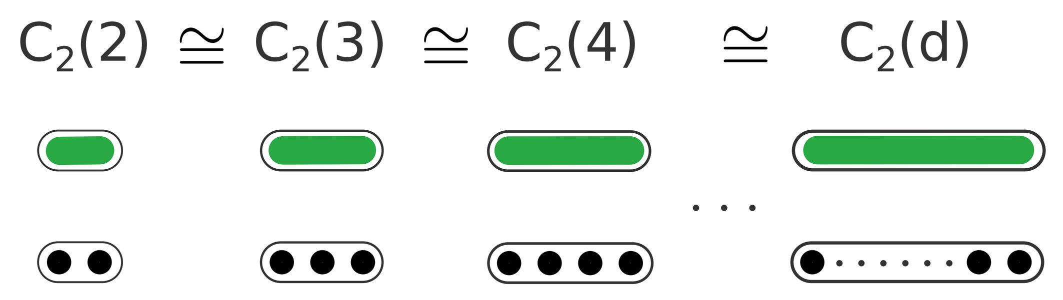
1. compare interaction measures
2. generate *new* interaction measures

Lattice Embeddings

Which lattices are associated with the existing information-theoretic interaction measures?

Two-Element Chain - Simple

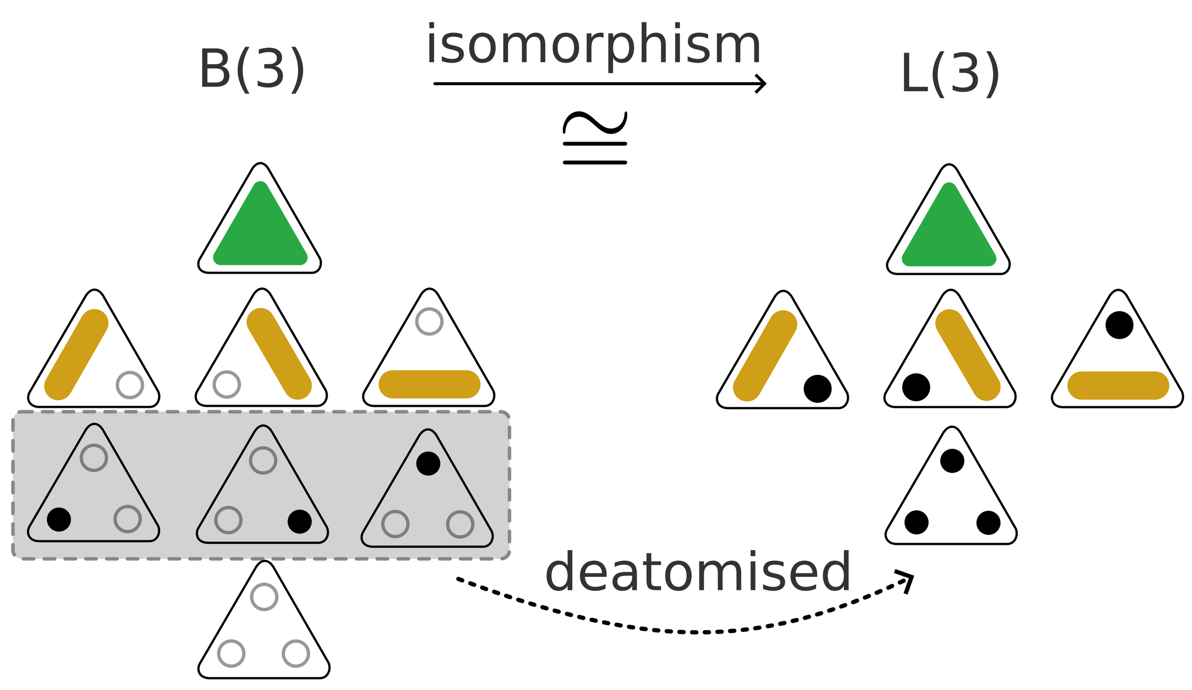
MI and TC are generated by the Two-Element Chain C_2, which are isomorphic for all d.



Boolean Lattice - Natural

Interaction information can be generated from the Boolean lattice $B(d)$.

