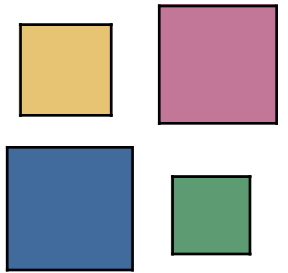


Grand Tour & Projection Pursuit

- **Problem bei PCP:**
Auch wenn alle Variablen gleichzeitig dargestellt werden, sind einige Features in den Daten nicht zu erkennen.
- **Problem bei Scatterplot:**
Orthogonale Achsen “enden” bei 3 Dimensionen
- **Lösung:**
Projektion von k -dimensionalen Daten auf 2-d möglich.
Während bei $k = 3$ nur eine “virtuelle” Dimension existiert, erhöht sich die Anzahl der “virtuellen” Dimensionen für $k > 3$.
- **Neues Problem:**
Für $k = 3$ sind die Rotationsachsen leicht zu erkennen.
Für $k > 3$ existieren diese Achsen nicht mehr
- **Neue Lösung: Grand Tour**



Grand Tour:

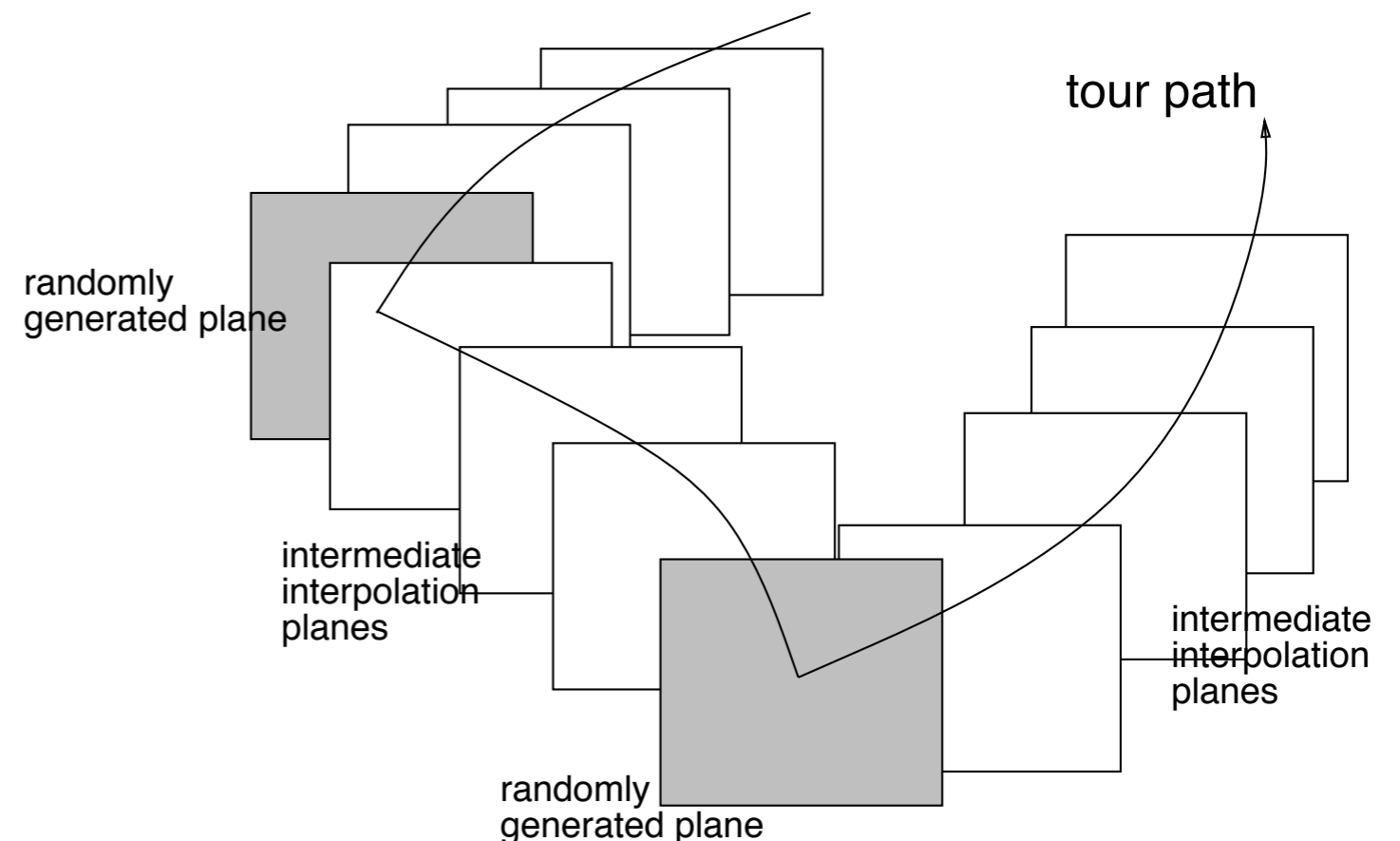
- **Definition** (Asimov 1985):

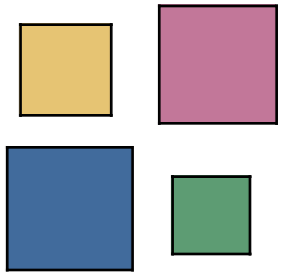
Eine *Grand Tour* ist definiert als eine stetige, einparametrische Familie von d -dimensionalen Projektionsebenen die dicht in der Menge aller d -dimensionalen Ebenen im p -dimensionalen Raum sind ($d < p$). Als Parameter wird meist die Zeit verwendet.

- **Implementierung:**

Der Pfad der Grand Tour interpoliert zwischen zufällig ausgewählten d -dimensionalen Ebenen.

⇒ Pseudorotation



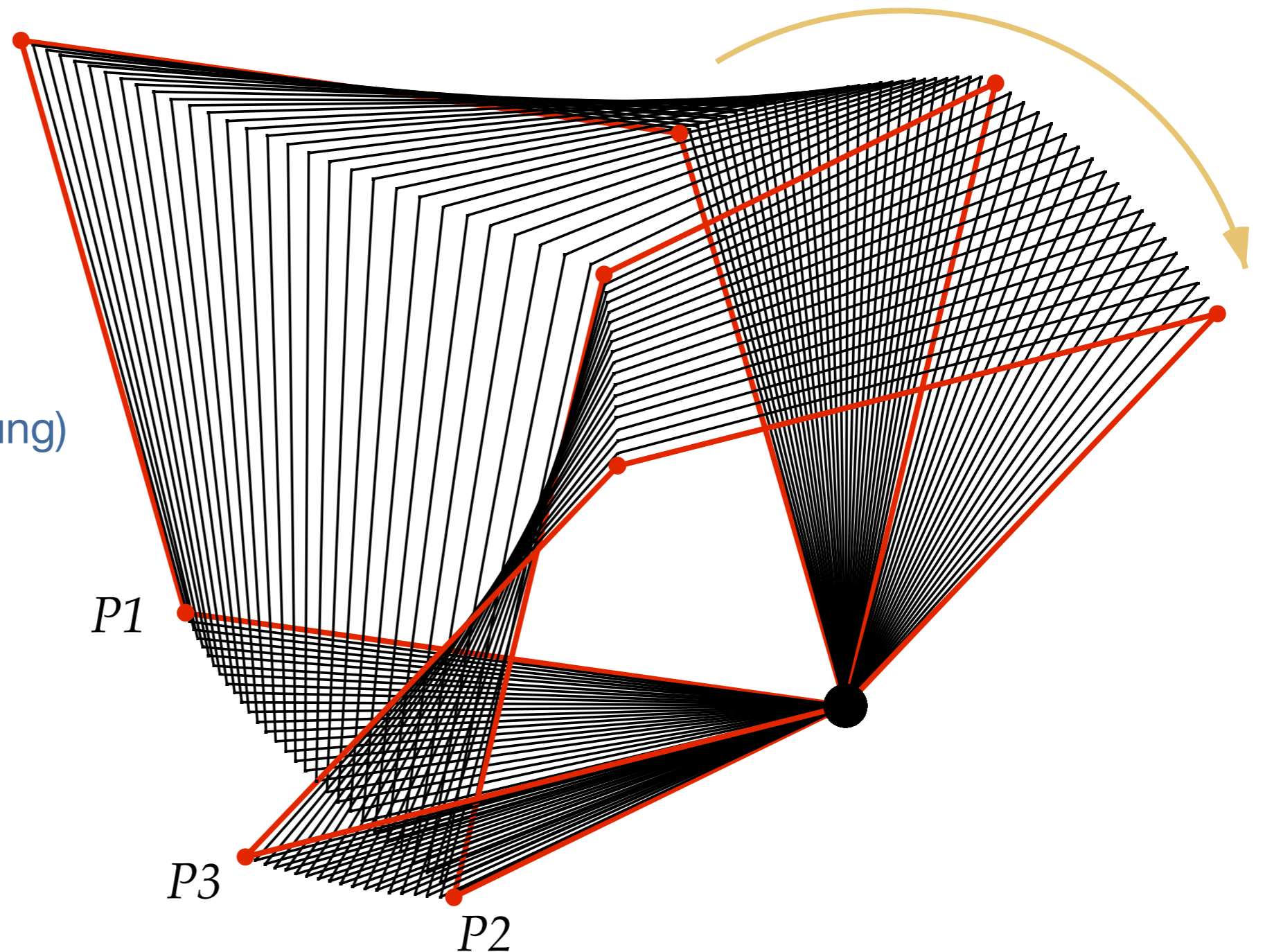


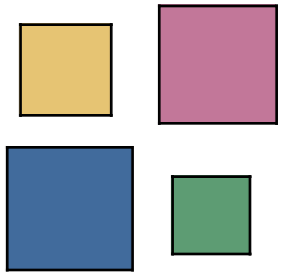
Grand Tour

- Beispiel

3 Frames mit
Interpolationen, je
definiert durch drei
Punkte.

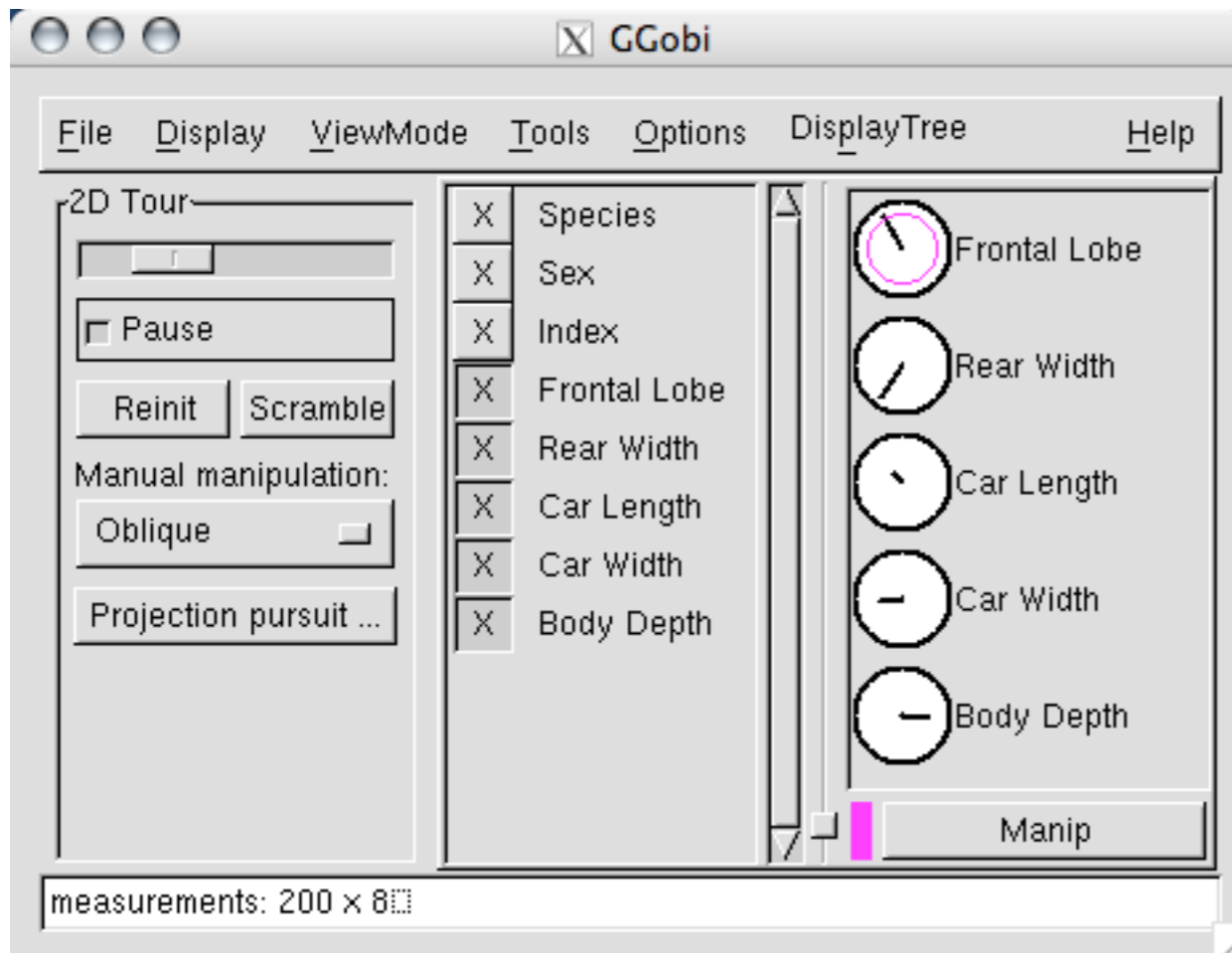
(Dicker Punkt ent-
spricht dem Ursprung)