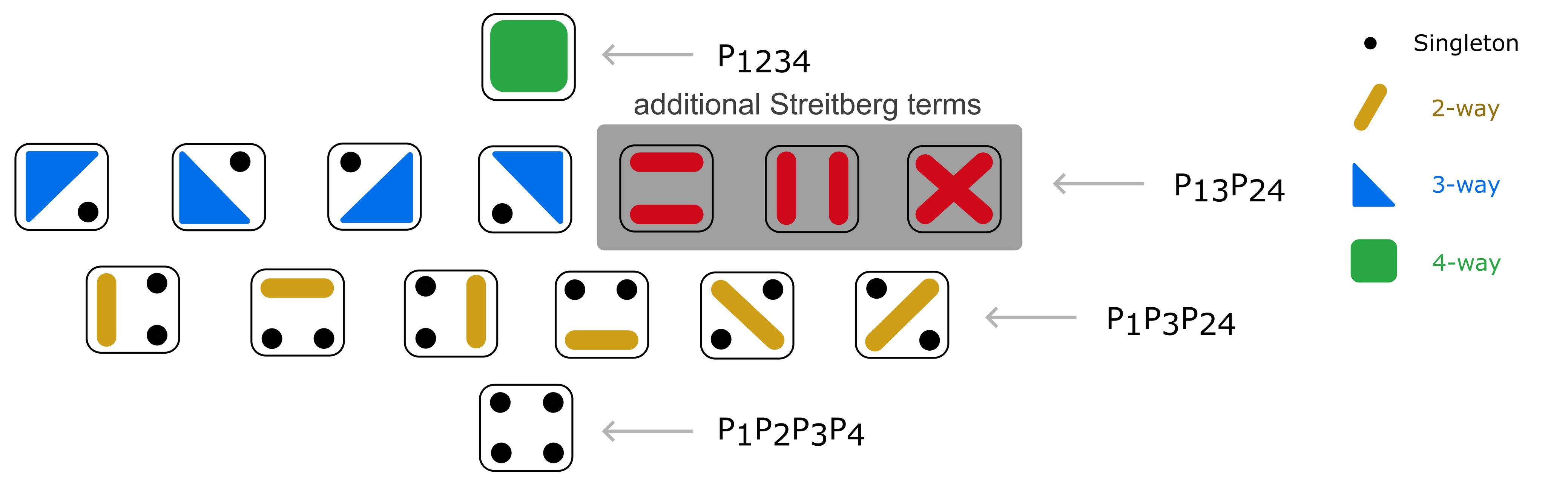
After removing the singletons, $B(d)$ is isomorphic to Lancaster lattice $L(d)$, a sublattice of partition lattice.



Partition Lattice - General

"All finite lattices can be embedded into a finite partition lattice."

This suggests an information measure induced by partition lattice can be more powerful.



Streitberg Information

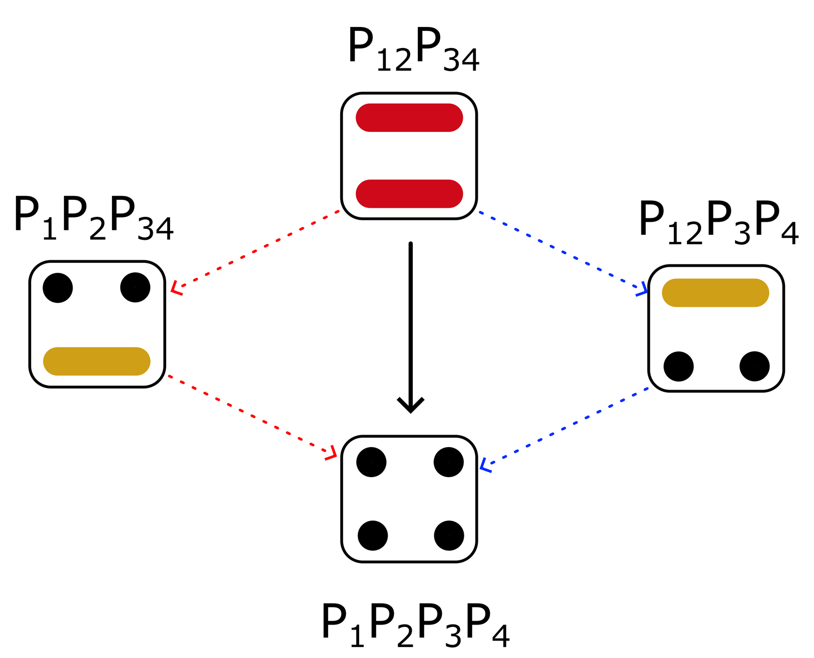
Streitberg interaction from partition lattice:

$$\Delta_S^d P = \sum_{\pi} (-1)^{|\pi|-1} (|\pi| - 1)!\, P_\pi$$

If the joint distribution factorises in any way, then $\Delta_S^d P$ becomes zero.

~~KL formulation~~

KL-divergence formulation is not effective as it destroys the structure of the lattice. E.g., $D_{\mathrm{KL}}(P_{12}P_{34} \| P_1P_2P_3P_4) = D_{\mathrm{KL}}(P_{34} \| P_3P_4) + D_{\mathrm{KL}}(P_{12} \| P_1P_2)$.



Tsallis Alpha Divergence

The Streitberg information $\mathrm{SI}(d)$ is defined as:

$$\mathrm{SI}(d) = \sum_{\pi} (|\pi| - 1)! (-1)^{|\pi|-1} D_{\mathrm{TA}} \left(p_\pi \Big\| \prod_{i=1}^{d} p_i \right)$$

Properties

Through experiments we validate:

1. Clear vanishing condition
2. Recursiveness
3. Monotonicity
4. Symmetry
5. Interpretability
6. Emergence

We provide a consistent estimator using kNN.

Applications

Finance

Identify the high-order interactions between the stocks in SP500, which can lead to more **diverse investment portfolio**.

Neuroscience

Outperforms pairwise-based methods at **decoding neural signals** as high-order information is prevalent in brain signals.

Machine Learning

Uncover the **true features** that generate the target variable through high-order interaction whereas standard feature selection methods like SHAP all fail.