Implementierung in ggobi

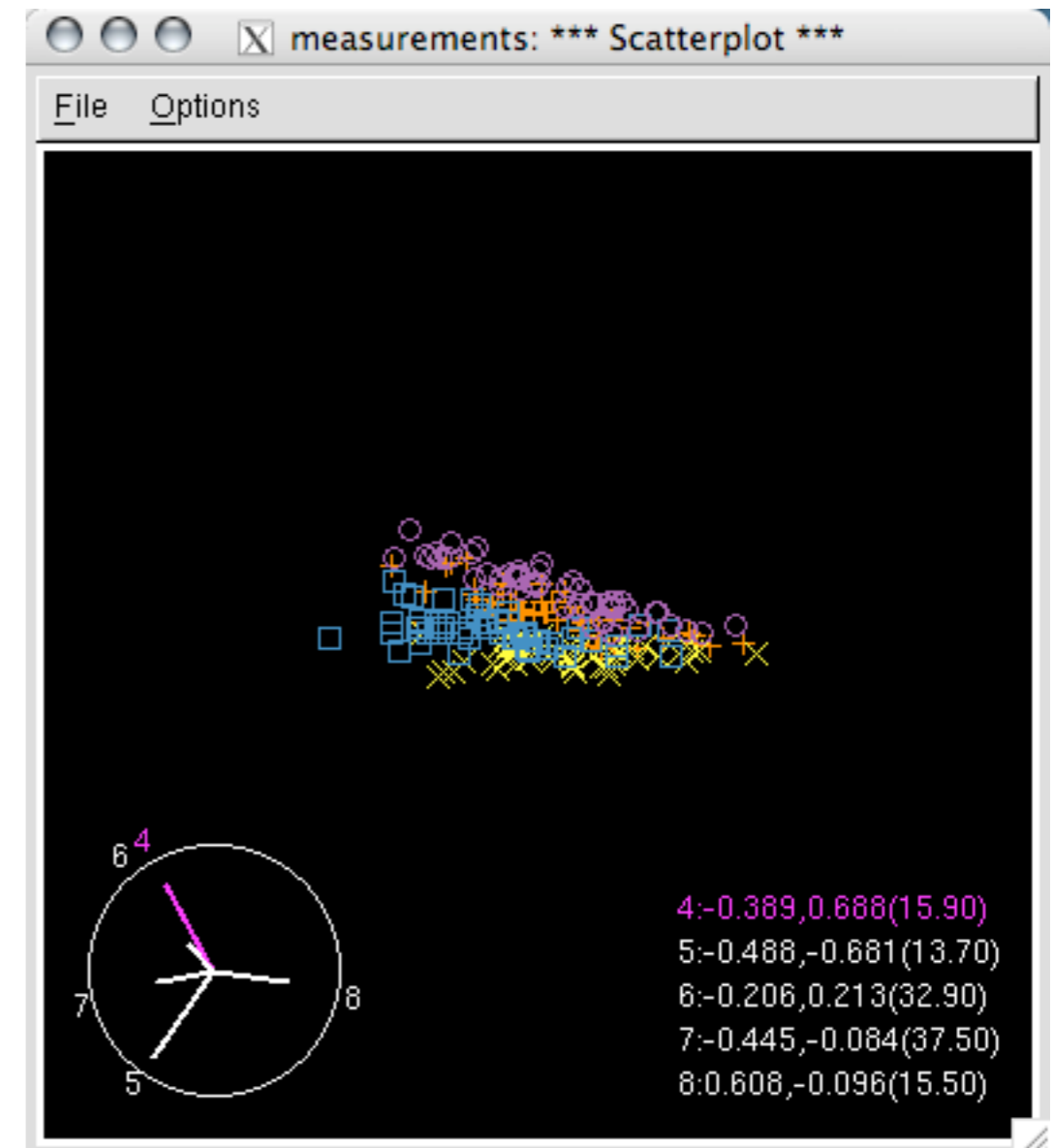
Variablen- und Kontroll-Fenster:

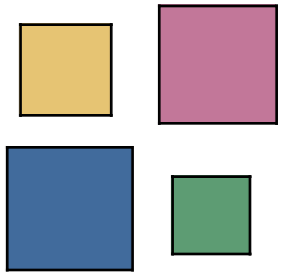


Manuelle "Rotation":

- frei
- entlang x oder y
- radial
- kreisförmig

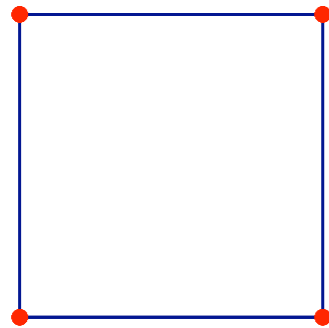
Tour Fenster



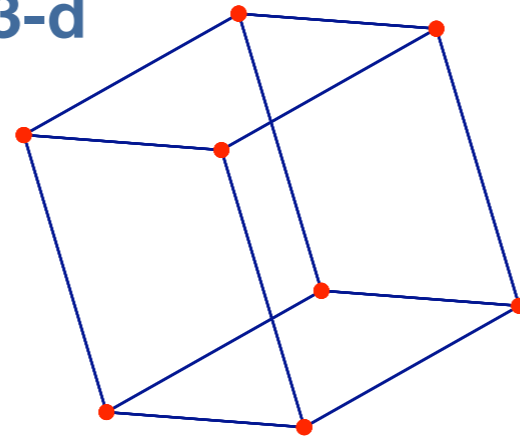


Beispiel: 6-Würfel

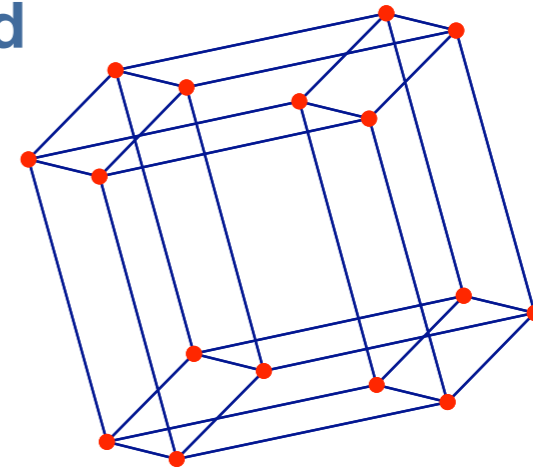
2-d



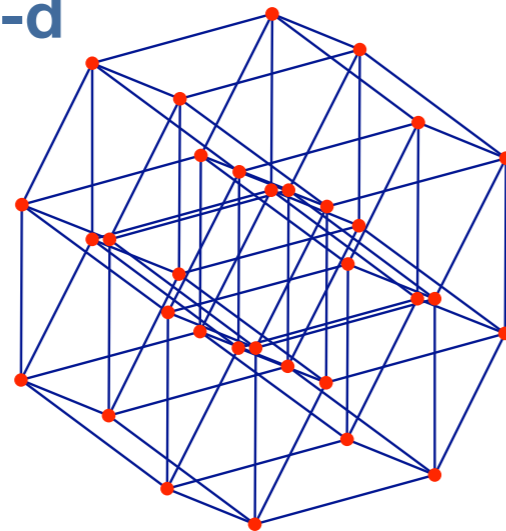
3-d



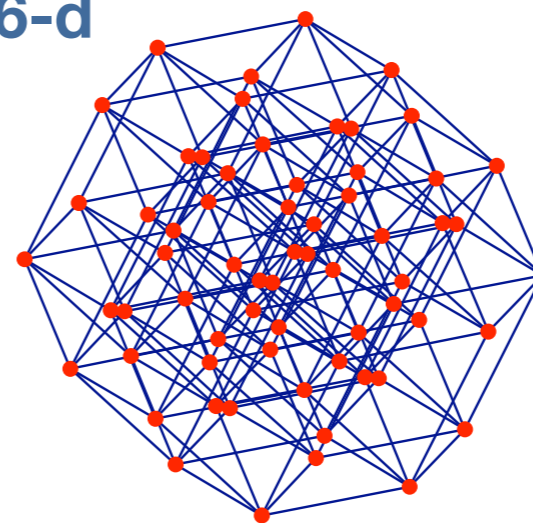
4-d

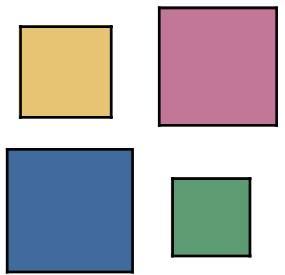


5-d



6-d

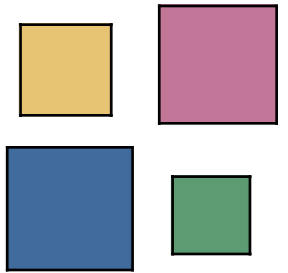




Case Study: Olive Oils

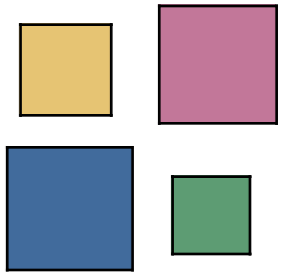
- Datenbeschreibung:
 - 10 Variablen
 - Region
 - ◆ North
 - ◆ South
 - ◆ Sardinia
 - Area (Hierarchische Verfeinerung)
 - ◆ North
 - East Liguria
 - West Liguria
 - Umbria
 - ◆ South
 - Calabria
 - North Apulia
 - South Apulia
 - Sicily
 - ◆ Sardinia
 - Coast Sardinia
 - Inland Sardinia
 - 8 prozentuale Anteile von Fettsäuren





Projection Pursuit

- **Grand Tour** “sucht” rein zufällig, d.h. ungerichtet, nach Features in den Daten.
- **Projection Pursuit** versucht bei der “Suche” ein bestimmtes Kriterium (Projection Pursuit Index, PPI) zu optimieren.
- Beispiele PPIs:
 - unsupervised
 - Holes
 - Central Mass
 - ...
 - supervised
 - LDA
 - Gini
 - Entropy
 - ...



Projection Pursuit

- Implementierung in ggobi

The screenshot displays the GGobi software interface. The main window, titled "olive oil: *** Scatterplot ***", shows a scatterplot of data points colored by a gradient from yellow to red. A "2D Tour Fenster" (2D Tour Window) is visible on the left, containing controls for "Pause", "Reinit", "Scramble", "Manual manipulation" (with "Oblique" checked), and "Projection pursuit ...". The status bar at the bottom left indicates "olive oil: 572 x 10".

A "Historie des PP Index" (History of the PP Index) window is open in the bottom right, showing a plot of the index value over iterations. The plot shows a curve that starts at 0.8, dips slightly, and then rises to a peak of 0.842 before ending at 0.9. The window also contains optimization settings: "Optimize" (unchecked), "Temp start: 1.00", "Cooling: 0.99", and "PP index: (0.8) 0.842 (0.9)". A "Holes" checkbox is also present.

At the bottom left of the scatterplot, a circular plot shows the current projection directions, with a pink line indicating the current direction. The following table lists the projection directions for dimensions 4 through 8:

Dimension	Value
4	-0.627
5	0.115
6	-0.308
7	-0.219
8	-0.671